Dr Oliver Mathematics
Diameter of a Conic

Contents

1 Introduction

"A chord of an ellipse is a line segment joining two points on the ellipse. An ellipse E has equation

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

The set of midpoints of the parallel chords of E with gradient m, where m is a constant, lie on the straight line l. Find an equation of l."

That, slightly paraphrased so that it can stand alone here, was question 8(c) of Edexcel's 2015 Further Pure Mathematics 3 paper. It was worth six marks and prompted (a) the usual flurry of complaints on social media that the examination was too hard and (b) the inevitable 'Hitler reacts to ...' video on YouTube.

A WHSB student's view as to whether this question was unfair turned largely on their ability to solve it; after all, no one who comes out of the exam hall having solved a problem that stumped both their local and national fellow students is likely to think that the question was unfair.

Effectively, the question is asking about a locus of points connected with an ellipse and, specifically, the diameter of an ellipse. (Yes, as we will see, it is not just circles that have diameters.) The Mathematics department therefore put together these notes of some properties of the conic sections in order to round out your knowledge of the subject. The vast majority of the content is not explicitly required by the specification but we think that all of it is potentially fair game as far as applications of your existing knowledge is concerned.

2 Diameter

2.1 What is a diameter?

Close your eyes and picture a 'diameter'. The chances are that you pictured a circle with a line going through the centre in some way, as shown in Figure 1.

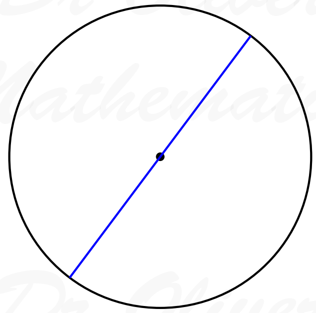

Figure 1: diameter of a circle

Now define what a diameter is (closing your eyes is optional here). The chances are that you came up with something along the lines of one of the following:

- a line joining two points on the circumference of the circle and that passes through the centre of the circle
- the longest possible chord of the circle,
- $\sup\{|\mathbf{x}-\mathbf{y}| : \mathbf{x}, \mathbf{y} \in \text{circle}\}$.

(Okay, well maybe not that last one although this is what we mean by the *generalised diameter* of a closed figure in n-dimensional Euclidean space. And if you are not sure what the sup function is then check out the appendix.)

Is there the equivalent of a 'diameter' for an ellipse? (Ellipses certainly have both a *major* and a *minor axis*.) What about for a parabola? (A parabola is not even a closed curve so could the concept of a diameter even make sense?) Or a hyperbola? (If there is such a thing as the 'diameter' of an ellipse and a parabola, shouldn't there be one for a hyperbola, even if it does come in two pieces?)

2.2 Diameter of an ellipse

Suppose that we have an ellipse E with standard equation

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

Consider a set of parallel chords of gradient m for the ellipse: see Figure 2. Can we say anything about the set of midpoints of these chords?

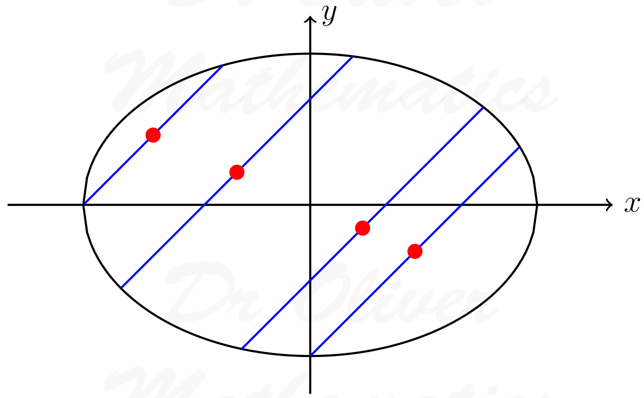

Figure 2: chords of constant gradient for an ellipse

You would have to be pretty lacking in imagination (and probably blind) not to guess that the midpoints of these chords all lie on a straight line, perhaps even a straight line that passes through the origin — and they do. So let's prove that this is the case.

Suppose that the chord has equation $y = mx + c$. Let $A(x_A, y_A)$ and $B(x_B, y_B)$ be the points where the chord intersects with the ellipse and $M(x_M, y_M)$ be the midpoint of the chord. Then we can find the intersection of the line and the ellipse as follows:

$$\begin{aligned}\frac{x^2}{a^2} + \frac{(mx+c)^2}{b^2} = 1 &\Rightarrow b^2x^2 + a^2(mx+c)^2 = a^2b^2 \\ &\Rightarrow (a^2m^2 + b^2)x^2 + 2a^2cmx + a^2(c^2 - b^2) = 0.\end{aligned}$$

x_A and x_B satisfy the quadratic equation above and, in particular,

$$x_A + x_B = -\frac{2a^2cm}{a^2m^2 + b^2}$$

giving

$$x_M = \tfrac{1}{2}(x_A + x_B) = -\frac{a^2cm}{a^2m^2 + b^2}$$

and

$$\begin{aligned}y_M &= mx_M + c \\ &= -\frac{a^2cm^2}{a^2m^2 + b^2} + c \\ &= \frac{-a^2cm^2 + c(a^2m^2 + b^2)}{a^2m^2 + b^2} \\ &= \frac{b^2c}{a^2m^2 + b^2}.\end{aligned}$$

So the locus of such midpoints is

$$b^2x + a^2my = 0 \text{ or } y = -\frac{b^2}{a^2m}x.$$

2.3 Diameter of a parabola

Let us apply the same idea to the parabola, i.e., consider the locus of the midpoints of parallel chords as shown in Figure 3.

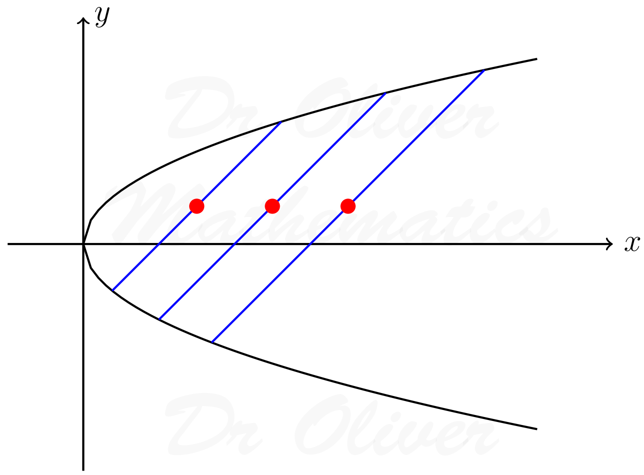

Figure 3: chords of constant gradient for a parabola

Suppose that P is a parabola with standard equation $y^2 = 4ax$. Suppose that the chord has equation $y = mx + c$. Let $A(x_A, y_A)$ and $B(x_B, y_B)$ be the points where the chord intersects with the parabola and $M(x_M, y_M)$ be the midpoint of the chord. Then we can find the intersection of the line and the parabola as follows:

$$\begin{aligned} y^2 = 4ax &\Rightarrow y^2 = 4a\left(\frac{y-c}{m}\right) \\ &\Rightarrow y^2 - \frac{4a}{m}y + \frac{4ac}{m}. \end{aligned}$$

Hence

$$y_A + y_B = \frac{4a}{m}$$

and so

$$y_M = \tfrac{1}{2}(y_A + y_B) = \frac{2a}{m}.$$

Note that this is a constant and so the diameter is a straight line that is parallel to the parabola's axis.

