The Conic Sections

C. Godsalve
email:seagods@hotmail.com

March 28, 2011

Contents

1 Introduction

The conic sections have been studied since the ancient Greek times, and are still important in physics today, as well as being quite pretty enough for them to be studied for themselves. What are they? They are the curves you get when a plane intersects with a cone. You can make a cone out of a piece of paper, and sketch the line that you think a plane might intersect with a cone to get a rough idea. Part of the reason for me putting this up is that students hardly ever see conic sections studied as planes intersecting cones. They are usually studied as the locus of a point, and the cone is entirely dispensed with. Here we shall give the cone centre stage, and introduce the locus of a point aspect later

If you take a right angle triangle with a hypotenuse of length one, as shown in Fig.1. The point P has coordinates (x, y), so Pythagoras' theorem tells us

$$x^2 + y^2 = 1. \tag{1}$$

In fact, if we consider all possible angles θ for all right angle triangles with hypotenuse of length one, we see that this equation holds true for each and every one. That is, eqn.1 is the equation of a circle of radius one. In fact, it is easy to see that if we want a circle of *any* radius, r say, then the equation of a circle of radius r is

$$x^2 + y^2 = r^2. \qquad (2)$$

Now suppose we have a three dimensional coordinate system shown in Fig.2. There is now a z axis perpendicular to both x and y. We can plot surfaces as well as lines in this coordinate system. In this example we have a cone. It looks like two cones point to point, with the axis of rotational symmetry being the z axis. In fact, this double cone is *the* cone in mathematics. Also, we have drawn in a base and a top to help us try to visualise the cone, but our mathematical cone goes on forever. Now we see that the circle has an equation in two dimensions. So can we expect the cone to have an equation in three dimensions?

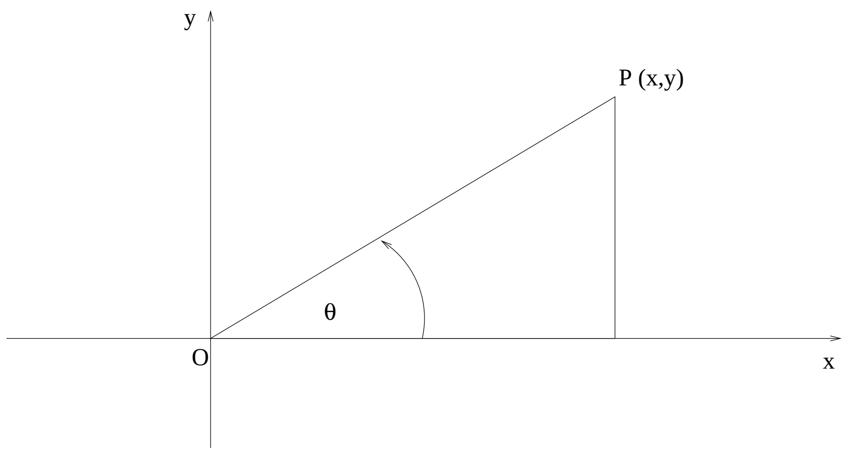

Figure 1: The point P gives us a right angle triangle with sides of length x and y. If the distance from the origin to P is one then $x^2 + y^2 = 1$. This is true for all points P having a distance of one from the origin.

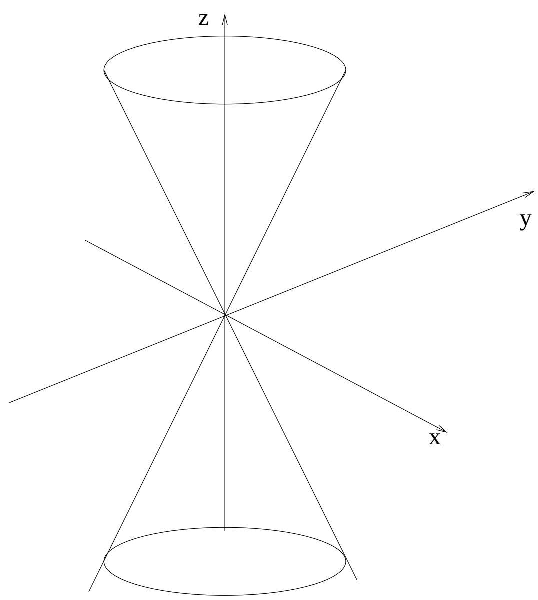

Figure 2: Here we see a the surface of a cone in three-dimensions. Every point is described by three numbers x, y, and z.

