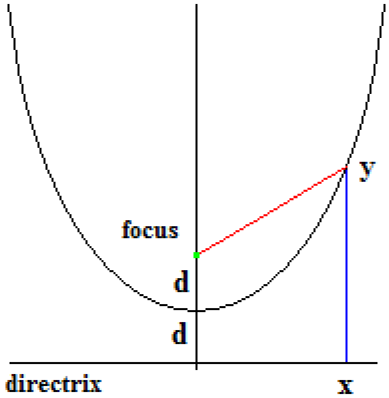
Conic Sections

A parabola is a curve defined so that every point along the curve is the same distance (R) from the focus and a line beneath the parabola called the directrix.



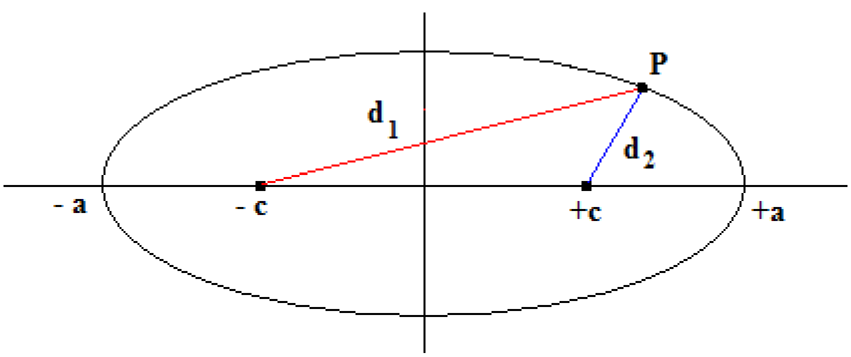
The lowest point of the curve, vertically beneath the focus, is the vertex of the parabola. For simplicity, we can set the vertex to have coordinates of (0,0). If the focus has coordinates of (0, d), then the directrix is a line with the equation, y = -d. We can see from the diagram that the blue line is equal in length to d + y. The red line is the hypotenuse of a triangle where the base has a length x and a height (y – d). The red line length is then $\sqrt{x^2 + (y-d)^2}$. If we set the red line length equal to the blue line length, as is true by definition for a parabola, we have:

$y + d = \sqrt{x^2 + (y-d)^2}$ which reduces to $x^2 = 4yd$

If we define $a = \frac{1}{4d}$ we have a simple equation for a parabola $y = a \cdot x^2$

when the parabola is aligned with the y-axis.

An ellipse is a curve defined so that every point along the curve is a distance from two points (called the foci of the ellipse) and the sum of the two distances is constant.



If we set the center of the ellipse at (0,0) between the two foci, then the foci will have coordinates of (-c, 0) and (+c, 0) and the vertices along the major axis will have coordinates

(-a, 0) and (+a, 0). We can see that when P is at the right vertex, $d_1 = c + a$ and $d_2 = a – c$. Therefore, $d_1 + d_2$ will always equal 2a.

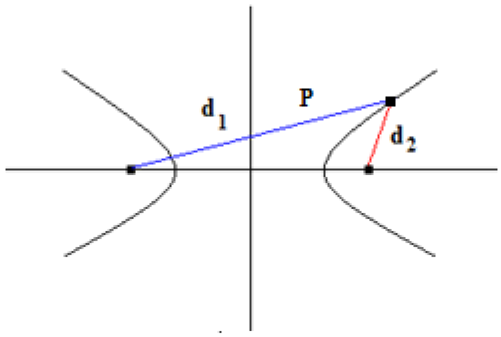
Take point P in the diagram above at coordinates (x, y) and we see $d_1 = \sqrt{(c + x)^2 + y^2}$ and $d_2 = \sqrt{(x - c)^2 + y^2}$. If $d_1 + d_2 = 2a$, then we have:

$\sqrt{(c + x)^2 + y^2} + \sqrt{(x - c)^2 + y^2} = 2a$ which can be simplified to

$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ where $b^2 = a^2 – c^2$ and b is the y-coordinate when x = 0 and

a > c, the vertices are outside the foci when the ellipse is aligned with the x-axis.

