Conic Sections in Cartesian and Polar Coordinates

The conic sections are a family of curves in the plane which have the property in common that they represent all of the possible intersections of a plane with a cone. This is related to the fact that they can all be represented by the same general equation

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0.$$ \[A\]

(In this discussion, we will set $B = 0$ to eliminate the so-called “mixed term”, which describes the way in which the axes of the curve are rotated with respect to the $x$- and $y$-coordinate axes. Here, the axes of the curves we survey will always be aligned “horizontally” or “vertically”.)

Take a straight line $D$ (called the “directrix”) and choose a point $P$ not on that line. We can form a curve for which each point at a distance $r$ from the point $P$ also lies at a perpendicular distance $kr$ from the line $D$, with $k$ being a positive constant. The shape of the curve depends on the value of $k$; the conic sections are the curves which have this property. We will consider the three possible cases in turn.

Case I -- $k = 1$

We’ll start by placing the directrix $D$ at $x = -d$ and the point $F$ not on this line at $(d, 0)$, so that the curve will pass through the origin $(0, 0)$. For this case, the perpendicular distance from a point on the curve to the line $D$ is the same as the distance from that point to the point $F$. This information will permit us to work out an expression to describe the curve. The distance from the point on the curve to the point $F$ is $r = \sqrt{(x-d)^2 + (y-0)^2}$, while the perpendicular distance from that point to the directrix is $r = \sqrt{(x-[-d])^2 + (y-y)^2}$. Consequently, we have

$$(x-d)^2 + y^2 = (x+d)^2 \quad \Rightarrow \quad x^2 - 2dx + d^2 + y^2 = x^2 + 2dx + d^2$$

$$\Rightarrow \quad y^2 = 4dx.$$
which is the form in Cartesian coordinates in which the equation of a (horizontal) parabola is typically presented. The point $F$ is the focus of the parabola and the distance we have called $d$ here is the focal distance. The point at the origin, labeled $V$, which is the closest point on the curve to the focus and to the directrix, is the vertex of the parabola.

![Diagram of a parabola with focus $F$, vertex $V$, and directrix $D$.](image)

We would now like to write an equation for this parabola using polar coordinates. While we could use the transformation equations $x = r \cos \theta$, $y = r \sin \theta$ to accomplish this, the result does not give the clearest expression for the conic section. So instead, we will shift the parabola horizontally so that the focus is at the origin (see graph below). We will then apply the usual definition for the location of a point on a polar curve to determine a description of the parabola in polar coordinates.

In this system, the distance of a point on the parabola from the focus at the origin is just $r$. This is the same as the perpendicular distance of that point from the directrix $D$. Since this point lies at a position $x = r \cos \theta$, and the directrix is $2d$ units to the left of the origin, the perpendicular distance is $r \cos \theta - [-2d] = r \cos \theta + 2d$. Hence,

$$r = r \cos \theta + 2d \quad \Rightarrow \quad r (1 - \cos \theta) = 2d \quad \Rightarrow \quad r = \frac{2d}{1 - \cos \theta} \quad .$$

Note that the denominator of this radius function can become zero at $\theta = 0$, which is consistent with the parabola being an open curve, here in the positive $x$-direction.
We return to working with Cartesian coordinates to develop a description of the conic sections for which the ratio of the perpendicular distance of a point on the curve to the directrix $D$ to the distance from the point to the focus $F$ is $kr : r = k$, taken to be a positive constant. Referring to the earlier diagram with the vertex of the curve at the origin, the directrix would now be found at $x = -kd$, making the perpendicular distance of interest $kr = \sqrt{(x - [-kd])^2 + (y - y)^2}$. This is $k$ times the distance between the point on the curve and the focus $F$, or $k \cdot r = k \sqrt{(x - d)^2 + (y - 0)^2}$.