2.4 Diameter of a hyperbola

Suppose that we have a hyperbola H with standard equation

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

Consider a set of parallel chords of gradient m for the hyperbola: see Figure 4 for the case where $|m| > \frac{b}{a}$ and Figure 5 for the case where $|m| < \frac{b}{a}$. (Why is there no picture for the case $|m| = \frac{b}{a}$?)

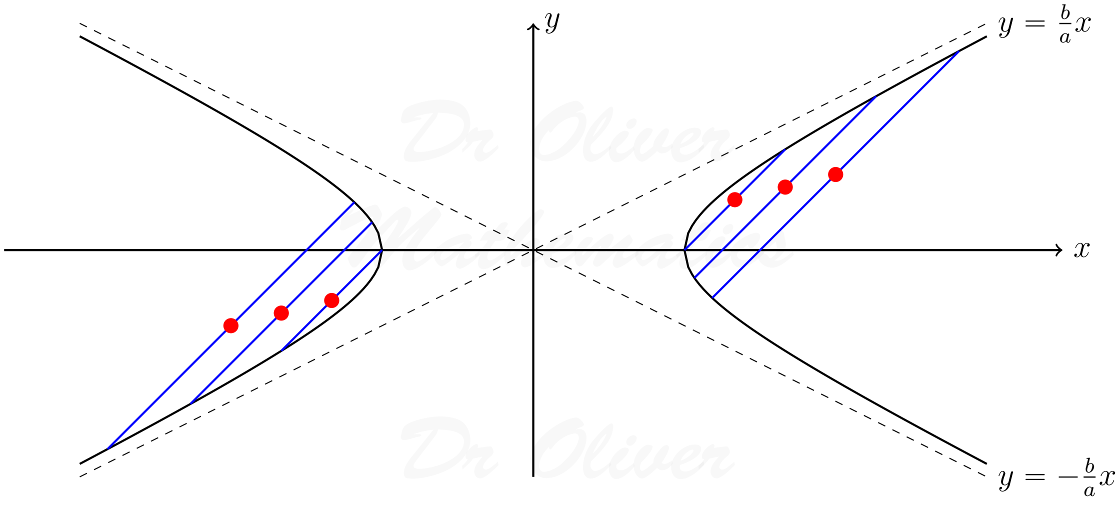

Figure 4: chords of constant gradient for a hyperbola, $|m| > \frac{b}{a}$

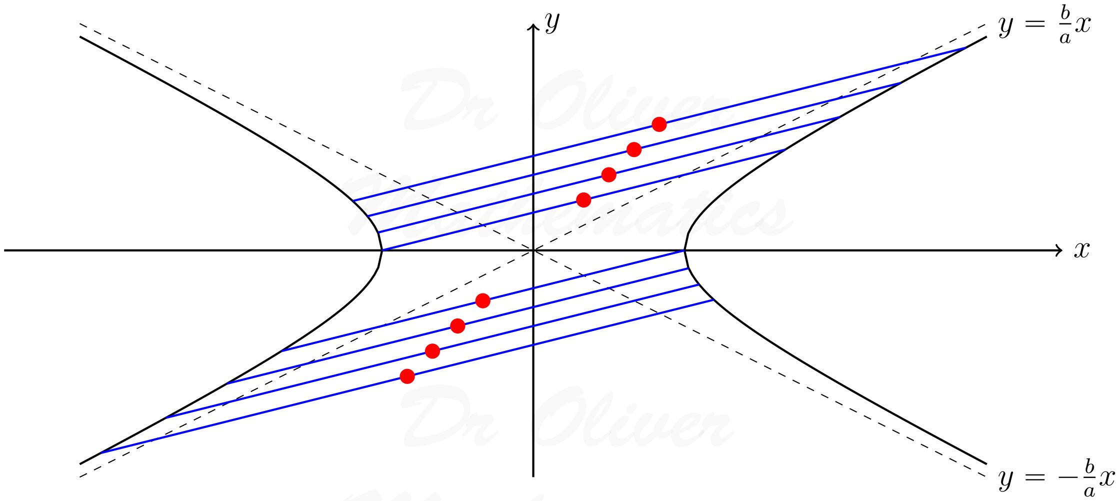

Figure 5: chords of constant gradient for a hyperbola, $|m| < \frac{b}{a}$

Suppose that the chord has equation $y = mx + c$. Let $A(x_A, y_A)$ and $B(x_B, y_B)$ be the points where the chord intersects with the hyperbola and $M(x_M, y_M)$ be the midpoint of the chord.

Then we can find the intersection of the line and the hyperbola as follows:

$$\begin{aligned}\frac{x^2}{a^2} - \frac{(mx+c)^2}{b^2} = 1 &\Rightarrow b^2x^2 - a^2(mx+c)^2 = a^2b^2 \\ &\Rightarrow (b^2 - a^2m^2)x^2 - 2a^2cmx - a^2(b^2+c^2) = 0.\end{aligned}$$

x_A and x_B satisfy the quadratic equation above and, in particular,

$$x_A + x_B = \frac{2a^2cm}{b^2 - a^2m^2}$$

giving

$$x_M = \tfrac{1}{2}(x_A + x_B) = \frac{a^2cm}{b^2 - a^2m^2}$$

and

$$\begin{aligned}y_M &= mx_M + c \\ &= \frac{a^2cm^2}{b^2 - a^2m^2} + c \\ &= \frac{a^2cm^2 + c(b^2 - a^2m^2)}{b^2 - a^2m^2} \\ &= \frac{b^2c}{b^2 - a^2m^2}.\end{aligned}$$

So the locus of such midpoints is

$$b^2x - a^2my = 0 \text{ or } y = \frac{b^2}{a^2m}x.$$

(Did you note that this is the same as for the ellipse, except for the sign? Spooky ...)

2.5 Diameter of a circle

Does this approach hold for a circle?

On a personal level, I don't think that it is a good idea to consider the circle as a degenerate case of an ellipse. If you are the kind of person who does, can you explain why the circle has eccentricity $e = 0$ — the consequence of setting $a = b$ — whereas for an ellipse $0 < e < 1$, why we only have one focus and not two foci, and where the directrix comes into this?

I prefer to think of the circle as a fourth type of conic section rather than a case of the ellipse where $a = b$ — Apollonius of Perga thought the same way and since it was he (in his book *The Conics*) who named the ellipse, the parabola, and the hyperbola as well as doing pioneering work on the conic sections I think I'm in good company.

But you know, however, from GCSE that this approach *does* hold: the perpendicular bisector of the chord passes through the centre of a circle, as shown in Figure 6.

