Bachelor of Science

(B.Sc.)

INTRODUCTION TO GEOMETRY

Semester-II

Author- Dr. Harsh Vardhan Harsh

SUREASH GYAN VIHAR UNIVERSITY

Centre for Distance and Online Education
Mahal, Jagatpura, Jaipur-302017

Published by:

S. B. Prakashan Pvt. Ltd.

WZ-6, Lajwanti Garden, New Delhi: 110046

Tel.: (011) 28520627 | Ph.: 9205476295

Email: info@sbprakashan.com | Web.: www.sbprakashan.com

Designed & Graphic by : S. B. Prakashan Pvt. Ltd.

Printed at :

CONTENTS

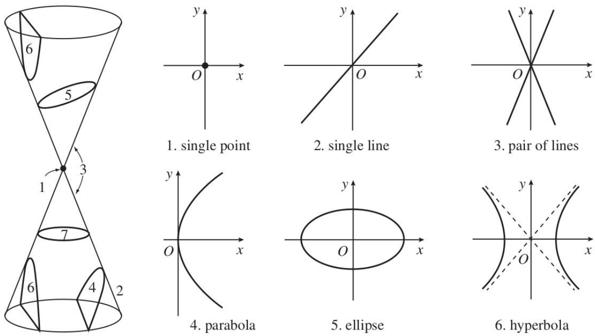
MODULE 1

CONICS

1.1 Conic Sections and Conics

1.1.1 Conic sections

Conic Section is the name given to the shapes that we obtain by taking different plane slices through a double cone. The shapes that we obtain from these cross sections are as drawn below.



1.1.2 Focus-Directrix Definition of the Non-Degenerate Conics

Earlier we defined the conic sections as the curves of intersection of a double cone with a plane. We have seen that the circle can be defined in a different way: as the set of points at a fixed distance from a fixed point. Here we give a method for constructing the other non-degenerate conics, the parabola, ellipse and hyperbola, as sets of points that satisfy a somewhat similar condition involving distances. Later we shall give a careful proof that each non-degenerate conic section is a non-degenerate (plane) conic, and vice-versa.

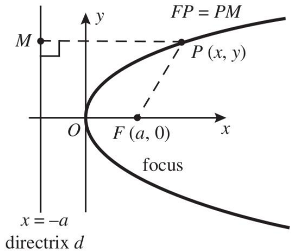
Eccentricity A non-degenerate conic is an ellipse if $0 \leq e < 1$, a parabola if $e = 1$, or a hyperbola if $e > 1$.

Parabola $(e = 1)$

A parabola is defined to be the set of points P in the plane whose distance from a fixed point F is equal to their distance from a fixed line d. We obtain a parabola in standard form if we choose

1. the focus F to lie on the x-axis, and to have coordinates $(a, 0), a > 0$;

2. the directrix d to be the line with equation $x = -a$.



> **Standard form of Parabola** A parabola in standard form has equation
>
> $$y^2 = 4ax, \qquad \text{where } a > 0$$
>
> It has focus $(a, 0)$ and directrix $x = -a$ and it can be described by the parametric equations
>
> $$x = at^2, \quad y = 2at \quad (t \in \mathbb{R}).$$

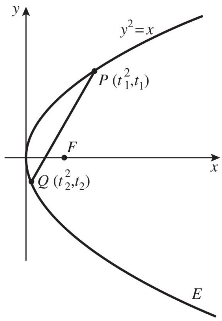
Example 1. Equation of a parabola E is $y^2 = 2x$ with parametric equations $x = \frac{1}{2}t^2, y = t (t \in \mathbb{R})$.

(a) Write down the focus, vertex, axis and directrix of E.

(b) Determine the equation of the chord that joins distinct points P and Q on E with parameters t_1 and t_2, respectively. Determine the condition on t_1 and t_2 such that the chord PQ passes through the focus of E.

Solution: (a) The parabola E is the parabola in standard form where $4a = 2$, or $a = \frac{1}{2}$.

It follows that the focus of E is $(\frac{1}{2}, 0)$, its vertex is $(0, 0)$, its axis is the $x-$ axis, and the equation of its directrix is $x = \frac{1}{2}$.



(b) The coordinates of P and Q are $(\frac{1}{2}t_1^2, t_1)$ and $(\frac{1}{2}t_2^2, t_2)$ respectively. So if $t_1^2 \neq t_2^2$. the slope (or gradient, as it is sometimes called) of PQ is given by

$$m = \frac{t_1 - t_2}{\frac{1}{2}t_1^2 - \frac{1}{2}t_2^2} = \frac{t_1 - t_2}{\frac{1}{2}[t_1^2 - t_2^2]} = \frac{2}{t_1 + t_2}$$

Since $(\frac{1}{2}t_1^2, t_1)$ lies on the line PQ, it follows that the equation of PQ is

$$y - t_1 = \frac{2}{t_1 + t_2}[x - \frac{1}{2}t_1^2].$$

Multiplying both sides by $t_1 + t_2$, we get

$$(t_1 + t_2)(y - t_1) = 2x - t_1^2,$$

So that

$$(t_1 + t_2)y - t_1^2 - t_1 t_2 = 2x - t_1^2$$

or

$$(t_1 + t_2)y = 2x + t_1 t_2. \tag{1}$$

If, however, $t_1^2 = t_2^2$, then since $t_1 \neq t_2$ we have $t_1 = -t_2$. Thus PQ parallel to the $y-$axis, and so has equation $x = \frac{1}{2}t_1^2$; so in this case too, PQ has equation given by (1).

The chord PQ with equation (1) passes through the focus $(\frac{1}{2}, 0)$ if $(t_1 + t_2)0 = 1 + t_1 t_2$ in other words, if $t_1 t_2 = -1$. $\square$

Problem 1. Consider the parabola E with equation $y^2 = x$ and parametric equations $x = t^2, y = t(t \in \mathbb{R})$.

(a) Write down the focus, vertex, axis and directrix of E.

(b) Determine the equation of the chord that joins distinct points P and Q on E with parameters t_1 and t_2, respectively.

(c) Determine the condition on t_1 and t_2 (and so on P and Q) that the focus of E is the midpoint of the chord PQ.

Solution:

(a) The parabola E is the parabola in standard form where $4a = 1$, or $a = \frac{1}{4}$. It follows that the focus of E is $\left(\frac{1}{4}, 0\right)$, the vertex is $(0, 0)$, the axis is the x -axis, and the equation of the directrix is $x = -\frac{1}{4}$.