Last of the three, a hyperbola is a curve defined so that every point along the curve is a distance from two points (the foci) and the difference between the two distances is constant.



If we center the hyperbola at the origin, the foci are again at (-c, 0) and (+c, 0). The diagram above shows $d_1 = \sqrt{(c + x)^2 + y^2}$ and $d_2 = \sqrt{(x - c)^2 + y^2}$

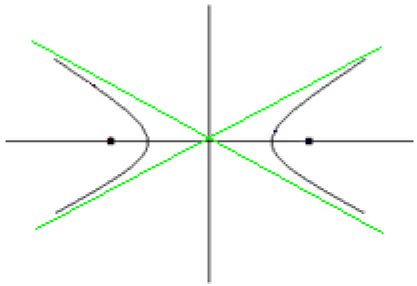
$\sqrt{(c + x)^2 + y^2} + \sqrt{(x - c)^2 + y^2} = 2a$ which can be simplified to

$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ where $b^2 = a^2 – c^2$ and b is the y-coordinate when x = 0

This looks like the same equation as the ellipse, but we write it as

$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ where $b^2 = c^2 – a^2$ because c > a, the foci are outside the vertices

and the hyperbola is aligned with the x-axis.



And if we rearrange $\frac{x^2}{a^2}$ - $\frac{y^2}{b^2} = 1$ for y, we get y = $\pm\frac{b}{a}\cdot x\cdot\sqrt{1-\frac{a^2}{x^2}}$

As x approaches infinity, $\frac{y}{x} = \pm\frac{b}{a}$ which are the slant asymptotes

---

Having written the conic section curves in Cartesian coordinates, we will now see all three forms can be written with a single equation in polar coordinates.

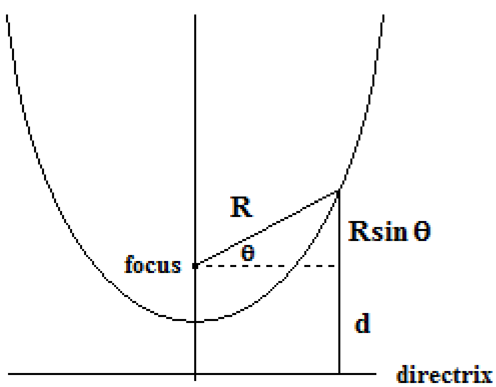
That equation is as follows:

$$R = \frac{e\cdot d}{1 \pm e\cdot sin\theta}$$

where d is the distance between the focus and the directrix and *e* is the eccentricity of the curve, which can be greater than, less than, or equal to one. If the directrix is right or above the focus, use the positive sign in the denominator, if left or below, use the negative sign. And if the conic section is aligned with the x-axis, replace sin with cos.

Let's take the case where *e* = 1.

$R = \frac{d}{1-sin\theta}$ d = R – Rsinθ or d + Rsinθ = R



As you see in the diagram above, that is exactly the condition we need for a parabola: the distance between any point on the curve and the focus (R) is equal to the distance between that same point and the directrix (d + Rsinθ).

We could have chosen the difference in the denominator; that is simply the same parabola with the directrix above the focus. Also, recall what the R and θ represent in polar coordinates. R indicates the distance of any point along the curve from the origin and θ is the corresponding angle as you travel from the origin to that point.