The answer is yes, and it is easy to find. If we look at the cone along the y axis, from a long way off, it will look like two straight lines at some angle to the z and x axis. This is the projection of the cone onto the (x, z) plane. As we move up the z axis from the origin, we can imagine a horizontal plane passing through our particular value of z. The plane will contain a circle where it cuts the cone. And the radius of this circle will be proportional to the z value. The radius of the circle is just the distance along the z axis times some number. We shall put $|z| = \beta r$. This is more convenient as it gives us the equations of the two lines that result when the cone is projected onto the (x, z) plane. We know the equation of the circle, so the equation of a cone is

$$\beta^2(x^2 + y^2) = z^2. \tag{3}$$

Now we want to study conic sections. A conic section is a curve where the cone intersects with a plane. We already knew the circle was a the conic section of an horizontal plane (constant z) and the cone to start with. This is because the cone has rotational symmetry around the axis by definition.

In order to simplify this we shall do two things. We shall shift the cone along the z axis by some distance α. This will simplify how we introduce the plane. Now, the general equation for a plane in three dimensional space is $z = ax + by + c$. We have done this shift so that the plane can pass through the origin, and still give is curves. (Without this a plane passing through the origin would just cut the cone at the single point where the two half cones meet.) Not only this, but the plane pass right through the whole of the y axis. This means the plane's equation is just $z = constant \times x$.

What is the equation of our shifted cone? It is

$$(z-\alpha)^2 = \beta^2(x^2+y^2). \tag{4}$$

To see this, when $z = \alpha$, the radius of the circle must be zero. Looking along the y axis, the plane cutting the cone looks something like Fig.3.

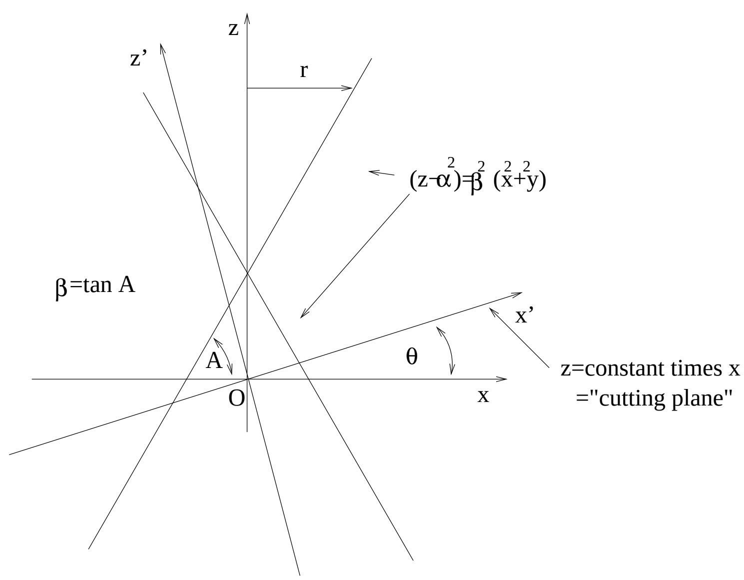

Figure 3: Here, a plane passing through the y axis cuts the cone $(z - \alpha^2)^2 = \beta^2(x^2 + y^2)$. New coordinates are to be found by rotating the x axis into the plane that cuts the cone.

Now though we have the plane $z = constant \times x$ we shall not be using it. What we are interested in is the angle θ, and the curve in the plane that cuts the cone. This is why it simplifies matters a lot if we choose a new x axis called the x' axis. This shall be perpendicular to the y axis. Not only that, but we shall use a new z' axis. This will be perpendicular to the x' axis and the y axis. Now the whole conic section will be contained in the (x', y) plane, and the problem will become two dimensional. It seems complicated, but it will simplify things a great deal.