(b) The coordinates of P and Q are $\left(t_1^2, t_1\right)$ and $\left(t_2^2, t_2\right)$, respectively. So, if $t_1^2 \neq t_2^2$, the slope of PQ is given by

$$\begin{aligned} m &= \frac{t_1 - t_2}{t_1^2 - t_2^2} \\ &= \frac{1}{t_1 + t_2} \end{aligned}$$

Since $\left(t_1^2, t_1\right)$ lies on the line PQ, it follows that the equation of PQ is

$$y - t_1 = \frac{1}{t_1 + t_2}\left(x - t_1^2\right).$$

Multiplying both sides by $t_1 + t_2$, we get

$$(t_1 + t_2)(y - t_1) = x - t_1^2$$

so that

$$(t_1 + t_2)\, y - t_1^2 - t_1 t_2 = x - t_1^2,$$

or

$$(t_1 + t_2)\, y = x + t_1 t_2 \qquad (*)$$

If, however, $t_1^2 = t_2^2$, then since $t_1 \neq t_2$ we have $t_1 = -t_2$. Thus PQ is parallel to the y -axis, and so has equation $x = t_1^2$; so in this case too, PQ has equation $(*)$.

(c) The midpoint of PQ is the point

$$\left(\frac{1}{2}\left(t_1^2 + t_2^2\right), \frac{1}{2}\left(t_1 + t_2\right)\right).$$

This is the focus $\left(\frac{1}{4}, 0\right)$ if

$$\left(\frac{1}{2}\left(t_1^2 + t_2^2\right), \frac{1}{2}\left(t_1 + t_2\right)\right) = \left(\frac{1}{4}, 0\right).$$

Comparing the second coordinates, we deduce that $t_2 = -t_1$. Comparing the first coordinates, we deduce that

$$\frac{1}{2}\left(t_1^2 + t_2^2\right) = \frac{1}{4}$$

so that $t_1^2 = \frac{1}{4}$. It follows that $t_1 = \pm\frac{1}{2}$, and so that $t_2 = \mp\frac{1}{2}$, respectively.

When $t = \frac{1}{2}$, the point $(t^2, t) = \left(\frac{1}{4}, \frac{1}{2}\right)$; and when $t = -\frac{1}{2}$, the point $(t^2, t) = \left(\frac{1}{4}, -\frac{1}{2}\right)$. It follows that the points P and Q must be $\left(\frac{1}{4}, \frac{1}{2}\right)$ and $\left(\frac{1}{4}, -\frac{1}{2}\right)$.

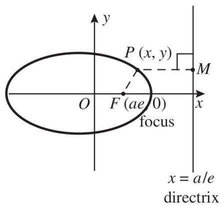
$\square$

Ellipse $(0 \leq e < 1)$

We define an ellipse with eccentricity zero to be a circle. Also we define an ellipse with eccentricity e (where $0 < e < 1$) to be the set of points P in the plane whose distance from a fixed point F is e times their distance from a fixed line d. We obtain such an ellipse in standard form if we choose

1. the focus F to lie on the x-axis, and to have coordinates $(ae, 0), a > 0$;

2. the directrix d to be the line with equation $x = a/e$.



Ellipse in Standard Form An ellipse in standard form has equation

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1, \quad \text{where } a \geq b > 0, b^2 = a^2(1-e^2), 0 \leq e < 1.$$

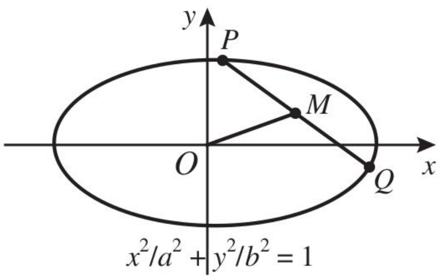
It can be described by the parametric equations $x = a\cos t, y = b\sin t, \quad (t \in (-\pi, \pi])$. If $e > 0$, it has foci $(\pm ae, 0)$ and directrices $x = \pm a/e$.

Example 2. Let PQ be an arbitrary chord of the ellipse with equation

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

Let M be the midpoint of PQ. Prove that the following expression is independent of the choice of P and Q :

$$\text{slope of } OM \times \text{slope of } PQ.$$



Solution: Let P and Q have the parametric coordinates $(acost_1, bsint_1)$ and $(acost_2, bsint_2)$, respectively. It follows that M has coordinates $(\frac{a}{2}(cost_1 + cost_2), \frac{b}{2}(sint_1 + sint_2))$. Now,

$$\text{the slope of } OM = \frac{b(sint_1 + sint_2)}{a(cost_1 + cost_2)}$$

and

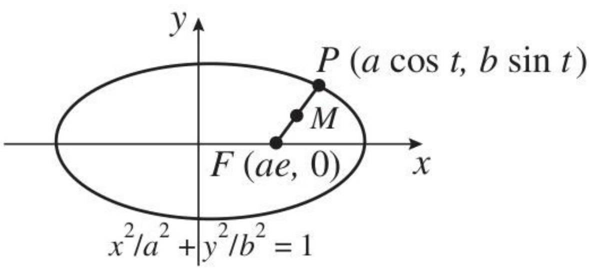
$$\text{the slope of } PQ = \frac{b(sint_1 - sint_2)}{a(cost_1 - cost_2)},$$

so,

$$\begin{aligned}
\text{slope of } OM \times \text{slope of } PQ &= \frac{b(sint_1 + sint_2)}{a(cost_1 + cost_2)} \cdot \frac{b(sint_1 - sint_2)}{a(cost_1 - cost_2)} \\
&= \frac{b^2}{a^2} \frac{(sin^2t_1 - sin^2t_2)}{(cos^2t_1 - cos^2t_2)} \\
&= \frac{b^2}{a^2} \frac{(sin^2t_1 - sin^2t_2)}{(1 - sin^2t_1) - (1 - sin^2t_2)} \\
&= \frac{-b^2}{a^2}.
\end{aligned}$$

which is independent of the values of t_1 and t_2. □

Problem 2. Let P be an arbitrary point on the ellipse with equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, and focus $F(ae, 0)$. Let M be the midpoint of FP. Prove that M lies on an ellipse whose centre is midway between the origin and F.