We thus obtain the equation

$$k^2 [(x - d)^2 + y^2] = (x + kd)^2 \Rightarrow k^2 x^2 - 2k^2 dx + k^2 d^2 + k^2 y^2 = x^2 + 2kd x + k^2 d^2$$

$$\Rightarrow k^2 x^2 - x^2 - 2k^2 dx - 2kd x + k^2 y^2 = 0$$

$$\Rightarrow (k^2 - 1)x^2 - 2(k^2 + k)dx + k^2 y^2 = 0 .$$

(We see in this equation and the equation for the parabola, $y^2 - 4dx = 0$, the “family resemblance” with the general equation $A$ for conic sections.)

We can “complete the square” for the terms related to powers of $x$, permitting us to write this equation in the standard form
\[
(k^2 - 1)x^2 - 2(k^2 + k)dx + K + k^2y^2 = K.
\]

"completing the square"

\[
\Rightarrow (x - h)^2 + \frac{k^2y^2}{(k^2 - 1)} = \frac{K}{(k^2 - 1)} \Rightarrow \frac{(x - h)^2}{K/(k^2 - 1)} + \frac{y^2}{K/k^2} = 1.
\]

to writing the binomial-square and dividing through

by \(k^2 - 1\); we are not concerned with the exact value of \(h\)

By the method of our calculation, \(K > 0\), so the second term on the left-hand side of this last equation is always positive. If \(k > 1\), as for the case we are presently discussing, both terms are positive, so we have the equation for an ellipse with its center at \((h, 0)\). For \(0 < k < 1\), the first term on the left-hand side becomes negative and we then have the equation of a hyperbola, the branches of which open along the \(y\)-axis.

We will not be concerned here with refining this equation further to investigate its properties in detail. Rather, we will place the center of the ellipse at the origin next and use other properties of ellipses to develop an equation in Cartesian coordinates.

From the equation we've already obtained, the presence of the terms involving \(x^2\) and \(y^2\) tells us that ellipses have symmetry about both the \(x\) and \(y\)-axes (referred to as "four-fold symmetry" or "symmetry about the origin"). This will mean that ellipses have two vertices and two foci. An additional property of the ellipse that will be useful to us is that the sum of the distances from a point on the curve to each of the foci is a constant. Upon referring to the positions shown in the graph above, we can write
\[ r + r' = \sqrt{(x-[f])^2 + (y-0)^2} + \sqrt{(x-f)^2 + (y-0)^2} = C \]

\[ \Rightarrow \sqrt{(x+f)^2 + y^2} = C - \sqrt{(x-f)^2 + y^2} \]

\[ \Rightarrow (x^2 + 2fx + f^2) + y^2 = C^2 - 2C\sqrt{(x-f)^2 + y^2} + (x^2 - 2fx + f^2) + y^2 \]

squaring both sides

\[ \Rightarrow 2C\sqrt{(x-f)^2 + y^2} = C^2 - 4fx \]

\[ \Rightarrow 4C^2\left[(x^2 - 2fx + f^2) + y^2\right] = \left(C^2 - 4fx\right)^2 = C^4 - 8C^2fx + 16f^2x^2 \]

again squaring both sides

\[ \Rightarrow 4C^2x^2 + 4C^2f^2 + 4C^2y^2 = C^4 + 16f^2x^2 \]

\[ \Rightarrow (4C^2 - 16f^2)x^2 + 4C^2y^2 = C^4 - 4C^2f^2 \]

\[ \Rightarrow (1 - 4\frac{f^2}{C^2})x^2 + y^2 = \frac{C^2}{4} - f^2. \quad [C] \]

Before we can simplify this relation any further, we will need to know what the constant \( C \) represents. If we use the distance-sum equation at one of the vertices of the ellipse, say, \( V'(a, 0) \), we obtain

\[ r + r' = \sqrt{(a+f)^2 + 0^2} + \sqrt{(a-f)^2 + 0^2} = C \Rightarrow C = (a+f) + (a-f) = 2a. \]

If we then apply the same equation at one of the points on the \( y \)-axis, say, \( M(0, b) \), we also find

\[ r + r' = \sqrt{(0+f)^2 + b^2} + \sqrt{(0-f)^2 + b^2} = C \Rightarrow C = 2\sqrt{f^2 + b^2}. \]