For an ellipse, we can again begin with $R = \frac{e \cdot d}{1 - e \cdot sin\theta}$ or R = e(d + sinθ)

Squaring both sides and converting to Cartesian coordinates yields:

$x^2 + y^2 = e^2(d + y)^2$ and then a fair amount of algebra converts this into:

$$(y - \frac{e^2 d}{1 - e^2})^2 + \frac{x^2}{1 - e^2} = \frac{e^2 d^2}{(1 - e^2)^2}$$

When $e < 1$, which is true for an ellipse, the second term is positive and we have the Cartesian form of an ellipse shifted upward a distance h:

$$\frac{(y-h)^2}{b^2} + \frac{x^2}{a^2} = 1$$

Dividing the equation

$(y - \frac{e^2 d}{1- e^2})^2 + \frac{x^2}{1- e^2} = \frac{e^2 d^2}{(1- e^2)^2}$ by $\frac{e^2 d^2}{(1- e^2)^2}$ shows

$h = \frac{e^2 d}{1- e^2}$ $a^2 = \frac{e^2 d^2}{1- e^2}$ $b^2 = \frac{e^2 d^2}{(1- e^2)^2}$

We can also use the equation for an ellipse $a^2 = b^2 - c^2$ and solve $c = \frac{e^2 d}{1- e^2}$

This shows the eccentricity $e = \frac{c}{a} = \frac{\sqrt{b^2 - a^2}}{a}$ and gives it some physical meaning. *c* is the distance from the center of the ellipse to a focus. *a* is the distance from the center of the ellipse to the farther side (a distance called the semi-major axis).

When $a = b$, the eccentricity is zero and the ellipse is a circle. *c* is zero because the lone focus is at the center of the circle. As *a* becomes smaller and smaller relative to *c*, the ellipse becomes narrower and "more eccentric".

We can also rearrange $a^2 = \frac{e^2 d^2}{1+e^2}$ into ed = $a(1 - e^2)$ so that the equation for an ellipse is:

$$R = \frac{a(1-e^2)}{1+e \cdot sin\theta}$$

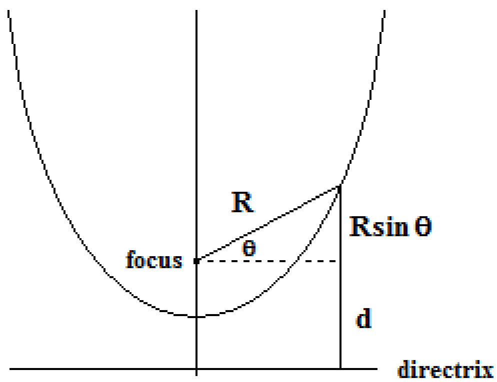
Going back to an intermediate step:

$$(y - \frac{e^2 d}{1-e^2})^2 + \frac{x^2}{1-e^2} = \frac{e^2 d^2}{(1-e^2)^2}$$

For a hyperbola, *e* > 1, which means the second term above is negative, and the equation takes the correct form of:

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

To give $R = \frac{e \cdot d}{1 \pm e \cdot sin\theta}$ a little more meaning, let's look at it applied to a parabola.



The directrix is below the focus, so we'll use the difference in the denominator. For a parabola, e = 1, so we can simplify the equation to $R = \frac{d}{1-sin\theta}$, noting that we are taking the focus as the origin.

When θ = 0, R = d and we travel from the focus to the right and plot a point at a distance d. When $\theta = \frac{\pi}{2}$, R = ∞ as it should be, because R increases as θ goes from 0 to $\frac{\pi}{2}$.

Going from $\frac{\pi}{2}$ to π, the sin function is symmetric to what it is from 0 to $\frac{\pi}{2}$ and so the left side of the parabola above the focus is similar to the right side.

From π to $\frac{3\pi}{2}$ the sin function is negative and increasing, giving us a larger and larger denominator, making the R value smaller and smaller. At $\frac{3\pi}{2}$, R is the shortest it will be, at $\frac{d}{2}$, which is should be for a parabola, where the vertex is directly between the focus and directrix.

Then again, from $\frac{3\pi}{2}$ to 2π, the symmetry of the sin function produces the symmetry of the parabola.

For an ellipse, $e < 1$ and that prevents the denominator of 1 – 1 = 0 when θ = $\frac{\pi}{2}$ and avoids the R value of infinity.

For a hyperbola, $e > 1$ and the infinite R arises even before θ = $\frac{\pi}{2}$.