Solution: Let P have coordinates $(a\cos t, b\sin t)$. Since the coordinates of F are $(ae, 0)$, the coordinates of M, the midpoint of FP,is

$$\left(\frac{1}{2}(a\cos t + ae), \frac{1}{2}(b\sin t + 0)\right).$$

Thus M lies on the curve in $\mathbb{R}^2$ with parametric equations

$$x = \frac{1}{2}(a\cos t + ae), \quad y = \frac{1}{2}b\sin t$$

We can rearrange these equations in the form

$$\cos t = \frac{2x - ae}{a}, \quad \sin t = \frac{2y}{b};$$

squaring and adding these, we get

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

We can rearrange this equation in the form

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

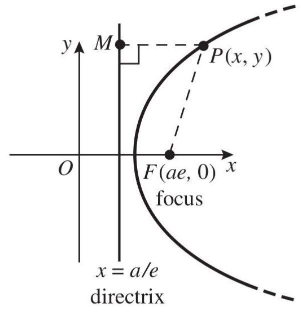
thus M lies on an ellipse with centre $\left(\frac{1}{2}ae, 0\right)$, the point midway between the origin and F. $\square$

Hyperbola $(e > 1)$

A hyperbola is the set of points P in the plane whose distance from a fixed point F, called focus of the hyperbola is e times their distance from a fixed line d,called directrix of the hyperbola, where $e > 1$. We obtain a hyperbola in standard form if we choose

1. the focus F to lie on the x -axis, and to have coordinates $(ae, 0), a > 0$;

2. the directrix d to be the line with equation $x = a/e$.



Let $P(x, y)$ be an arbitrary point on the hyperbola, and let M be the foot of the perpendicular from P to the directrix. Since $FP = e \cdot PM$, by the definition of the hyperbola, it follows that $FP^2 = e^2 \cdot PM^2$; we may rewrite this equation in terms of coordinates as

$$\begin{aligned}(x - ae)^2 + y^2 &= e^2 \left(x - \frac{a}{e}\right)^2 \\ &= (ex - a)^2\end{aligned}$$

Multiplying out the brackets we get

$$x^2 - 2aex + a^2e^2 + y^2 = e^2x^2 - 2aex + a^2$$

which simplifies to

$$x^2\left(e^2 - 1\right) - y^2 = a^2\left(e^2 - 1\right)$$

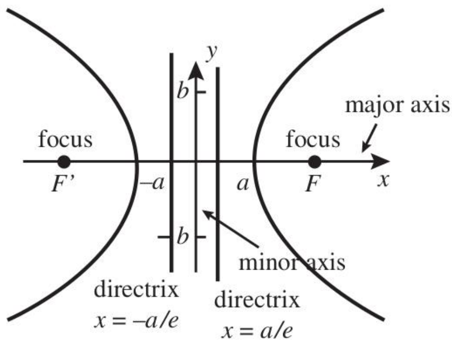
or

$$\frac{x^2}{a^2} - \frac{y^2}{a^2\left(e^2-1\right)} = 1.$$

Substituting b for $a\sqrt{e^2-1}$, so that $b^2 = a^2\left(e^2-1\right)$, we obtain the standard form of the equation of the hyperbola

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

Notice that this equation is symmetrical in x and symmetrical in y, so that the hyperbola also has a second focus $F'(-ae, 0)$ and a second directrix d' with equation $x = -a/e$.



The hyperbola intersects the x -axis at the points $(\pm a, 0)$. We call the segment joining the points $(\pm a, 0)$ the major axis or transverse axis of the hyperbola, and the segment joining the points $(0, \pm b)$ the minor axis or conjugate axis of the hyperbola (notice that this is NOT a chord of the hyperbola). The origin is the centre of this hyperbola.

Notice also that each point with coordinates $(a \sec t, b \tan t)$, where t is not an odd multiple of $\pi/2$, lies on the hyperbola, since

$$\frac{a^2 \sec^2 t}{a^2} - \frac{b^2 \tan^2 t}{b^2} = 1$$

Then, just as for the parabola, we can check that

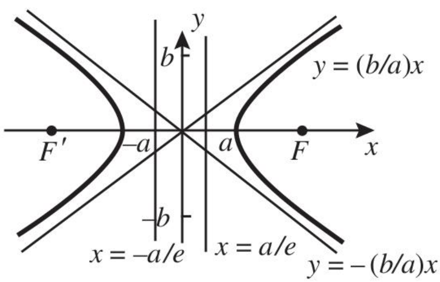
$$x = a \sec t, \quad y = b \tan t \quad (t \in (-\pi/2, \pi/2) \cup (\pi/2, 3\pi/2))$$

gives a parametric representation of the hyperbola.

Two other features of the shape of the hyperbola stand out. Firstly, the hyperbola consists of two separate curves or branches. Secondly, the lines with equations

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0, \quad \text{or} \quad y = \pm\frac{b}{a}x$$

divide the plane into two pairs of opposite sectors; the branches of the hyperbola lie in one pair. As $x \to \pm\infty$ the branches of the hyperbola get closer and closer to these two lines. We call the lines $y = \pm(b/a)x$ the asymptotes of the hyperbola.



We summarize the above facts as follows.

Hyperbola in Standard Form. A hyperbola in standard form has equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \qquad \text{where } b^2 = a^2\left(e^2 - 1\right), \qquad a > 0, e > 1$$

It has foci $(\pm ae, 0)$ and directrices $x = \pm a/e$; and it can be described by the parametric equations

$$x = a\sec t, \quad y = b\tan t \quad (t \in (-\pi/2, \pi/2) \cup (\pi/2, 3\pi/2)).$$

Problem 3. Let P be a point $\left(\sec t, \frac{1}{\sqrt{2}}\tan t\right)$, where $t \in (-\pi/2, \pi/2) \cup (\pi/2, 3\pi/2))$, on the hyperbola E with equation $x^2 - 2y^2 = 1$

(a) Determine the foci F and F' of E.

(b) Determine the slopes of FP and $F'P$, when these lines are not parallel to the y -axis.

(c) Determine the point P in the first quadrant on E for which FP is perpendicular to $F'P$.

Solution:

(a) This hyperbola is of the form $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ with $a = 1$ and $b^2 = \frac{1}{2}$, so that $b = 1/\sqrt{2}$.

If e denotes the eccentricity of the hyperbola E, so that

$b^2 = a^2(e^2 - 1)$, we have

$$\frac{1}{2} = e^2 - 1$$

it follows that $e^2 = \frac{3}{2}$ and so $e = \sqrt{\frac{3}{2}}$.

In general the foci are $(\pm ae, 0)$; it follows that here the foci are $\left(\pm\sqrt{\frac{3}{2}}, 0\right)$.

(b) Let F and F' be $\left(\sqrt{\frac{3}{2}}, 0\right)$ and $\left(-\sqrt{\frac{3}{2}}, 0\right)$, respectively. (It does not really matter which way round these are chosen.) Then the slope of FP is

$$\frac{\frac{1}{\sqrt{2}}\tan t - 0}{\sec t - \sqrt{\frac{3}{2}}} = \frac{\tan t}{\sqrt{2}\sec t - \sqrt{3}}$$

where we may assume that $\sec t \neq \sqrt{\frac{3}{2}}$, since FP is not parallel to the y -axis; and the slope of $F'P$ is

$$\frac{\frac{1}{\sqrt{2}}\tan t - 0}{\sec t + \sqrt{\frac{3}{2}}} = \frac{\tan t}{\sqrt{2}\sec t + \sqrt{3}}$$

where we may assume that $\sec t \neq -\sqrt{\frac{3}{2}}$, since $F'P$ is not parallel to the y -axis.