This leads us to an important characteristic of ellipses, that \( a = \sqrt{f^2 + b^2} \)

\[ \Rightarrow f^2 = a^2 - b^2. \]

The equation under development thus becomes

\[ (\frac{C^2 - 4f^2}{C^2})x^2 + y^2 = \frac{C^2}{4} - f^2 \]

\[ \Rightarrow (\frac{[2a]^2 - 4[a^2-b^2]}{[2a]^2})x^2 + y^2 = \frac{(2a)^2}{4} - (a^2-b^2) \]

\[ \Rightarrow (\frac{4b^2}{4a^2})x^2 + y^2 = b^2 \Rightarrow \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad [D] \]

which is the familiar equation for the ellipse.
We will follow the approach we used for the parabola to determine the equation of the ellipse in polar coordinates by placing one focus at the origin; we will then use the distance-sum equation for the ellipse and some trigonometry to analyze the arrangement.

Here is one approach to developing the polar curve for the ellipse. We have already worked out that \( r + r' = 2a \), so we have \( (r')^2 = (2a - r)^2 \). We can construct a right triangle \( \Delta FPQ \) by dropping a perpendicular line from a point \( P \) on the ellipse down to the \( x \)-axis to mark a point \( Q \). The point \( P \) at \((x, y)\) has the position \((r \cos \theta, r \sin \theta)\) in polar coordinates, so the point \( Q \) lies at \((r \cos \theta, 0)\).

Since \( \Delta FPQ \) is a right triangle, we can use the Pythagorean Theorem to find

\[
(r')^2 = (r \sin \theta)^2 + (2f - r \cos \theta)^2 \Rightarrow (r \sin \theta)^2 + (2f - r \cos \theta)^2 = (2a - r)^2
\]

\[
\Rightarrow r^2 \sin^2 \theta + 4f^2 - 4fr \cos \theta + r^2 \cos^2 \theta = 4a^2 - 4ar + r^2
\]

\[
\Rightarrow r^2 (\sin^2 \theta + \cos^2 \theta) + 4f^2 - 4fr \cos \theta = 4a^2 - 4ar + r^2
\]

\[
\Rightarrow ar - fr \cos \theta = a^2 - f^2 \Rightarrow r (a - f \cos \theta) = a^2 - (a^2 - b^2)
\]

applying the Pythagorean Identity, rearranging terms, and dividing by 4

\[
\Rightarrow r = \frac{b^2}{a - f \cos \theta}.
\]
We have arrived at a polar equation for the ellipse. In preparation for what is yet to come, it will be convenient to define the quantity \( e = \frac{f}{a} \). Our polar equation then becomes
\[
r = \frac{b^2}{a - f \cos \theta} \cdot \frac{1}{a} = \frac{b^2}{a - e \cos \theta};
\]
since \( f^2 = a^2 - b^2 \), we have \( e^2 = \frac{f^2}{a^2} = \frac{a^2 - b^2}{a^2} = 1 - \frac{b^2}{a^2} \) and thus
\[
r = \frac{a \cdot (b^2/a^2)}{1 - e \cos \theta} = \frac{a(1 - e^2)}{1 - e \cos \theta}.
\]
Because \( a, b, \) and \( f \) must be positive lengths, the quantity \( e \) must be positive; it is possible, though, for \( a \) to equal \( b \) (in which case, the ellipse becomes a circle), so that \( f = 0 \), so \( e \) may also be zero. Since \( r \), the distance from the focus at the origin to a point on the ellipse, must also be positive, we require that \( 1 - e^2 > 0 \). No point on the ellipse is at the origin, so \( r \neq 0 \Rightarrow e \neq 1 \). For an ellipse, then, \( 0 \leq e < 1 \).