So, how do we do the rotation? We could use matrices, which are covered a little bit in one of my "Math School" articles. But here we shall just write down the transformation to the new coordinates. Given any point (x, z) in the original coordinate system,

$$x' = \cos\theta x + \sin\theta = cx + sz,$$

$$z' = -\sin\theta x + \cos\theta = -sx + cz. \tag{5}$$

Here we have introduced $c = \cos\theta$, and $s = \sin\theta$ to simplify how the algebra looks. The other way round we have

$$x = cx' - sz',$$

$$z = sx' + cz'. \tag{6}$$

Now, lets put these values for x and z in the eqn.4. That is into the equation for the cone shifted along the z axis.

$$[(sx' + cz') - \alpha]^2 = \beta^2(cx' - sz')^2 + \beta^2 y^2. \tag{7}$$

Now, the whole curve lies in the plane $z' = 0$. This gives us the equation of the curve in the x', y plane as

$$[sx' - \alpha]^2 = \beta^2 c^2 x'^2 + \beta^2 y^2, \tag{8}$$

or after a little rearrangement, and dropping the prime on the x,

$$(\beta^2 c^2 - s^2)x^2 + 2s\alpha x + \beta^2 y^2 = \alpha^2. \tag{9}$$

Now we can begin to investigate the properties of this curve. It cuts the new x axis at y=0. If we apply the formula for solving the quadratic we get

$$x = \frac{-2s\alpha \pm \sqrt{4s^2\alpha^2 + 4\alpha^2(\beta^2c^2 - s^2)}}{2(\beta^2c^2 - s^2)}$$

after a little work

$$x = -\alpha\left(\frac{s \pm \beta c}{\beta^2 c^2 - s^2}\right). \tag{10}$$

(Remember the "c" in the usual quadratic formula is -α^2.)

Now remember that c and s are the cosine and sine of the angle at which the plane cuts the cone, so that we are not surprised that this angle turns out to important. If $\beta c = s$ the slope of the plane matches that of cone. To simplify matters a little we assume $\beta = \alpha = 1$. Then the cone angle is 45 degrees.

$$\cos 45 = \sin 45 = \frac{\sqrt{2}}{2}$$

so

$$c^2 - s^2 = \frac{1}{2} - \frac{1}{2} = 0. \tag{11}$$

This tells us that something very important happens close to 45 degrees, where the cone angle nearly matches the plane slope angle. We can see this in Fig.4. In the following x is *not* the x axis of the figure, but the x axis in the coordinates of the cutting plane passing through the origin. The meaning of the solution for x is that it is the point where the lines cross the 45 degree cone lines. First we examine the line AB, here $c > s$ since the slope of the line is less than 45 degrees. We shall call ϵ the difference between c and s, so that $c = s + \epsilon$. Then the two solutions for x are

$$x = \frac{\epsilon}{2s\epsilon + \epsilon^2}$$

$$x = \frac{-2s - \epsilon}{\epsilon + \epsilon^2}. \tag{12}$$

There is a positive solution, and a negative solution. (Remember that if the angle is less than 45 degrees $\sqrt{2}/2 < c < 1$ and $0 < s < \sqrt{2}/2$.) The solution must form a closed loop, which will be an ellipse (or a circle for $\theta = 0$).

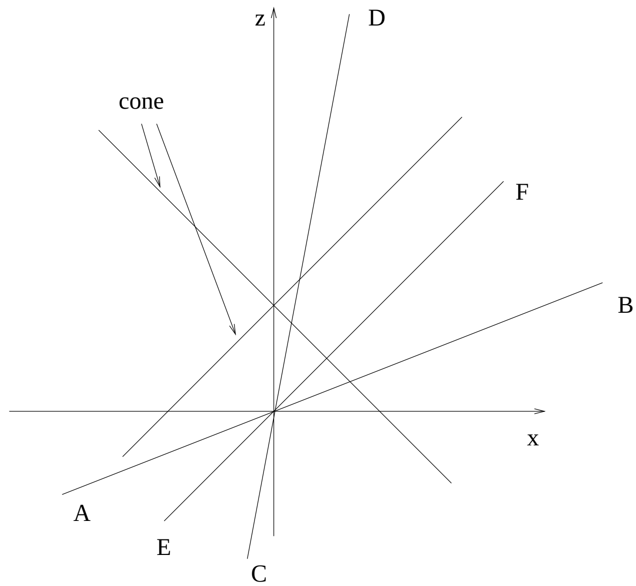

Figure 4: Here, a plane three planes cut the cone. One is inclined at less than 45 degrees (AB), one at more than 45 degrees (CD) and one at 45 degrees (45 degrees is the cone angle.