(c) When FP is perpendicular to $F'P$, we have that

$$\frac{\tan t}{\sqrt{2}\sec t - \sqrt{3}} \cdot \frac{\tan t}{\sqrt{2}\sec t + \sqrt{3}} = -1$$

We may rewrite this in the form

$$\frac{\tan^2 t}{2\sec^2 t - 3} = -1$$

so that $2\sec^2 t - 3 + \tan^2 t = 0$; since $\sec^2 t = 1 + \tan^2 t$, it follows that we must have $3\tan^2 t = 1$. Since we are looking for a point P in the first quadrant, we choose $\tan t = 1/\sqrt{3}$.

When $\tan t = 1/\sqrt{3}$, we have $\sec^2 t = 1 + \frac{1}{3} = \frac{4}{3}$. Since we are looking for a point P in the first quadrant, we choose $\sec t = 2/\sqrt{3}$.

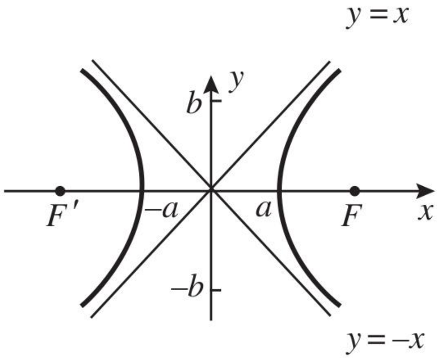
It follows that the required point P has coordinates

$$\left(\frac{2}{\sqrt{3}}, \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{3}}\right) = \left(\frac{2}{\sqrt{3}}, \frac{1}{\sqrt{6}}\right)$$

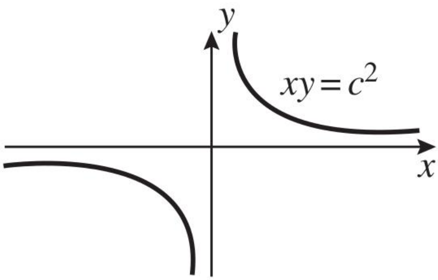
$\square$

Rectangular Hyperbola ($e = \sqrt{2}$)

When the eccentricity e of a hyperbola takes the value $\sqrt{2}$, then $e^2 = 2$ and $b = a$. Then the asymptotes of the hyperbola have equations $y = \pm x$, so that in particular they are at right angles. A hyperbola whose asymptotes are at right angles is called a rectangular hyperbola.



Then, if we use the asymptotes as new x - and y -axes (instead of the original x - and y -axes), it turns out that the equation of the hyperbola can be written in the form $xy = c^2$, for some positive number c.



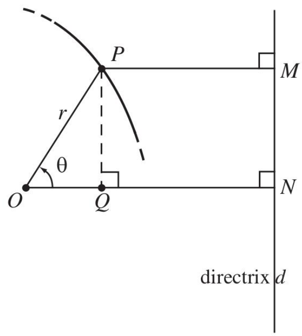
The rectangular hyperbola with equation $xy = c^2$ has the origin as its centre, and the x - and y -axes as its asymptotes. Also, each point on it can be uniquely represented by the parametric representation

$$x = ct, \quad y = \frac{c}{t} \quad \text{where } t \neq 0.$$

We shall use rectangular hyperbolas later on.

1.1.3 Polar Equation of a Conic

For many applications it is useful to describe the equation of a non-degenerate conic in terms of polar coordinates r and θ. A point $P(x, y)$ in the plane has polar coordinates (r, θ) if r is the distance OP (where O is the origin) and θ is the anticlockwise angle between OP and the positive direction of the x -axis.



Take the origin O to be the focus of the conic, d the directrix, M the foot of the perpendicular from a point P on the conic to d, N the foot of the perpendicular from O to d, and Q the foot of the perpendicular from P to ON.

Then by the definition of the conic, we have $OP = e \cdot PM$.

We can rewrite this as

$$\begin{aligned} r &= e(ON - OQ) \\ &= e \cdot ON - er\cos\theta \\ &\quad or \\ r(1 + e\cos\theta) &= e \cdot ON \\ &= l, \qquad \text{a constant.} \end{aligned}$$

It follows that the equation of the conic can be expressed in the form

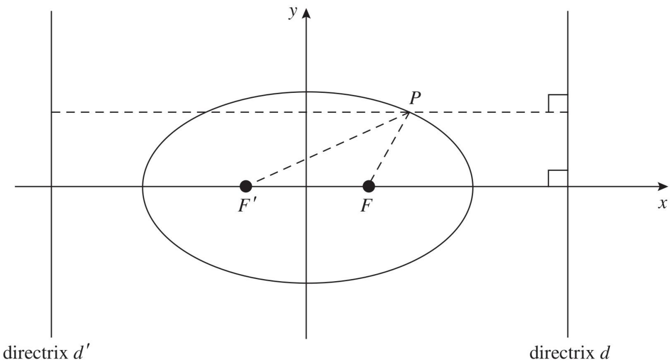
$$r = \frac{l}{1 + e\cos\theta}.$$

The polar form of the equation of a conic is often used in problems in Dynamics: for example, in determining the motion of a planet or of a comet round the Sun.

1.1.4 Focal Distance Properties of Ellipse and Hyperbola

We now prove two simple but surprising results. We deal with the ellipse first.

Theorem 1. (Sum of Focal Distances of Ellipse) Let E be an ellipse with major axis $(-a, a)$ and foci F and F'. Then, if P is a point on the ellipse, $FP + PF' = 2a$. In particular, $FP + PF'$ is constant for all points P on the ellipse.



Proof: Let d and d' be the directrices of the ellipse that correspond to the foci F and F', respectively. Then, since

$$PF = e \times \text{ (distance from } P \text{ to } d)$$

and

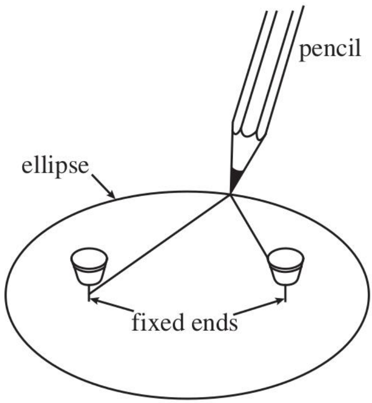
$$PF' = e \times \text{ (distance from } P \text{ to } d'),$$

it follows that

$$\begin{aligned} PF + PF' &= e \times (\text{ distance between } d \text{ and } d')\,. \\ &= 2a \end{aligned}$$

which is a constant. $\square$

The result of Theorem 5 can be used to draw an ellipse, using a piece of string fixed at both ends. A pencil is used to pull the string taut; then, as we move the pencil round, the shape that it traces out is an ellipse whose foci are the two ends of the string.