**Case III \(- 0 < k < 1\)**

We have already said earlier that this case produces a hyperbola. We will work, with a hyperbola whose branches open along the \( x \)-axis, in a manner analogous to what we did for the ellipse. We will again develop the equation for the curve in Cartesian coordinates, now for a hyperbola centered at the origin. In place of the “distance-sum relation” of the ellipse, for the hyperbola we have that the difference between the distances from each focus to a point on the curve is constant. Therefore,
\[ r - r' = \sqrt{(x-[-f])^2 + (y-0)^2} - \sqrt{(x-f)^2 + (y-0)^2} = C \]

\[ \Rightarrow \sqrt{(x+f)^2 + y^2} = C + \sqrt{(x-f)^2 + y^2} \]

\[ \Rightarrow 2C \sqrt{(x-f)^2 + y^2} = 4fx - C^2 \Rightarrow (1 - 4 \frac{f^2}{C^2})x^2 + y^2 = \frac{C^2}{4} - f^2 \]

The two times we must square the equation in order to solve it leads to a result which appears to be just the same as we have in Equation \( C \) for the ellipse. What is more, taking the distances from the focus \( F(-f,0) \) to the vertex \( V(a,0) \) and from the focus \( F'(f,0) \) to \( V \) yields \( r - r' = (a-[-f]) - (f-a) = 2a \), so our equation in Cartesian coordinates is again

\[ \frac{(C^2 - 4f^2)}{C^2} x^2 + y^2 = \frac{C^2}{4} - f^2 \Rightarrow \frac{4a^2 - 4f^2}{4a^2} x^2 + y^2 = \frac{4a^2}{4} - f^2 \]

\[ \Rightarrow \frac{(a^2 - f^2)}{a^2} x^2 + y^2 = a^2 - f^2 \]

Now, however, \( a < f \), so we will multiply the equation through by \(-1\) to keep all the factors in the equation positive, giving us

\[ \Rightarrow \frac{x^2}{a^2/(f^2 - a^2)} - \frac{y^2}{f^2 - a^2} = \frac{x^2}{a^2} - \frac{y^2}{(f^2 - a^2)} = 1 \]

\[ [F] \]

the equation for a hyperbola with branches that open along the \( x \)-axis. (There is no property of the curve itself which provides us with information about the value of \( b \), so we will have to show later what the positive factor \( f^2 - a^2 \) equals.)
We will also repeat our development of the polar equation for this curve in a way analogous to that for the ellipse. We place one focus of the hyperbola at the origin and label the significant points in the same fashion as before (shown in the graph above). The “distance-difference equation” gives us \( r - r' = 2a \Rightarrow (r')^2 = (r - 2a)^2 \). We will use the Law of Cosines on the triangle which has vertices at the point on the curve and the two foci \( F \) and \( F' \) to write

\[
(r')^2 = r^2 + (2f)^2 - 2 \cdot r \cdot 2f \cdot \cos \theta \Rightarrow r^2 + 4f^2 - 4fr \cos \theta = (r - 2a)^2
\]

\[
\Rightarrow r^2 + 4f^2 - 4fr \cos \theta = r^2 - 4ar + 4a^2
\]

\[
\Rightarrow r (a - f \cos \theta) = a^2 - f^2
\]

If we again use our quantity \( e = f / a \), this may be expressed as

\[
r = \frac{a^2 (1 - \frac{f^2}{a^2})}{a - f \cos \theta} = \frac{a (1 - e^2)}{1 - e \cos \theta},
\]