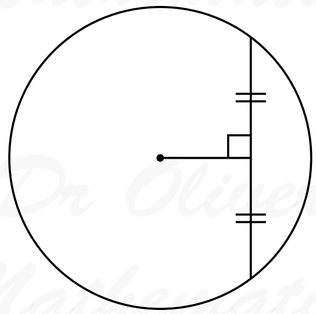

Figure 6: perpendicular bisector of a chord of a circle

You can prove this with simple congruence arguments. It is also the way that you can find the centre of a circle that passes through three non-concurrent points: find the point of intersection of any two of the three perpendicular bisectors of the three points. Hence the locus of midpoints of parallel chords of a circle all lie on a straight line that passes through the centre of the circle.

2.6 Summary

The results can be summarised in Table 1 as follows.

Conic	**Standard equation**	**Diameter for** $y = mx + c$
Ellipse	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$	$y = -\frac{b^2}{a^2 m} x$
Parabola	$y^2 = 4ax$	$y = \frac{2a}{m}$
Hyperbola	$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$	$y = \frac{b^2}{a^2 m} x$

Table 1: summary table

3 Chords

3.1 Focal chord of a parabola

Let $P(ap^2, 2ap)$ and $Q(aq^2, 2aq)$ be distinct points on the parabola $y^2 = 4ax$ such that the chord PQ passes through the focus $S(a, 0)$. The the tangent to the parabola at P is perpendicular to the tangent to the parabola at Q. Moreover, if T is the point of intersection of the tangents, then T lies on the directrix $x = -a$ and the line segment ST is perpendicular to the chord PQ, as shown in Figure 7.

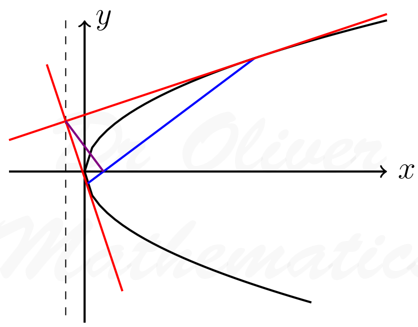

Figure 7: tangents from a focal chord of a parabola

The gradient of the chord PQ is

$$\frac{2aq - 2ap}{aq^2 - ap^2} = \frac{2a(q-p)}{a(q+p)(q-p)} = \frac{2}{q+p};$$

note that the denominator is zero if $p = -q$ and that this happens only if the chord PQ is vertical. Otherwise, the equation of the chord PQ is

$$y - 2aq = \frac{2}{q+p}(x - aq^2).$$

Suppose that this chord passes through the focus of the parabola $S(a, 0)$. Then

$$\begin{aligned} -2aq = \frac{2}{q+p}(a - aq^2) &\Rightarrow -2aq^2 - 2apq = 2a - 2aq^2 \\ &\Rightarrow pq = -1. \end{aligned}$$

Now, recall how to derive the tangent to the curve:

$$y^2 = 4ax \Rightarrow 2y\frac{\mathrm{d}y}{\mathrm{d}x} = 4a \Rightarrow \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2a}{y}$$

and so the gradients at P and Q are $\frac{1}{p}$ and $\frac{1}{q}$ respectively. The product of these gradients is

$$\frac{1}{p} \times \frac{1}{q} = \frac{1}{pq} = -1,$$

and hence the tangents are perpendicular.

The equations of the tangents are

$$\begin{aligned}\text{tangent at } P: \quad y - 2ap &= \frac{1}{p}(x - ap^2)\\ \text{tangent at } Q: \quad y - 2aq &= \frac{1}{q}(x - aq^2),\end{aligned}$$

and to find the point of intersection we simply subtract:

$$\begin{aligned}&\Rightarrow \quad -2ap - (-2aq) = \frac{1}{p}(x - ap^2) - \frac{1}{q}(x - aq^2)\\ &\Rightarrow \quad 2a(q-p) = \left(\tfrac{1}{p} - \tfrac{1}{q}\right)x - ap + aq\\ &\Rightarrow \quad 2a(q-p) = \tfrac{q-p}{pq}x + a(q-p)\\ &\Rightarrow \quad 2a = \tfrac{1}{pq}x + a \text{ (since } p \neq q)\\ &\Rightarrow \quad a = -x \text{ (since } pq = -1)\\ &\Rightarrow \quad x = -a,\end{aligned}$$

and so T lies on the directrix, as required.

Finally, the y-coordinate of T is given by

$$y = \frac{1}{p}(-a - ap^2) + 2ap = \frac{(-a + ap^2) + 2ap^2}{p} = \frac{a(p^2-1)}{p};$$

the expression that you get if you use the tangent at Q is identical since $pq = -1$. The gradient of ST is

$$\frac{\frac{a(p^2-1)}{p} - 0}{-a - a} = \frac{1-p^2}{2p} = \frac{-pq - p^2}{2p} = \frac{-p(p+q)}{2p} = -\frac{p+q}{2};$$

since the gradient of the chord PQ is $\frac{2}{p+q}$, ST and PQ are perpendicular.

3.2 Chord of an ellipse

Let $A(a\cos\theta, b\sin\theta)$ and $B(a\cos\phi, b\sin\phi)$ be distinct points on the ellipse

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

The gradient of the chord AB is given by

$$\begin{aligned}\frac{b\sin\theta - b\sin\phi}{a\cos\theta - a\cos\phi} &= \frac{2b\cos\left(\frac{\phi+\theta}{2}\right)\sin\left(\frac{\phi-\theta}{2}\right)}{-2a\sin\left(\frac{\phi+\theta}{2}\right)\sin\left(\frac{\phi-\theta}{2}\right)} \text{ (using the factor formulae)}\\ &= -\frac{b\cos\left(\frac{\phi+\theta}{2}\right)}{a\sin\left(\frac{\phi+\theta}{2}\right)}.\end{aligned}$$

and hence the equation of the chord is

$$
\begin{aligned}
& y - b\sin\theta = -\frac{b\cos\left(\frac{\phi+\theta}{2}\right)}{a\sin\left(\frac{\phi+\theta}{2}\right)}(x - a\cos\theta) \\
\Rightarrow\quad & ay\sin\left(\tfrac{\phi+\theta}{2}\right)\cos\theta - ab\sin\left(\tfrac{\phi+\theta}{2}\right)\sin\theta = -bx\cos\left(\tfrac{\phi+\theta}{2}\right) + ab\cos\left(\tfrac{\phi+\theta}{2}\right)\cos\theta \\
\Rightarrow\quad & bx\cos\left(\tfrac{\phi+\theta}{2}\right) + ay\sin\left(\tfrac{\phi+\theta}{2}\right)\cos\theta = ab\left[\cos\left(\tfrac{\phi+\theta}{2}\right)\cos\theta + \sin\left(\tfrac{\phi+\theta}{2}\right)\sin\theta\right] \\
\Rightarrow\quad & bx\cos\left(\tfrac{\phi+\theta}{2}\right) + ay\sin\left(\tfrac{\phi+\theta}{2}\right)\cos\theta = ab\cos\left(\tfrac{\phi-\theta}{2}\right),
\end{aligned}
$$

and note that we simply need the appropriate sum and differences of the parameters θ and ϕ to form the equation.