Notice that, if we are given any three points F, F' and P (not on the line segment $F'F$) in the plane, then there is only one ellipse through P with F and F' as its foci. Its centre is the midpoint, O, of the segment $F'F$, its axes are the line along $F'F$ and the line through O perpendicular to $F'F$, and its major axis has length $PF + PF'$.

Also, if we are given any two points F and F' in the plane, the locus of points P (not on the line segment $F'F$) in the plane for which $PF+PF'$ is a constant is necessarily an ellipse. Thus the converse of Theorem 5 holds.

There is an analogous result for the hyperbola.

Theorem 2. (Difference of Focal Distances of Hyperbola) Let H be a hyperbola with major axis $(-a, a)$

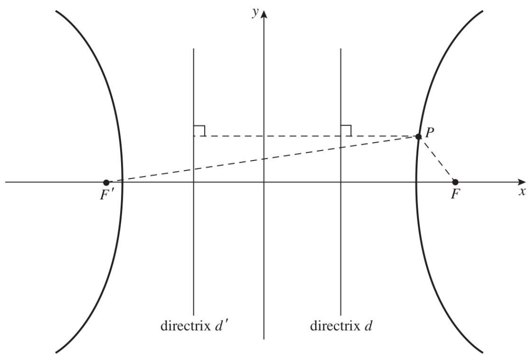
and foci F and F'. Then, if P is a point on the branch of the hyperbola that is closer to F,

$$PF' - PF = 2a;$$

and, if P is a point on the branch of the hyperbola closer to F',

$$PF' - PF = -2a.$$

In particular, $|PF' - PF|$ is constant for all points P on the hyperbola.



Proof: We shall prove only the first formula; the proof of the second is similar. Let d and d' be the directrices of the hyperbola that correspond to the foci F and F' respectively, and let P be a point on the branch of the hyperbola that is closer to F. Then, since

$$PF = e \times (\text{ distance from } P \text{ to } d)$$

and

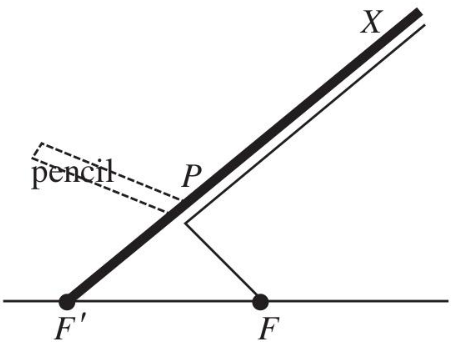
$$PF' = e \times (\text{ distance from } P \text{ to } d')$$

it follows that

$$\begin{aligned} PF' - PF &= e \times (\text{ distance between } d \text{ and } d') \\ &= 2a, \end{aligned}$$

which is a constant. □

The result of Theorem 6 can be used to draw a hyperbola, this time using piece of string and a stick. Choose two points F and F' on the x -axis, equidistant from and on opposite sides of the origin. Hinge one end of a movable stick $F'X$ at the focus F'; attach one end of a string of length ℓ (where ℓ is less than the length of $F'X$) to the end X of the stick and the other end of the string to F, and keep the string taut by holding a pencil tight against the stick, as shown.



Then, as we move the pencil along the stick, the shape that it traces out is part of one branch of a hyperbola with foci F

and F'. For,

$$\begin{aligned} PF' - PF &= XF' - (XP + PF) \\ &= XF' - \ell \\ &= \text{ a constant independent of } P \end{aligned}$$

We obtain the other branch of the hyperbola by interchanging the roles of F and F' in the construction.

Notice that, if we are given any three points F, F' and P (not on the line through $F'F$ or its perpendicular bisector) in the plane, then there is only one hyperbola through P with F and F' as its foci. Its centre is the midpoint, O, of the segment $F'F$, its axes are the line along $F'F$ and the line through O perpendicular to $F'F$, and its major axis has length $|PF' - PF|$.

Also, if we are given any two points F and F' in the plane, the locus of points P (not on the line segment $F'F$) in the plane for which $PF' - PF$ is a non-zero constant is necessarily one branch of a hyperbola. Thus the converse of Theorem 6 holds, in the following sense: Given any three points F, F' and P (where P must lie strictly between F and F' if it lies on the line through $F'F$) in the plane for which $PF' - PF \neq 0$, the locus of points Q in the plane for which $QF' - QF = \pm\, |PF' - PF|$ is a hyperbola.

1.2 Properties of Conics

1.2.1 Tangents

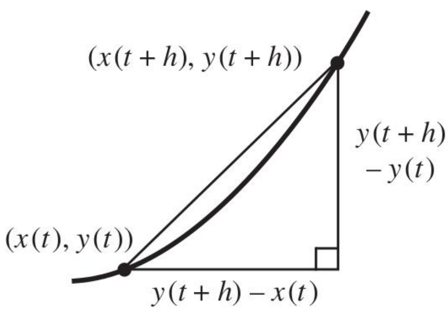
In the previous section you met the parametric equations of the parabola, ellipse and hyperbola in standard form.

We now tackle a rather natural question: given parametric equations $x = x(t), y = y(t)$ describing a curve, what is the slope of the tangent to the curve at the point with parameter t? This information will enable us to determine the equation of the tangent to the curve at that point.

Theorem 1. The slope of the tangent to a curve in $\mathbb{R}^2$ with parametric equations $x = x(t), y = y(t)$ at the point with parameter t is

$$\frac{y'(t)}{x'(t)}$$

provided that $x'(t) \neq 0$.



Proof: The points on the curve with parameters t and $t + h$

have coordinates $(x(t), y(t))$ and $(x(t+h), y(t+h))$, respectively. Then, if $h \neq 0$, the slope of the chord joining these two points is

$$\frac{y(t+h) - y(t)}{x(t+h) - x(t)}$$

which we can write in the form

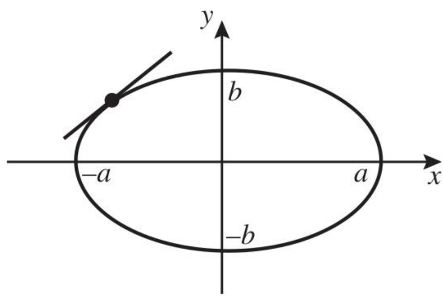
$$\frac{(y(t+h) - y(t))/h}{(x(t+h) - x(t))/h}$$

We then take the limit of this ratio as $h \to 0$. The slope of the chord tends to the slope of the tangent, namely $y'(t)/x'(t)$. $\square$

Example 1. (a) Determine the equation of the tangent at the point with parameter t to the ellipse with parametric equations

$$x = a\cos t, \quad y = b\sin t,$$

where $t \in (-\pi, \pi], t \neq 0, \pi$



(b) Hence determine the equation of the tangent to the ellipse with parametric equations $x = 3\cos t, y = \sin t$ at the point with parameter $t = \pi/4$ Deduce the coordinates of the point of intersection of this tangent with the x -axis.