Now what happens when we make ϵ very small. A small number (much less than 1) when squared is far smaller than the original, which means as ϵ gets very small, we can ignore it altogether. Then

$$x \approx \frac{1}{2s}$$

$$x \approx -\frac{1}{\epsilon} - \frac{1}{2s}. \tag{13}$$

As ϵ becomes very small, the far end of the ellipse gets further and further from origin. The near end gets closer and closer to

$$x = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}. \tag{14}$$

So, when the angle is 45 degrees, there is only one solution where the plane cuts the cone.

Now what if $s = c + \epsilon$. Again ϵ is a positive small number. Then the plane angle is bigger than 45 degrees. If we substitute this into the solution for x. Then we get

$$x = \frac{-\epsilon}{-2c\epsilon - \epsilon^2} = \frac{\epsilon}{2c\epsilon + \epsilon^2}$$

$$x = \frac{-2c - \epsilon}{-2c\epsilon - \epsilon^2} . x = \frac{2c + \epsilon}{2c\epsilon + \epsilon^2}. \tag{15}$$

Now we see a switch in behaviour. Both the roots are to the right of the origin. For very small ϵ we have

$$x = \frac{1}{2c}$$

$$x = \frac{1}{\epsilon} + \frac{1}{2c}. \tag{16}$$

Now, as ϵ gets smaller and smaller, one solution gets larger and larger, and the other solution gets closer to $\sqrt{2}/2$ as before.

To sum up. If θ is less than 45 degrees (the cone angle) we see a closed loop called an ellipse which can roughly be described as an elongated circle. At 45 degrees, there is only one intersection, and the curve goes on forever. This curve is called a parabola. For θ greater than the cone angle we see that we have two curves, both unclosed and going on forever. This type of curve is called an hyperbola.

2 The Ellipse

We shall look at all these curves in some detail, but we shall pay most of our attention to the ellipse. Suppose $x1$ is the x value for the positive solution and $x2$ is the negative solution. The distance between the two points is the *major axis* of the ellipse, and the *semi-major axis* is half this value. If we denote it as a then it is easily seen that

$$a = \frac{\alpha\beta c}{\beta^2 c^2 - s^2}. \tag{17}$$

Now, the distance C from the origin to the centre of the semi major axis is easily found to be

$$c = \frac{\alpha\beta s}{\beta^2 c^2 - s^2}. \tag{18}$$

So if we make the substitution

$$x = x' - \frac{\alpha\beta s}{\beta^2 c^2 - s^2}, \tag{19}$$

we find the equation in the new coordinates is

$$(\beta^2 c^2 - s^2)x'^2 + \beta^2 y^2 = \frac{\alpha^2\beta^2 c^2}{\beta^2 c^2 - s^2}, \tag{20}$$

or

$$\left[\frac{(\beta^2c^2 - s^2)^2}{\alpha^2\beta^2c^2}\right]x'^2 + \left[\frac{\beta^2c^2 - s^2}{\alpha^2c^2}\right]y^2 = 1. \tag{21}$$

Clearly, we can write this as

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \tag{22}$$

where a is the semi major axis, and b is the semi minor axis. We have dropped the primes as we shall keep these coordinates. Clearly, we have $a > b$.

Suppose we have some point on the x axis that is within the ellipse. That is $x = \lambda a$ where $\lambda < 1$. What is the distance r from this point to the ellipse.

$$r^2 = (x - \lambda a)^2 + y^2 = (x - \lambda a)^2 + b^2\left(1 - \frac{x^2}{a^2}\right). \tag{23}$$

This is re-written

$$r^2 = \left(1 - \frac{b^2}{a^2}\right)x^2 - 2\lambda a x + b^2. \tag{24}$$

There is a particular value of λ for which this simplifies greatly. Choose

$$\lambda^2 = 1 - \frac{b^2}{a^2}, \tag{25}$$

then

$$r^2 = \lambda^2\left(x^2 - \frac{2ax}{\lambda} + \frac{a^2}{\lambda^2}\right)^2 \tag{26}$$