3.3 Focal chord of an ellipse

Suppose that PQ is a chord on the ellipse

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

which passes through the focus $S(ae, 0)$. Let T be the point where the tangent to the ellipse at P and the tangent to the ellipse at Q meet. Then T lies on the directrix $x = \frac{a}{e}$ of the ellipse and the line segment ST is perpendicular to the chord PQ.

Suppose that $P(a\cos\theta, b\sin\theta)$ and $Q(a\cos\phi, b\sin\phi)$ be distinct points on the ellipse and that the chord PQ passes through the focus $S(ae, 0)$. The equations of the tangents are

$$
\begin{aligned}
\text{tangent at } P: \quad & bx\cos\theta + ay\sin\theta = ab \Rightarrow bx\cos\theta\sin\phi + ay\sin\theta\sin\phi = ab\sin\phi \\
\text{tangent at } Q: \quad & bx\cos\phi + ay\sin\phi = ab \Rightarrow bx\cos\phi\sin\theta + ay\sin\phi\sin\theta = ab\sin\theta,
\end{aligned}
$$

and to find their point of intersection we just need to subtract:

$$
\begin{aligned}
& bx(\cos\theta\sin\phi - \cos\phi\sin\theta) = ab(\sin\phi - \sin\theta) \\
\Rightarrow\quad & x\sin(\phi - \theta) = a(\sin\phi - \sin\theta) \\
\Rightarrow\quad & x = \frac{a(\sin\phi - \sin\theta)}{\sin(\phi - \theta)}.
\end{aligned}
$$

Now the chord PQ has equation

$$y - b\sin\theta = \frac{b\sin\theta - b\sin\phi}{a\cos\theta - a\cos\phi}(x - a\cos\theta).$$

Since this passes through the focus $S(ae, 0)$,

$$
\begin{aligned}
& -b\sin\theta = \frac{b\sin\theta - b\sin\phi}{a\cos\theta - a\cos\phi}(ae - a\cos\theta) \\
\Rightarrow\quad & -\sin\theta\cos\theta + \sin\theta\cos\phi = e\sin\theta - \sin\theta\cos\theta - e\sin\phi + \sin\phi\cos\theta \\
\Rightarrow\quad & \sin\theta\cos\phi - \sin\phi\cos\theta = e\sin\theta - e\sin\phi \\
\Rightarrow\quad & \sin(\theta - \phi) = e(\sin\theta - \sin\phi),
\end{aligned}
$$

and hence the x-coordinate of T, the point of intersection of the tangents is $\frac{a}{e}$ and so the tangents meet on the directrix, as required.

Finally, the y-coordinate of T is given by

$$\begin{aligned} \tfrac{ab}{e}\cos\theta + ay\sin\theta = ab &\Rightarrow b\cos\theta + aey\sin\theta = be \\ &\Rightarrow aey\sin\theta = b(e-\cos\theta) \\ &\Rightarrow y = \frac{b(e-\cos\theta)}{ae\sin\theta}. \end{aligned}$$

3.4 Focal chord of a hyperbola

Suppose that $P(a\sec p, b\tan p)$ and $Q(a\sec q, b\tan q)$ are points on different branches of the hyperbola

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

such that the chord PQ, when extended, passes through the focus $S(ae, 0)$. Let T be the point where the tangent to the hyperbola at P and the tangent to the hyperbola at Q meet. Then T lies on the directrix $x = \frac{a}{e}$ of the hyperbola and the line segment ST is perpendicular to the chord PQ.

The equations of the tangents are

$$\begin{aligned} \text{tangent at } P: ay\sinh p + ab = bx\cosh p &\Rightarrow ay\sinh p\sinh q + ab\sinh q = bx\cosh p\sinh q \\ \text{tangent at } Q: ay\sinh q + ab = bx\cosh q &\Rightarrow ay\sinh q\sinh p + ab\sinh p = bx\cosh q\sinh p, \end{aligned}$$

and to find the point of intersection we just need to subtract:

$$\begin{aligned} & ab\sinh q - ab\sinh p = bx\cosh p\sinh q - bx\cosh q sinhp \\ \Rightarrow\quad & a(\sinh q - \sinh p) = x(\cosh p\sinh q - \cosh q sinhp) \\ \Rightarrow\quad & a(\sinh q - \sinh p) = x\sinh(q-p) \\ \Rightarrow\quad & x = \frac{a(\sinh q - \sinh p)}{\sinh(q-p)}. \end{aligned}$$

Now the chord PQ has equation

$$y - b\sinh p = \frac{b\sinh p - b\sinh q}{a\cosh p - a\cosh q}(x - a\cosh t).$$

Since this passes through the focus $(ae, 0)$,

$$
\begin{aligned}
& -b\sinh p = \frac{b\sinh p - b\sinh q}{a\cosh p - a\cosh q}(ae - a\cosh p) \\
\Rightarrow\quad & -ab\sinh p(\cosh p - \cosh q) = ab(\sinh p - \sinh q)(e - \cosh p) \\
\Rightarrow\quad & -\sinh p(\cosh p - \cosh q) = (\sinh p - \sinh q)(e - \cosh p) \\
\Rightarrow\quad & -\sinh p\cosh p + \sinh p\cosh q = e(\sinh p - \sinh q) - \sinh p\cosh p + \sinh q\cosh q \\
\Rightarrow\quad & \sinh p\cosh q - \sinh q\cosh q = e(\sinh p - \sinh q) \\
\Rightarrow\quad & \sinh(p - q) = e(\sinh p - \sinh q),
\end{aligned}
$$

and hence the x-coordinate of T, the point of intersection of the tangents is $\frac{a}{e}$ and so the tangents meet on the directrix, as required.

4 Focal properties

4.1 Focal property of a parabola

If a ray of light is parallel to the axis of the parabola then it will be reflected to the focus, as shown in Figure 8.

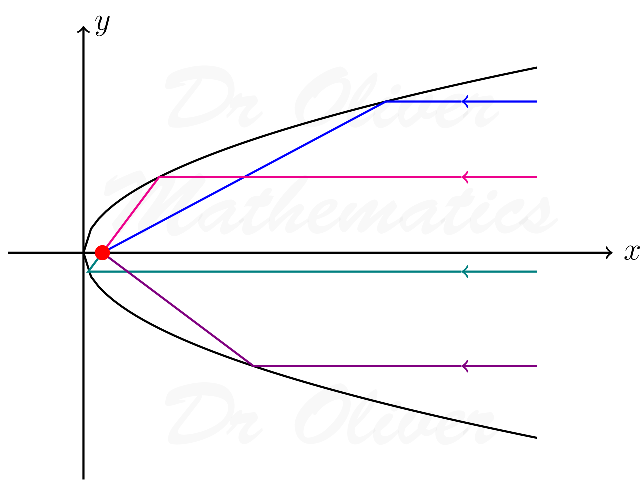

Figure 8: focal property of a parabola

Conversely, a source of light positioned at the focus will produce a beam of light that travels parallel to the parabola's axis.

Let $P(at^2, 2at)$ be a point on the parabola $y^2 = 4ax$ with focus $S(a, 0)$. Suppose that the normal to the parabola at P cuts the x-axis at N: then it will suffice to show that PN bisects the angle $Q\hat{P}S$, as shown in Figure 9.