which gives us the same polar equation as for the ellipse (Equation E). Because \( f > a \), we find that \( e > 1 \) for a hyperbola and that there is no upper limit for this quantity. However, now it is possible for \( r \) to take on any real value, which makes the hyperbola an open curve. In this form of the equation, the numerator is always negative, so we have \( r > 0 \) for \( \cos \theta < 1/e \) and \( r < 0 \) for \( \cos \theta > 1/e \). Because cosine is an even function \( [\cos \theta = \cos (-\theta)] \), the hyperbola is symmetrical about the \( x \)-axis. The denominator in our polar equation can become zero at \( \cos \theta = 1/e \), so the radius becomes undefined along two directions, the so-called asymptotes of the hyperbola. We will be able to use these to resolve two of our standing questions about conic sections.
In the graph above, we have drawn the ellipse centered on the origin, \( \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \), and the hyperbola also centered on the origin, \( \frac{x^2}{a^2} - \frac{y^2}{b^2} = \frac{1}{f^2 - a^2} \), along with the vertices \( V \) of both the ellipse and hyperbola, the endpoints \( M \) of the vertical axis of the ellipse, the asymptotes of the hyperbola (the two diagonal lines through the origin), and the two lines radiating from the focus \( F \) of the hyperbola in the directions specified by \( \cos \theta = \frac{1}{e} \) (where the radius for the hyperbola is undefined).

We see that the asymptotes passes through the corners of a box marked by the points \( A \), which lie at \((a, b), (a, -b), (-a, -b), \) and \((-a, b)\), thereby neatly enclosing the ellipse. The slopes of the two asymptote lines are thus \( \frac{b}{a} \) and \(-\frac{b}{a}\). (We can confirm this by solving the equation of the ellipse for \( y \) and taking the limit \( \lim_{x \to \pm \infty} \pm b \sqrt{1 - \frac{x^2}{a^2}} \).)

But we also see that the directions in which the hyperbola is undefined are parallel to these asymptotic lines. They lie at angles \( \theta \) to the positive \( x \)-axis given by \( \cos \theta = \frac{1}{e} \). From the relations among the trigonometric functions, this gives

\[
\tan \theta = \pm \sqrt{\sec^2 \theta - 1} = \pm \sqrt{\frac{1}{\cos^2 \theta} - 1} = \pm \sqrt{e^2 - 1} \quad \text{or} \quad \pm \sqrt{\left(\frac{f}{a}\right)^2 - 1} = \pm \sqrt{\frac{f^2 - a^2}{a}}.
\]

These are the slopes of the lines extending away from the focus of the hyperbola, on which the curve is undefined and which are parallel to the asymptotes. Hence,

\[
\pm \sqrt{\frac{f^2 - a^2}{a}} = \pm \frac{b}{a} \quad \Rightarrow \quad f^2 - a^2 = b^2 \quad \text{or} \quad a^2 + b^2 = f^2
\]

for the hyperbola. Referring back to Equation \( F \), we now see that the hyperbola does have a Cartesian equation, \( \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \), analogous to that for the ellipse.

We also find that both the ellipse and the hyperbola can be described by the same polar equation, \( r = \frac{a (1 - e^2)}{1 - e \cos \theta} \), with \( 0 \leq e < 1 \) for the ellipse and \( e > 1 \) for the hyperbola. What of the parabola, for which we found \( r = \frac{2d}{1 - \cos \theta} \) (equation \( B \))? This suggests that \( e = 1 \), but the ellipse-hyperbola equation we have developed would be useless in this case.

One approach that can be taken is to note the value of the radius for the curves at \( \theta = \frac{\pi}{2} \), which is \( |r| = |a (1 - e^2)| \) for the ellipse or hyperbola and \( r = 2d \) for the parabola. In a conic section, this line segment running from the focus to the curve in the directions \( \theta = \pm \frac{\pi}{2} \) is referred to as the semi-latus rectum, \( p \). We can then write a single polar equation for all three conic sections,
\[ r = \frac{p}{1 - e \cos \theta}, \quad \text{where} \quad p = \begin{cases} a (1 - e^2) & 0 \leq e < 1 \\ 2d & e = 1 \\ a (e^2 - 1) & e > 1 \end{cases}. \]

This unification of the conic sections under one expression simplifies the proof that an object, subject to a force which varies with the inverse square of the distance from the agent of that force (such as gravity or the electric force of attraction or repulsion), will follow a path which is one of the conic sections. (While that proof is a bit sophisticated, it is completely within the means of the methods we’ve learned in Calculus II.)

-- G. Ruffa
6 – 14 July 2010