One root gives

$$\frac{r}{\lambda} = \frac{a}{\lambda} - x. \tag{27}$$

Now, this is something special. If P is on the ellipse, the distance from the point λa to P is proportional to the *horizontal* distance from P to a vertical line passing through the $x = a/\lambda$! This remarkable property at this particular value of λ gives it a special name, the *eccentricity*, denoted by e which we shall use from now on.

$$e = \sqrt{1 - \frac{b^2}{a^2}} \tag{28}$$

This means we can write b in terms of a and e. The point $(ae, 0)$ is called a *focus* of the ellipse, and there is another focus at $(-ae, 0)$. The vertical line at $x = a/e$ is called a *directrix* of the ellipse, again, there is another directrix at $x = -a/e$.

This strange property of the ellipse has an immediate consequence arising out of the symmetry of the ellipse. Now, from Fig.5 we have

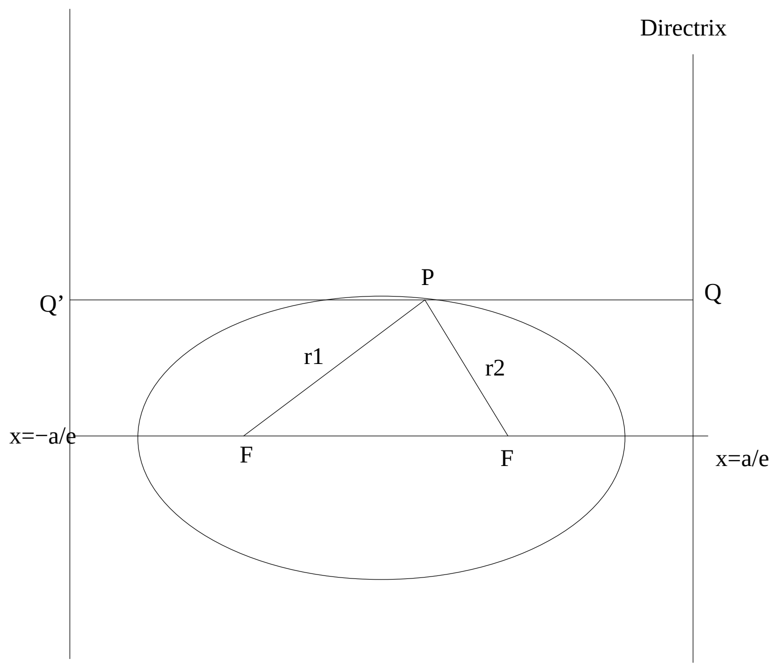

Figure 5: The distance from the left focus to P plus the distance from P to the right focus is a constant. Also, if we have a vertical line a distance a/e from the right focus, then the ratio of the distance from the right focus to P to the horizontal distance PQ is constant, and is equal to e. This line is called the directrix.

$$\frac{r1}{e} = PQ$$

and

$$\frac{r2}{e} = PQ'. \tag{29}$$

From this we see that

$$\frac{r1 + r2}{e} = \frac{2a}{e}. \tag{30}$$

That is the distance from one focus of the ellipse to any point P on the ellipse, added to the distance from P to the other focus is twice the semi-major axis.

Now we have seen that the ellipse is also the locus of a point P such that the distance from a point $(x = -c, y = 0)$ to P plus the distance from the point $(x = +c, y = 0)$ to P is constant. Imagine two pins stuck into the x axis, both a distance c from the origin, and stretching a string of length L between them. We might place a pen at the point where the string makes a corner, and if we move the pen so that the string is always stretched, then we shall trace out an ellipse. (This is called a Durer compass after the renaissance artist Albrecht Durer.) We shall proceed to re-derive the equation of the ellipse from this property, as if we had no idea that the ellipse was a section of a cone.