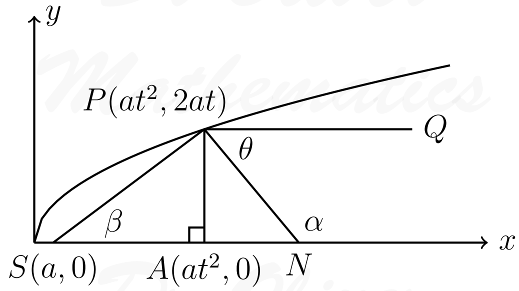

Figure 9: normal to the parabola

The equation of the normal at P has equation

$$tx + y = at^3 + 2at;$$

hence the gradient of the line is $-t$ and so $\tan\alpha = -t$. Next,

$$\tan\theta = \tan(\pi - \alpha) = -\tan\alpha = t.$$

In addition, by considering triangle PAS,

$$\tan\beta = \frac{2at}{at^2 - a} = \frac{2t}{t^2 - 1}.$$

Finally,

$$\tan Q\hat{P}S = \tan(\pi - \beta) = -\tan\beta = \frac{2t}{1 - t^2} = \tan 2\theta.$$

Hence $Q\hat{P}S = 2Q\hat{P}N$ and so PN bisects $Q\hat{P}S$, as required.

4.2 Focal property of an ellipse

If a ray of light is emitted from one focus then it will be reflected off the ellipse to the other focus, as shown in Figure 10.

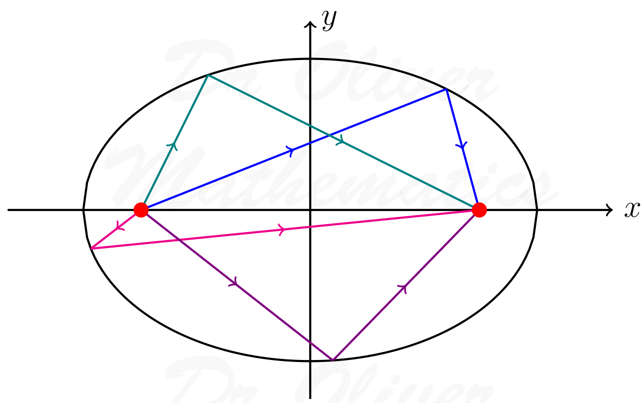

Figure 10: focal property of an ellipse

Note, that since all of the four paths shown are precisely the same lengths, light rays emitted from one focus will all arrive at the other focus at exactly the same time.

In order to prove this result, it will suffice to show that the normal at the perimeter of the ellipse bisects the angle as this establishes that the angle of incidence is equal to the angle of reflection.

Let $P(a\cos t, b\sin t)$ be a point on the ellipse

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

and let $A(-ae, 0)$ and $B(ae, 0)$ be the two foci. The normal to the ellipse at P is

$$ax\sin t - by\cos t = (a^2 - b^2)\sin t\cos t$$

(this is a standard result that you should be able to derive both quickly and correctly) and, since $a^2 - b^2 = a^2e^2$, this crosses the x-axis at the point $C(e^2\cos t, 0)$; this is summarised in Figure 11.

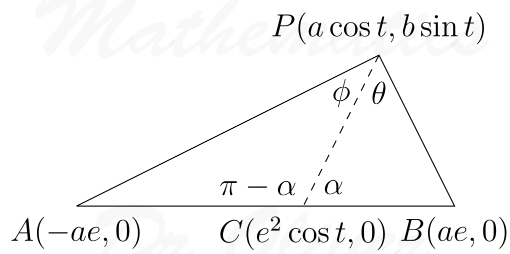

Figure 11: the normal to the ellipse

$$\begin{aligned}
PB^2 &= (ae - a\cos t)^2 + (b\sin t)^2 \\
&= a^2e^2 - 2a^2e\cos t + a^2\cos^2 t + b^2\sin^2 t \\
&= a^2e^2 - 2a^2e\cos t + (a^2 - b^2)\cos^2 t + b^2\cos^2 t + b^2\sin^2 t \\
&= a^2e^2 - 2a^2e\cos t + (a^2 - b^2)\cos^2 t + b^2 \\
&= a^2e^2\cos^2 t - 2a^2e\cos t + a^2 \\
&= a^2(e^2\cos^2 t - 2e\cos t + 1) \\
&= a^2(e\cos t - 1)^2,
\end{aligned}$$

and hence $PB = a(1 - e\cos t)$. Since $AP + PB = 2a$, this gives $PA = a(1 + e\cos t)$.

We now apply the sine rule to triangle APC:

$$\begin{aligned}\sin\phi &= \frac{AC\sin(\pi-\alpha)}{AP}\\ &= \frac{(ae+e^2\cos t)\times\sin\alpha}{a(1+e\cos t)}\\ &= e\sin\alpha.\end{aligned}$$

We now apply the sine rule to triangle CPB:

$$\begin{aligned}\sin\theta &= \frac{BC\sin\alpha}{PB}\\ &= \frac{(ae-e^2\cos t)\sin\alpha}{a(1-e\cos t)}\\ &= e\sin\alpha.\end{aligned}$$

So $\sin\phi=\sin\theta$ and, since the angles cannot be supplementary, $\phi=\theta$.

4.3 Focal property of a hyperbola

If a ray is directed at one focus of a hyperbola but strikes the other branch of the hyperbola first then the ray will be reflected to the other focus, as shown in Figure 12.

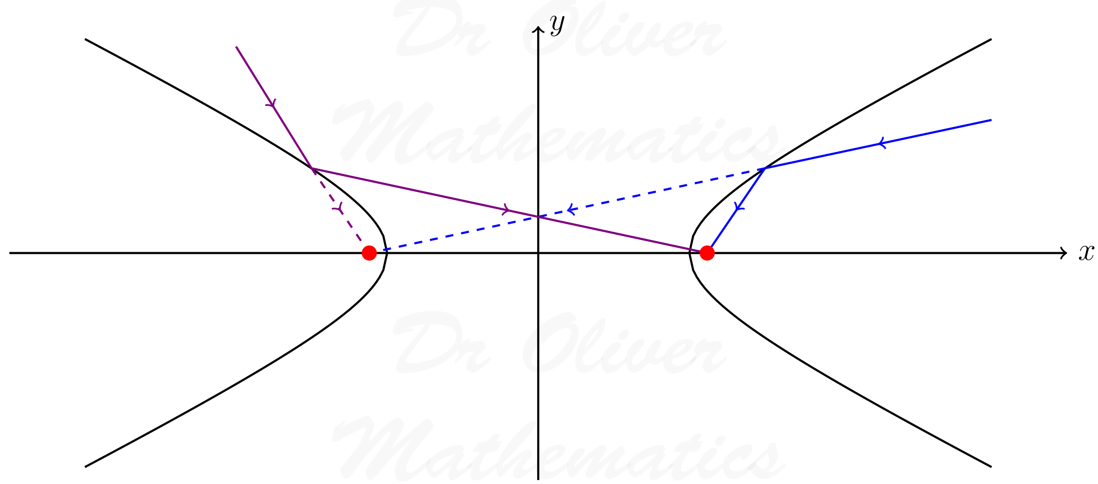

Figure 12: focal property of the hyperbola

In order to prove this result, it will suffice to show that the tangent at the perimeter of the hyperbola bisects the angle as this establishes that the angle of incidence is equal to the angle of reflection.