Solution:

(a) Now, $y'(t) = b\cos t$ and $x'(t) = -a\sin t$ for $t \in (-\pi, \pi]$; it follows that, for $t \neq 0$ or π, the slope of the tangent at the point with parameter t is

$$\frac{y'(t)}{x'(t)} = \frac{b\cos t}{-a\sin t}$$

Hence the equation of the tangent at the point $(a\cos t, b\sin t), t \neq 0\ \pi$, is

$$y - b\sin t = -\frac{b\cos t}{a\sin t}(x - a\cos t)$$

Multiplying both sides and rearranging terms, we get

$$xb\cos t + ya\sin t = ab\cos^2 t + ab\sin^2 t = ab$$

and dividing both sides by ab gives the equation

$$\frac{x}{a}\cos t + \frac{y}{b}\sin t = 1 \tag{1}$$

The point on the ellipse where $t = 0$ is $(a, 0)$, at which the tangent has equation $x = a$. Similarly, the point on the ellipse where $t = \pi$ is $(-a, 0)$, at which the tangent has equation $x = -a$. It follows that equation (1) covers these cases also.

(b) Here the curve is the ellipse in part (a) in the particular case that $a = 3$, $b = 1$. When $t = \pi/4$, it follows from equation (1) that the equation of the tangent at the point with parameter $t = \pi/4$ is

$$\frac{x}{3} \cdot \frac{1}{\sqrt{2}} + y \cdot \frac{1}{\sqrt{2}} = 1 \ ,$$

or

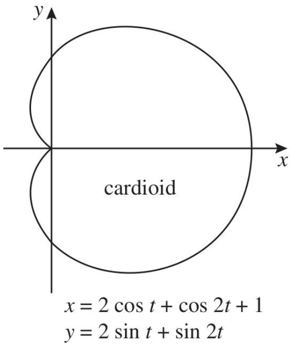
$$\frac{1}{3}x + y = \sqrt{2}$$

Hence, at the point T where the tangent crosses the x -axis, $y = 0$ and so $x = 3\sqrt{2}$. Thus, T is the point $(3\sqrt{2}, 0)$. □

Problem 1. Determine the slope of the tangent to the curve in $\mathbb{R}^2$ with parametric equations

$$x = 2\cos t + \cos 2t + 1, \quad y = 2\sin t + \sin 2t$$

at the point with parameter t, where t is not a multiple of π. Hence determine the equation of the tangent to this curve at the points with parameters $t = \pi/3$ and $t = \pi/2$.



Solution: We use the formula of Theorem 1 to find the slope of the tangent to the curve at the point with parameter t, where t is not a multiple of π. Since

$$\begin{aligned} x(t) &= 2\cos t + \cos 2t + 1 \qquad \text{and} \\ y(t) &= 2\sin t + \sin 2t \end{aligned}$$

we have

$$\begin{aligned} x'(t) &= -2\sin t - 2\sin 2t \qquad \text{and} \\ y'(t) &= 2\cos t + 2\cos 2t \end{aligned}$$

Hence the slope of the curve at this point is

$$\begin{aligned} \frac{y'(t)}{x'(t)} &= \frac{2\cos t + 2\cos 2t}{-2\sin t - 2\sin 2t} \\ &= -\frac{\cos t + \cos 2t}{\sin t + \sin 2t} \end{aligned}$$

In particular, at the point with parameter $t = \pi/3$, this slope is

$$-\frac{\cos \pi/3 + \cos 2\pi/3}{\sin \pi/3 + \sin 2\pi/3} = -\frac{\frac{1}{2} - \frac{1}{2}}{\sqrt{3}/2 + \sqrt{3}/2} = 0$$

Thus the tangent at this point is horizontal. Also,

$$\begin{aligned} y(\pi/3) &= 2\sin(\pi/3) + \sin(2\pi/3) \\ &= 2 \cdot \frac{\sqrt{3}}{2} + \frac{\sqrt{3}}{2} = \frac{3}{2}\sqrt{3} \end{aligned}$$

It follows that the equation of the tangent at the point with parameter $t = \pi/3$ is $y = \frac{3}{2}\sqrt{3}$. Next, at the point with parameter $t = \pi/2$, the slope of the curve is

$$-\frac{\cos\pi/2 + \cos\pi}{\sin\pi/2 + \sin\pi} = -\frac{0-1}{1+0} = 1.$$

Also,

$$x(\pi/2) = 2\cos(\pi/2) + \cos(\pi) + 1 = -1 + 1 = 0$$

and

$$y(\pi/2) = 2\sin(\pi/2) + \sin(\pi) = 2 + 0 = 2$$

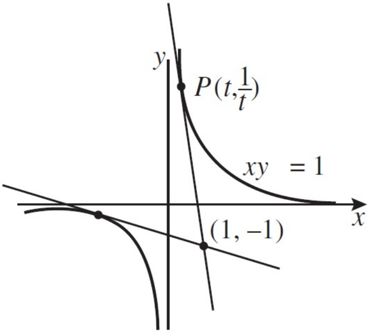
It follows that the equation of the tangent at the point with parameter $t = \pi/2$ is

$$y - 2 = 1(x - 0), \quad \textit{or} \quad y = x + 2$$

□

Problem 2. (a) Determine the equation of the tangent at a point P with parameter t on the rectangular hyperbola with parametric equations $x = t$, $y = 1/t$

(b) Hence determine the equations of the two tangents to the rectangular hyperbola from the point $(1, -1)$



Solution:

(a) Here $x'(t) = 1$ and $y'(t) = -1/t^2$; it follows that the slope of the tangent at the point with parameter t is

$$\begin{aligned}\frac{y'(t)}{x'(t)} &= \frac{-1/t^2}{1}\\ &= -\frac{1}{t^2}\end{aligned}$$