At some point P we have

$$L = r_1 + r_2 = \sqrt{(x-c)^2 + y^2} + \sqrt{(x+c)^2 + y^2}$$

or

$$\sqrt{(x+c)^2 + y^2} = L - \sqrt{(x-c)^2 + y^2}. \tag{31}$$

If we square both sides we arrive at

$$2L\sqrt{(x-c)^2 + y^2} = L^2 - 4xc$$

or

$$\sqrt{(x-c)^2 + y^2} = \frac{L}{2} - \frac{2xc}{L} \tag{32}$$

On squaring both sides it is found that

$$\frac{4x^2}{L^2} + \frac{4y^2}{L^2 - 4c^2} = 1. \tag{33}$$

Now one half of L is the semi major axis a so this equation can be written

$$\frac{x^2}{a^2} + \frac{y^2}{a^2 - c^2} = 1. \tag{34}$$

Denoting the semi-minor axis as b, we see that

$$c = a\sqrt{1 - \frac{b^2}{a^2}} = ae. \tag{35}$$

Picking the focus $x = c$, then the distances from c to the $y = 0$ points on the ellipse are $a(1-e)$ and $a(1+e)$.

Now, examine

$$r^2 = (x-c)^2 + y^2 = (x-ae)^2 + b^2\left(1 - \frac{x^2}{a^2}\right). \tag{36}$$

We use $b^2 = a^2 - c^2$ so

$$\frac{r^2}{e^2} = x^2 - 2\frac{a}{e}x + \frac{a^2}{e^2}. \tag{37}$$

Now, suppose we have a vertical line at some value of x. The distance from the point on the ellipse to this line is $A = D - x$, so

$$\frac{r^2}{A^2e^2} = \frac{x^2 - 2\frac{a}{e}x + \frac{a^2}{e^2}}{x^2 - 2Dx + D^2}. \tag{38}$$

Now if we put $D = a/e$, the right hand side is just one, i.e. for this particular line

$$\frac{r}{A} = e. \tag{39}$$

So, we have re-derived the focus-directrix property of the ellipse.

Now, how does the eccentricity relate to the original cone? Equations 21 ans 22 give us a and b in terms of the angles θ and A as defined in Fig.3. From this, we can use eqn.27 to find the eccentricity.

$$e^2 = 1 - \frac{a^2}{b^2} = \frac{s^2}{\beta^2}(1 + \beta^2). \tag{40}$$

Remembering that $s = \sin\ \theta$ and $\beta = tan\ A$, and that $1 + tan^2 A = 1/\cos^2 A$ we find that the eccentricity of the ellipse is

$$e = \frac{\sin\ \theta}{\sin\ A}. \tag{41}$$

It seems obvious that there must be some relationship between the focus-directrix definition of the ellipse (and other conic sections) and the intersection of the cone and the plane. However, this is by no means obvious.

There *really is* a direct relation between the cone and the cutting plane and the focci and the directrices of the ellipse. (From symmetry, there are two directrices, one for each focus.) Though the focus/directrix definitions of the conic sections were known before 300BC, it took until 1822 for these relations to be discovered by the French mathematician Dandelin. I shall give a brief description of this at the end.

3 The Ellipse for Astronomy

Kepler discovered that the planets move in elliptical orbits around the sun, and that the sun was at a focus of an ellipse. In astronomy, the best coordinates to use to describe this motion are polar coordinates centred at the focus.

Translating coordinates to the focus we have

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

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

$$(1 - e^2)(x^2 + 2aex^2e^2) + y^2 = a^2(1 - e^2). \tag{42}$$

Now, we put $x = r\ \cos\ \theta$ and $y = r\ \sin\ \theta$, and remember that $\cos^2\theta + \sin^2\theta = 1$. This gives us

$$-e^2\cos^2\theta + (1 - e^2)2aer\ \cos\ \theta + a^2e^2 + r^2 = a^2(1 - e^2). \tag{43}$$

Noting that

$$a^2(1-e^2) - a^2e^2(1-e^2) = a^2(1-e^2)^2, \tag{44}$$

we have

$$\begin{aligned} r^2 &= a^2(1-e^2)^2 - 2aer\cos\theta + e^2r^2\cos^2\theta \\ &= (er\ \cos\theta - a(1-e^2))^2. \end{aligned} \tag{45}$$

Since r is positive, we have

$$r = a(1-e^2) - er\ \cos\theta, \tag{46}$$

so gathering terms in r gives

$$r = \frac{a(1-e^2)}{1+e\ \cos\ \theta}. \tag{47}$$

This is the equation of the ellipse with the focus at the origin in polar coordinates.