Let $P(a \sec t, b \tan t)$ be a point on the hyperbola

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

and let $A(-ae, 0)$ and $B(ae, 0)$ be the two foci. (Note: this means that the proof is valid for only the right-hand branch of the parabola but you can repeat the argument with the point $(-a \sec t, b \tan t)$ for the other branch.) The tangent to the hyperbola at P is

$$bx \sec t - ay \tan t = ab$$

and this crosses the x-axis at the point $C(a \cos t, 0)$; this is summarised in Figure 13.

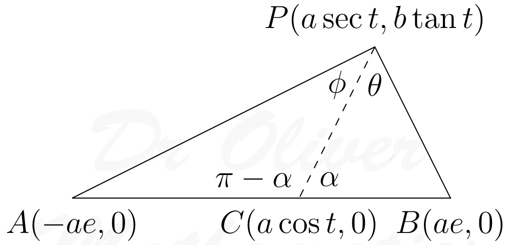

Figure 13: the tangent to the hyperbola

$$\begin{aligned}
PB^2 &= (ae - a \sec t)^2 + (b \tan t)^2 \\
&= a^2e^2 - 2a^2e \sec t + a^2 \sec^2 t + b^2 \tan^2 t \\
&= a^2e^2 - 2a^2e \sec t + a^2 \sec^2 t + b^2(\sec^2 t - 1) \\
&= (a^2e^2 - b^2) - 2a^2e \sec t + (a^2 + b^2) \sec^2 t \\
&= a^2 - 2a^2e \sec t + a^2e^2 \sec^2 t \\
&= a^2(1 - 2e \sec t + e^2 \sec^2 t) \\
&= a^2(1 - e \sec t)^2,
\end{aligned}$$

and hence $PB = a(e \sec t - 1)$. Since $|AP - PB| = 2a$, this gives $PA = a(e \sec t + 1)$.

We now apply the sine rule to triangle APC:

$$\begin{aligned}
\sin \phi &= \frac{AC \sin(\pi - \alpha)}{AP} \\
&= \frac{(ae + a \cos t) \times \sin \alpha}{a(e \sec t + 1)} \\
&= \frac{(e + \cos t) \sin \alpha}{e \sec t + 1}.
\end{aligned}$$

We now apply the sine rule to triangle CPB:

$$\begin{aligned}\sin\theta &= \frac{BC\sin\alpha}{PB}\\ &= \frac{(ae-a\cos t)\sin\alpha}{a(e\sec t-1)}\\ &= \frac{(e-\cos t)\sin\alpha}{e\sec t-1}\end{aligned}$$

Unlike in the case of the ellipse we are not quite finished: we need to show that

$$\frac{e+\cos t}{e\sec t+1} = \frac{e-\cos t}{e\sec t-1}$$

in order to establish that the two angles have the same sine. To do this, simply divide one by the other:

$$\begin{aligned}\frac{e+\cos t}{e\sec t+1} \div \frac{e-\cos t}{e\sec t-1} &= \frac{e+\cos t}{e\sec t+1} \times \frac{e\sec t-1}{e-\cos t}\\ &= \frac{e^2\sec t-e+e-\cos t}{e^2\sec t-e+e-\cos t}\\ &= 1.\end{aligned}$$

So $\sin\phi = \sin\theta$ and, since the angles cannot be supplementary, $\phi = \theta$.

5 Tangent property of a hyperbola

Suppose that the tangent to a hyperbola at the point P cuts the asymptotes are the points A and B, as shown in Figure 14. Then $PA = PB$.

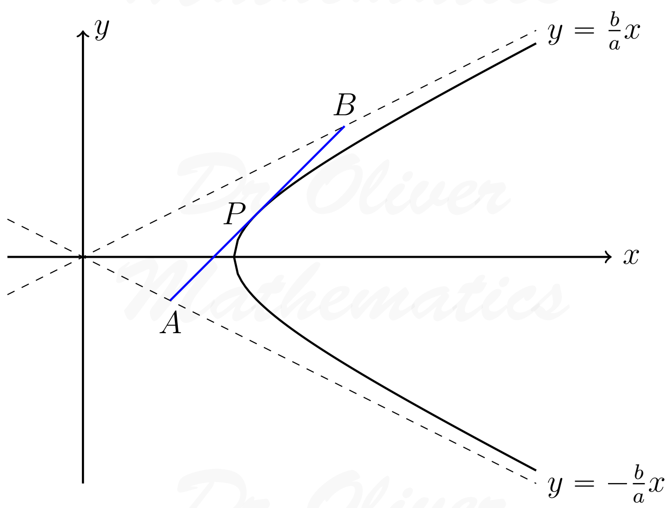

Figure 14: tangent to a hyperbola

Suppose that $P(a \sec\theta, b\tan\theta)$ is a point on the hyperbola

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

Then the equation of the tangent is

$$bx\sec\theta - ay\tan\theta = ab.$$

If A is the point of intersection between the tangent and the asymptote $y = \frac{b}{a}x$ then the x-coordinate of A is given by

$$\begin{aligned} bx_A\sec\theta - a\left(\frac{b}{a}x_A\right)\tan\theta = ab &\Rightarrow bx_A\sec\theta - bx_A\tan\theta = ab \\ &\Rightarrow bx_A(\sec\theta - \tan\theta) = ab \\ &\Rightarrow x_A = \frac{a}{\sec\theta - \tan\theta}. \end{aligned}$$

In the same way, if B is the point of intersection between the tangent and the asymptote $y = \frac{b}{a}x$ then the x-coordinate of B is given by

$$\begin{aligned} bx_B\sec\theta - a\left(-\frac{b}{a}x_B\right)\tan\theta = ab &\Rightarrow bx_B\sec\theta + bx_B\tan\theta = ab \\ &\Rightarrow bx_B(\sec\theta + \tan\theta) = ab \\ &\Rightarrow x_B = \frac{a}{\sec\theta + \tan\theta}. \end{aligned}$$

So the x-coordinate of the midpoint of AB is given by

$$\begin{aligned} \tfrac{1}{2}(x_A + x_B) &= \tfrac{1}{2}\left[\frac{a}{\sec\theta - \tan\theta} + \frac{a}{\sec\theta + \tan\theta}\right] \\ &= \frac{a(\sec\theta + \tan\theta) + a(\sec\theta - \tan\theta)}{2(\sec^2\theta - \tan^2\theta)} \\ &= a\sec\theta, \end{aligned}$$

and this is simply the x-coordinate of P. Hence P is the midpoint of AB, as required.