It follows that the equation of the tangent at the point P is

$$y - \frac{1}{t} = -\frac{1}{t^2}(x - t)$$

or

$$y = -\frac{x}{t^2} + \frac{2}{t}\ .$$

(b) The line with equation $y = -\frac{x}{t^2} + \frac{2}{t}$ passes through the point $(1, -1)$ if

$$-1 = -\frac{1}{t^2} + \frac{2}{t}$$

We can rewrite this equation in the form

$$t^2 + 2t - 1 = 0$$

or

$$(t+1)^2 = 2$$

it follows that the values of t at the two points on the hyperbola for which the tangents pass through $(1, -1)$ are

$$t = -1 \pm \sqrt{2}$$

When $t = -1 + \sqrt{2}$, the equation of the tangent is

$$\begin{aligned} y &= -\frac{x}{(\sqrt{2}-1)^2} + \frac{2}{\sqrt{2}-1} \\ &= -\frac{x}{3 - 2\sqrt{2}} + \frac{2}{\sqrt{2}-1} \end{aligned}$$

When $t = -1 - \sqrt{2}$, the equation of the tangent is

$$\begin{aligned} y &= -\frac{x}{(-1-\sqrt{2})^2} + \frac{2}{-1-\sqrt{2}} \\ &= -\frac{x}{3 + 2\sqrt{2}} - \frac{2}{1+\sqrt{2}} \end{aligned}$$

$\square$

We can modify the result of Example 1(a) to find the equation of the tangent at the point (x_1, y_1) on the ellipse with equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. We take $x = a\cos t, y = b\sin t$ as parametric equations for the ellipse, and let $x_1 = a\cos t_1$ and $y_1 = b\sin t_1$. Then it follows from equation (1) above that the equation of the tangent is

$$\frac{x}{a}\cos t_1 + \frac{y}{b}\sin t_1 = 1$$

which we can rewrite in the form $\frac{xx_1}{a^2} + \frac{yy_1}{b^2} = 1$.

We can determine the equations of tangents to the hyperbola and the parabola in a similar way; the results are given in the following theorem.

Theorem 2. The equation of the tangent at the point (x_1, y_1) to a conic in standard form is as follows.

Conic	Tangent
Ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$	$\frac{xx_1}{a^2} + \frac{yy_1}{b^2} = 1$
Hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$	$\frac{xx_1}{a^2} - \frac{yy_1}{b^2} = 1$
Parabola $y^2 = 4ax$	$yy_1 = 2a(x + x_1)$

Problem 3. Prove that the equation of the tangent at the point (x_1, y_1) to the rectangular hyperbola $xy = 1$ is $\frac{1}{2}(xy_1 + x_1y) = 1$

Solution: The rectangular hyperbola $xy = 1$ has parametric equations $x = t, y = 1/t$ (where $t \neq 0$). You found in Problem 2 (a) that the slope of the tangent at the point with parameter

t is

$$\frac{y'(t)}{x'(t)} = -\frac{1}{t^2}$$

Since $-\frac{1}{t^2} = -\frac{y_1}{x_1}$, the slope of the tangent at the point (x_1, y_1) may be written in a convenient form as $-\frac{y_1}{x_1}$. (The slope may be expressed in many other forms involving x_1 and y_1, but this particular form saves some algebra later in the calculation.)

Then the equation of the tangent to the hyperbola $xy = 1$ at the point (x_1, y_1) is

$$y - y_1 = -\frac{y_1}{x_1}(x - x_1)$$

Multiplying both sides by x_1, we may express this as

$$x_1 y - x_1 y_1 = -x y_1 + x_1 y_1$$

so that

$$\begin{aligned} x_1 y + x y_1 &= 2x_1 y_1 \\ &= 2 \end{aligned}$$

dividing this by 2 , we obtain the required equation. $\square$

Problem 4. For each of the following conics, determine the equation of the tangent to the conic at the indicated point.

(a) The unit circle $x^2 + y^2 = 1$ at $\left(-\frac{1}{2}, \frac{1}{2}\sqrt{3}\right)$.

(b) The hyperbola $xy = 1$ at $\left(-4, -\frac{1}{4}\right)$.

(c) The parabola $y^2 = x$ at $(1, -1)$.

Solution: The required equations may be obtained by simply substituting numbers into the appropriate equation in Theorem 2 or Problem 3 .

(a) The equation of the tangent to the unit circle $x^2 + y^2 = 1$ at $\left(-\frac{1}{2}, \frac{1}{2}\sqrt{3}\right)$ is

$$x\left(-\frac{1}{2}\right) + y\left(\frac{1}{2}\sqrt{3}\right) = 1,$$

which may be written in the form

$$\sqrt{3}y = x + 2 \ .$$

(b) The equation of the tangent to the rectangular hyperbola $xy = 1$ at $\left(-4, -\frac{1}{4}\right)$ is

$$\frac{1}{2}\left(x\left(-\frac{1}{4}\right) - 4y\right) = 1,$$

which may be written in the form $x + 16y = -8$

(c) The equation of the tangent to the parabola $y^2 = x$ at $(1, -1)$ is

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

which may be written in the form

$$x + 2y = -1$$

$\square$

We can deduce a useful fact from the equation $xx_1 + yy_1 = 1$ for the tangent at the point (x_1, y_1) to the unit circle $x^2 + y^2 = 1$. Let (a, b) be some point on this tangent, so that

$$ax_1 + by_1 = 1 \tag{2}$$

Next, let the other tangent to the unit circle through the point (a, b) touch the circle at the point (x_2, y_2); it follows that

$$ax_2 + by_2 = 1 \tag{3}$$

From equations (2) and (3) we deduce that the points (x_1, y_1) and (x_2, y_2) both satisfy the equation $ax + by = 1$. Since this is the equation of a line, it must be the equation of the line through the points (x_1, y_1) and (x_2, y_2). For historical reasons, this line is called *polar of* (a, b) *with respect to the unit circle.*

Theorem 3. Let (a, b) be a point outside the unit circle, and let the tangents to the circle from (a, b) touch the circle at P_1 and P_2. Then the equation of the line through P_1 and P_2 is

$$ax + by = 1$$

For example, the polar of $(2, 0)$ with respect to the unit circle is the line $2x = 1$

Problem 5. Determine the equation of the polar of the point

$(2, 3)$ with respect to the unit circle.

Solution: Since $2^2 + 3^2 > 1$, the point $(2, 3)$ lies outside the unit circle. Hence, by Theorem 3 , the polar of the point $(2, 3)$ with respect to the unit circle has the equation

$$2x + 3y = 1$$

□

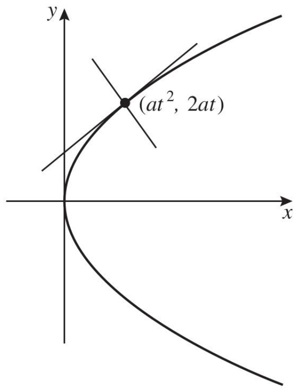
In the next example we meet the idea of the normal to a curve.