4 The Hyperbola and Parabola

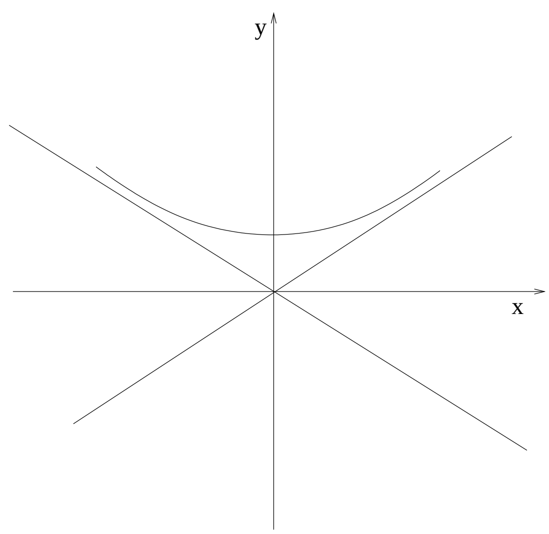

Figure 6: The hyperbola approaches two straight lines as $|x|$ be comes large. These straight lines, which the hyperbola becomes closer and closer to, but never touches are called *asymptotes.* There is a mirror image for negative y which is not drawn here.

The ellipse is the star of this little article, and I will just briefly outline the parabola and the hyperbola. Fist of all, I shall simply state that given a point, the focus, and a line, the directrix, we can construct the parabola and the hyperbola in the same manner as we did the ellipse. In the case of the parabola, the ratio e=1, and in the case of the hyperbola $e > 1$.

Generally, the parabola is often represented as $y^2 = 4ax$, (note that there is no x^2 term in eqn.9 if the eccentricity is one). It is the usual $y = x^2$ graph that we first come across in early secondary school. For the hyperbola, we can find coordinates such that eqn.9 looks like

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

$$y = \pm\sqrt{1 + \frac{a^2}{b^2}x^2}. \tag{48}$$

Note, that if x is very large we have

$$y \approx \pm\frac{a}{b}x. \tag{49}$$

For large values of x, the shape is close to a straight line. We saw the shape in the figure at the beginning of this section.

5 Dandelin's Spheres

In Fig 7, we see a "sidelong" view of the intersection of a plane and a cone resulting in an ellipse.

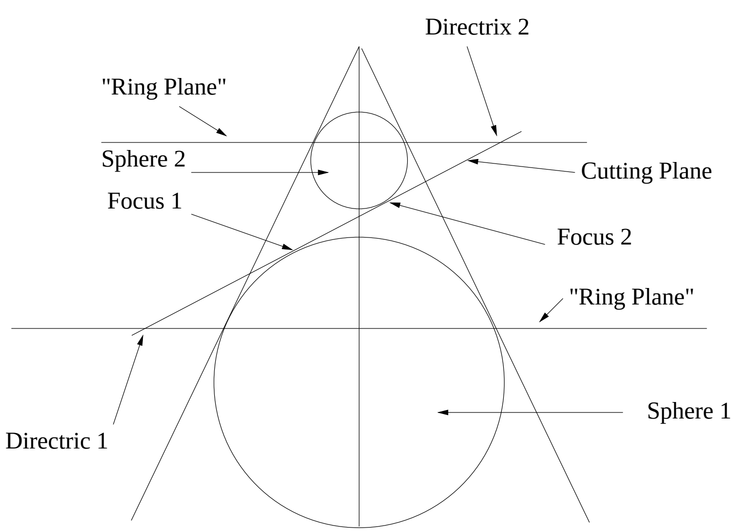

Figure 7: Dandelin's geometrical construction for the foccii and directrices.

We can pick any point on the axis of the cone, and find a sphere that just touches the surface of the cone. The sphere touches the cone in a horizontal circular ring (looking like a straight line in our sidelong view). We shall call the plane containing the ring a "ring plane" (this is not formal language).

However, there are only two spheres that just touch the cone *and* just touch the cutting plane. These are Dandelin's spheres. Dandelin showed that the points where these spheres touch the cutting plane are the foccii of the ellipse. Not only that, Dandelin showed that the lines where the cutting plane intersects the ring planes are the directrices of the ellipse.

So this is the relationship between the actual cone and cutting plane with the focus-directrix locus of a point definition.