6 Arc length

6.1 Arc length of a parabola

What is the length of the parabola $y = x^2$ between the points $(0, 0)$ and $(1, 1)$?

$$y = x^2 \Rightarrow \frac{\mathrm{d}y}{\mathrm{d}x} = 2x \Rightarrow \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2} = \sqrt{1 + (2x)^2} = \sqrt{1 + 4x^2}$$

and hence the arc length l is given by

$$l = \int_0^1 \sqrt{1 + 4x^2}\,\mathrm{d}x.$$

As is so often the case, a relatively simple function will give rise to a very complicated integral for its arc length. There are three standard methods of solving the problem:

(a) use the substitution $\theta = \arctan 2x$,

(b) use the substitution $t = \operatorname{arsinh} 2x$,

(c) use integration by parts.

We will derive the result in each of these three ways.

(a) $\theta = \arctan 2x \Rightarrow x = \frac{1}{2}\tan\theta$ and so

$$\sqrt{1 + 4x^2} = \sqrt{1 + \tan^2\theta} = \sec\theta;$$

note that we do not need to worry about signs here: the function $\arctan\theta$ is defined on the interval $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$ and, since $\cos\theta$ is always positive in this interval, $\sec\theta$ is also positive.

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = \tfrac{1}{2}\sec^2 u \Rightarrow \mathrm{d}\theta = \tfrac{1}{2}\sec^2\theta\,\mathrm{d}\theta.$$

We also need to change the limits:

$$x = 0 \Rightarrow \theta = 0 \text{ and } x = 1 \Rightarrow \theta = \arctan 2.$$

Hence

$$\begin{aligned} l &= \int_0^1 \sqrt{1 + 4x^2}\,\mathrm{d}x \\ &= \tfrac{1}{2}\int_0^{\arctan 2} \sec^3\theta\,\mathrm{d}\theta \\ &= \tfrac{1}{2}\int_0^{\arctan 2} \sec\theta\sec^2\theta\,\mathrm{d}\theta \end{aligned}$$

and we now use integration by parts:

$$u = \sec\theta \Rightarrow \frac{\mathrm{d}u}{\mathrm{d}\theta} = \sec\theta\tan\theta \text{ and } \frac{\mathrm{d}v}{\mathrm{d}\theta} = \sec^2\theta \Rightarrow v = \tan\theta.$$

Hence

$$\begin{aligned} l &= \tfrac{1}{2}\left[\sec\theta\tan\theta\right]_{\theta=0}^{\arctan 2} - \tfrac{1}{2}\int_0^{\arctan 2} \sec\theta\tan^2\theta\,\mathrm{d}\theta \\ &= \tfrac{1}{2}\left[\sec\theta\tan\theta\right]_{\theta=0}^{\arctan 2} - \tfrac{1}{2}\int_0^{\arctan 2} \sec\theta(\sec^2\theta - 1)\,\mathrm{d}\theta \\ &= \tfrac{1}{2}\left[\sec\theta\tan\theta\right]_{\theta=0}^{\arctan 2} - l + \tfrac{1}{2}\int_0^{\arctan 2} \sec\theta\,\mathrm{d}\theta \end{aligned}$$

and hence

$$2l = \tfrac{1}{2}\left[\sec\theta\tan\theta + \ln|\sec\theta + \tan\theta|\right]_{\theta=0}^{\arctan 2};$$

hence

$$l = \tfrac{1}{4}\left[\sec\theta\tan\theta + \ln|\sec\theta + \tan\theta|\right]_{\theta=0}^{\arctan 2}.$$

Now,

$$\sec^2\theta = 1 + \tan^2\theta \Rightarrow \sec(\arctan 2) = \sqrt{5};$$

as explained above there is no need to worry about signs here. Hence

$$l = \tfrac{1}{4}\left[2\sqrt{5} + \ln\left(\sqrt{5} + 2\right)\right].$$

(b) $t = \operatorname{arsinh} 2x \Rightarrow x = \frac{1}{2}\sinh t$ and so

$$\sqrt{1 + 4x^2} = \sqrt{1 + \sinh^2 t} = \cosh t,$$

and, again, we do not need to worry about signs (why?).

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \tfrac{1}{2}\cosh t \Rightarrow \mathrm{d}x = \tfrac{1}{2}\cosh t\,\mathrm{d}t.$$

We also need to change the limits:

$$x = 0 \Rightarrow t = 0 \text{ and } x = 1 \Rightarrow t = \operatorname{arsinh} 2.$$

Hence

$$\begin{aligned} l &= \int_0^1 \sqrt{1 + 4x^2}\,\mathrm{d}x \\ &= \tfrac{1}{2}\int_0^{\operatorname{arsinh} 2} \cosh^2 t\,\mathrm{d}t \\ &= \tfrac{1}{4}\int_0^{\operatorname{arsinh} 2} (1 + \cosh 2t)\,\mathrm{d}t \\ &= \tfrac{1}{4}\left[t + \tfrac{1}{2}\sinh 2t\right]_{t=0}^{\operatorname{arsinh} 2}. \end{aligned}$$

Now $\operatorname{arsinh} 2 = \ln(2 + \sqrt{5})$ and

$$\sinh 2t = 2\sinh t\cosh t = 2\sinh t\sqrt{1 + \sinh^2 t}$$

to give $\sinh 2(\operatorname{arsinh} 2) = 2 \times 2 \times \sqrt{5} = 4\sqrt{5}$. Hence

$$l = \tfrac{1}{4}\left[\ln(2 + \sqrt{5}) + 2\sqrt{5}\right].$$

(c)

$$u = \sqrt{1+4x^2} \Rightarrow \frac{\mathrm{d}u}{\mathrm{d}x} = \frac{4x}{\sqrt{1+4x^2}} \text{ and } \frac{\mathrm{d}v}{\mathrm{d}x} = 1 \Rightarrow v = x$$

gives

$$\begin{aligned}
\int_0^1 \sqrt{1+4x^2}\,\mathrm{d}x &= \left[x\sqrt{1+4x^2}\right]_{x=0}^1 - \int_0^1 \frac{4x^2}{\sqrt{1+4x^2}}\,\mathrm{d}x \\
&= (\sqrt{5}-0) - \int_0^1 \frac{(1+4x^2)-1}{\sqrt{1+4x^2}}\,\mathrm{d}x \\
&= \sqrt{5} - \int_0^1 \sqrt{1+4x^2}\,\mathrm{d}x + \int_0^1 \frac{1}{\sqrt{1+4x^2}}\,\mathrm{d}x
\end{aligned}$$

so

$$\begin{aligned}
2\int_0^1 \sqrt{1+4x^2}\,\mathrm{d}x &= \sqrt{5} + \int_0^1 \frac{1}{\sqrt{1+4x^2}}\,\mathrm{d}x \\
&= \sqrt{5} + \left[\tfrac{1}{2}\operatorname{arsinh} 2x\right]_{x=0}^1 \\
&= \sqrt{5} + \tfrac{1}{2}\ln(2+\sqrt{5}),
\end{aligned}$$

and dividing by two will now give the answer as before:

$$l = \tfrac{1}{2}\sqrt{5} + \tfrac{1}{4}\ln(2+\sqrt{5}).$$

6.2 Arc length of an ellipse

Let $P(a\cos\theta, b\sin\theta)$ be a point on the ellipse

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

Then

$$\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\theta}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}\theta}\right)^2} = \sqrt{(-a\sin\theta)^2 + (b\cos\theta)^2}$$

and so the integral that we need for the arc length has the form

$$\int \sqrt{a^2\sin^2\theta + b^2\cos^2\theta}\,\mathrm{d}\theta.$$

How hard can this be? Well, pretty tough, in fact. This is an example of what is called an *incomplete elliptic integral of the second kind* and there is no antiderivative with a simple closed form. (You could easily get a degree in Mathematics without ever coming across ways of trying to tackle such integrals.)

6.3 Arc length of a hyperbola

Surely this case is easier than that of an ellipse? After all, we know a lot about the function $y = \frac{1}{x}$ and we specifically created the natural logarithm in order to find the area under the curve so how hard can it be? So what is the arc length between $(1, 1)$ and $(2, \frac{1}{2})$?

$$y = \frac{1}{x} \Rightarrow \frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{1}{x^2} \Rightarrow \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2} = \sqrt{1 + \frac{1}{x^4}}$$

and so the arc length l is given by

$$l = \int_1^2 \sqrt{1 + \frac{1}{x^4}}\,\mathrm{d}x.$$

Although your calculator will give you a very good numerical approximation to this integral (mine gives 1.132 090 039 3 (FCD) in a matter of moments) this is another integral that simply cannot be done using your existing skills.

7 Polar coordinates

In order to get a neat expression for the polar equation of a conic section we make a change to the usual set up: consider the fixed point to be the origin $(0, 0)$ and the fixed line to be $x = -a$, as shown in Figure 15.

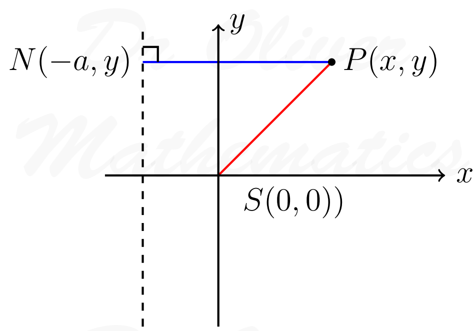

Figure 15: the standard set-up for the polar form of a conic section

Using the standard set up, if the conic has eccentricity e, then

$$\begin{aligned} PS = ePN &\Rightarrow r = e(r\cos\theta + a) \\ &\Rightarrow r = er\cos\theta + ae \\ &\Rightarrow r - er\cos\theta = ae \\ &\Rightarrow r(1 - e\cos\theta) = ae \\ &\Rightarrow r = \frac{ae}{1 - e\cos\theta}. \end{aligned}$$

Hence every equation of the form

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

where $e > 0$, $k > 0$ describes a conic section of eccentricity e and a focus at the origin. It is possible to recast the derivation show that

$$\begin{aligned} r &= \frac{k}{1 + e\cos\theta} \text{ (using } x = a\text{)}, \\ r &= \frac{k}{1 - e\sin\theta} \text{ (using } y = a\text{), and} \\ r &= \frac{k}{1 + e\sin\theta} \text{ (using } y = -a\text{)} \end{aligned}$$

also describe conic sections with a focus at the origin.

8 Appendix

8.1 Roots of a quadratic

Consider the quadratic equation $ax^2+bx+c = 0$ where $a \neq 0$. The solutions to this equation, x_1 and x_2, are given by

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \text{ and } x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}.$$

Then

$$x_1 + x_2 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} + \frac{-b - \sqrt{b^2 - 4ac}}{2a} = -\frac{b}{a}.$$

If $x = 3$ and $x = 4$ are the roots of a quadratic expression, then the factor theorem tells us that $(x - 3)$ and $(x - 4)$ are factors of the quadratic and so the quadratic must be a scalar mutliple of

$$(x - 3)(x - 4) = x^2 - (3 + 4)x + (3 \times 4) = x^2 - 7x + 12.$$

In general, α and β are roots of a quadratic equation if and only if the quadratic is a scalar multiple of

$$x^2 - (\alpha + \beta)x + \alpha\beta = 0;$$

so the sum of the roots is the negative of the coefficient of x and the product of the roots is the constant term (which is, if you recall, how you learned to factorise simple quadratics in Year 9).

8.2 Supremum

The third suggested definition of a diameter was

$$\sup \{|\mathbf{x} - \mathbf{y}| : \mathbf{x}, \mathbf{y} \in \text{circle}\}.$$

What is this?

The *supremum* function — sup — is used when we want to find the maximum value of a set of numbers but in circumstances where there may not be a maximum. For example, consider the set of numbers $\{x \in \mathbb{R} : 0 \leqslant x < 1\}$. What is the maximum value to be found in this set? Well, it is not 0.97, since 0.98 is bigger. It's not 0.999 999 since 0.999 999 999 is bigger. And it's certainly not $0.\dot{9} = 0.999\,999\,999\ldots$ since this is equal to 1 and hence is not a member of the set. So there is no maximum value.

The supremum gets us out of this difficulty. What it does is take the least upper bound that works. So, in this example, 16 is an upper bound for the maximum value (just not a very good one). 2 is a better upper bound. 1.04 is a better upper bound. 1 is an even better upper bound. But no number smaller than 1 can function as an upper bound. Hence

$$\max \{x \in \mathbb{R} : 0 \leqslant x < 1\} \text{ does not exist}$$

but

$$\sup \{x \in \mathbb{R} : 0 \leqslant x < 1\} = 1.$$

There is a corresponding function, called the *infimum* (inf), which acts as a greatest lower bound.

8.3 Closed-form expressions

In mathematics we break the types of expressions that we use into a variety of categories, depending on their complexity, whether they can be completed in finitely many step, involve limits, and so on. There is no universal agreement as to where the boundaries of these categories are drawn but Table 2 will give at least some sense of where plausible boundaries exist.

	Arithmetic expression	Polynomial expression	Algebraic expression	Closed-form expression	Analytic expression	Mathematical expression
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Variable	Yes	Yes	Yes	Yes	Yes	Yes
Elementary arithmetic operation	Yes	Yes	Yes	Yes	Yes	Yes
Factorial	Yes	Yes	Yes	Yes	Yes	Yes
Integer exponent	No	Yes	Yes	Yes	Yes	Yes
n^{th} root	No	No	Yes	Yes	Yes	Yes
Rational exponent	No	No	Yes	Yes	Yes	Yes
Irrational exponent	No	No	No	Yes	Yes	Yes
Logarithm	No	No	No	Yes	Yes	Yes
Trigonometric function	No	No	No	Yes	Yes	Yes
Inverse trigonometric function	No	No	No	Yes	Yes	Yes
Hyperbolic function	No	No	No	Yes	Yes	Yes
Inverse hyperbolic function	No	No	No	Yes	Yes	Yes
Gamma function	No	No	No	No	Yes	Yes
Bessel function	No	No	No	No	Yes	Yes
Special function	No	No	No	No	Yes	Yes
Continued fraction	No	No	No	No	Yes	Yes
Infinite series	No	No	No	No	Yes	Yes
Formal power series	No	No	No	No	No	Yes
Limit	No	No	No	No	No	Yes
Derivative	No	No	No	No	No	Yes
Integral	No	No	No	No	No	Yes

Table 2: a summary of expressions