Definition. The normal to a curve C at a point P on C is the line through P that is perpendicular to the tangent to C at P.

Example 2. (a) Determine the equation of the tangent at the point with parameter t to the parabola with parametric equations

$$x = at^2, \quad y = 2at \quad (t \in \mathbb{R})$$

(b) Hence determine the equations of the tangent and the normal to the parabola with parametric equations $x = 2t^2, y = 4t$ at the point with parameter $t = 3$.



Solution:

(a) Since $y'(t) = 2a$ and $x'(t) = 2at$, it follows that, for $t \neq 0$, the slope of the tangent at this point is

$$\frac{y'(t)}{x'(t)} = \frac{2a}{2at} = \frac{1}{t}$$

Hence the equation of the tangent at the point $(at^2, 2at)$, $t \neq 0$, is

$$y - 2at = \frac{1}{t}\left(x - at^2\right)$$

which can be rearranged in the form

$$ty = x + at^2 \tag{4}$$

The point on the parabola at which $t = 0$ is $(0, 0)$; there the tangent to the parabola is the y -axis, with equation $x = 0$. It follows that equation (4) covers this case also.

(b) Here the curve is the parabola in part (a) in the particular case that $a = 2$. When $t = 3$, it follows from equation (4) that the equation of the tangent is $3y = x + 2 \cdot 3^2$, or $3y = x + 18$.

To find the equation of the normal, we must find its slope and the coordinates of the point on the parabola at which $t = 3$.

When $t = 3$, it follows from the equation of the tangent that the slope of the tangent is $\frac{1}{3}$. Since the tangent and normal are perpendicular to each other, it follows that the slope of the normal must be -3. Also, when $t = 3$, we have that $x = 2 \cdot 3^2 = 18$ and $y = 4 \cdot 3 = 12$; so the corresponding point on the parabola has coordinates $(18, 12)$.

It follows that the equation of the normal to the parabola at the point $(18, 12)$ is

$$\begin{aligned} y - 12 &= -3(x - 18) \\ &= -3x + 54 \end{aligned}$$

or

$$y = -3x + 66$$

□

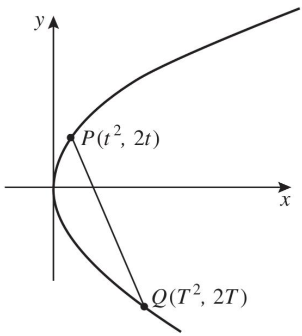
Problem 6. The normal to the parabola with parametric equations $x = t^2, y = 2t (t \in \mathbb{R})$ at the point P with parameter

$t, t \neq 0$, meets the parabola at a second point Q with parameter T.

(a) Prove that the slope of the normal to the parabola at P is $-t$.

(b) Find the equation of the normal to the parabola at P

(c) By substituting the coordinates of Q into your equation from part (b), prove that $T = -\frac{2}{l} - t$.



Solution:

(a) We saw in Example 2 (a) that the slope of the tangent at the point P with parameter t (where $t \neq 0$) is $1/t$. Since the normal and the tangent at P are perpendicular to each other, it follows that the slope m of the normal at P must satisfy the equation $m \cdot (1/t) = -1$. Hence $m = -t$.

(b) The normal at P is thus the line through the point $(t^2, 2t)$ with slope $-t$, and so has equation

$$y - 2t = -t\left(x - t^2\right),$$

or

$$y = -tx + 2t + t^3 \qquad (*)$$

(c) Let Q be the point on the parabola with parameter T, say; thus its coordinates are $(T^2, 2T)$. Since Q lies on the line with equation $(*)$, it follows that

$$2T = -tT^2 + 2t + t^3$$

we can rearrange this equation in the form

$$2(T - t) = -t\left(T^2 - t^2\right)$$

Since $T \neq t$, we may divide through by $T - t$, to get

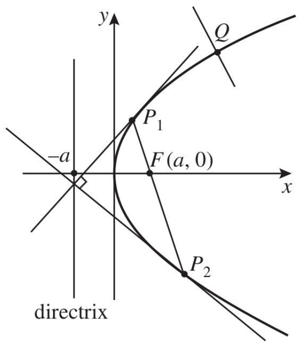
$$\begin{aligned} 2 &= -t(T + t) \\ &= -tT - t^2 \end{aligned}$$

so that $tT = -2 - t^2$; it follows that $T = -\frac{2}{t} - t$

□

Problem 7. This question concerns the parabola with parametric equations $x = at^2, y = 2at (t \in \mathbb{R})$

(a) Determine the equation of the chord joining the points P_1 and P_2 on the parabola with parameters t_1 and t_2, respectively, where t_1 and t_2 are unequal and non-zero.



Now assume that the chord P_1P_2 passes through the focus $(a, 0)$ of the parabola.

(b) Prove that $t_1 t_2 = -1$.

(c) Use the result of Example 2 (a) to write down the equations of the tangents to the parabola at P_1 and P_2, and to prove that these tangents are perpendicular.

(d) Find the point of intersection P of the two tangents in part (c), and verify that it lies on the directrix $x = -a$ of the parabola.

(e) Find the equation of the normal at the point $Q\,(at^2, 2at)$ to the parabola. Hence prove that if the normal at Q passes through the focus $F(a, 0)$, then Q is the vertex of the parabola.

Solution:

(a) P_1 has coordinates $\left(at_1^2, 2at_1\right)$ and P_2 has coordinates $\left(at_2^2, 2at_2\right)$. So, if $t_1^2 \neq t_2^2$ the slope of the chord P_1P_2 is

$$\begin{aligned} \frac{2at_2 - 2at_1}{at_2^2 - at_1^2} &= 2\frac{t_2 - t_1}{t_2^2 - t_1^2} \\ &= \frac{2}{t_2 + t_1} \end{aligned}$$

It follows that the equation of P_1P_2 is

$$y - 2at_1 = \frac{2}{t_1 + t_2}\left(x - at_1^2\right)$$

or

$$(t_1 + t_2)(y - 2at_1) = 2\left(x - at_1^2\right). \qquad (**)$$

If, however, $t_1^2 = t_2^2$ we must have $t_1 = -t_2$ since P_1 and P_2 are distinct. The chord P_1P_2 is then parallel to the y -axis, so that we can write its equation as $x = at_1^2$. Thus the equation of the chord is given by equation $(**)$ in this case too.

(b) If the chord P_1P_2 passes through the focus $(a, 0)$, the coordinates of $(a, 0)$ must satisfy equation (2); hence

$$(t_1 + t_2)(-2at_1) = 2\left(a - at_1^2\right),$$

so that

$$-t_1^2 - t_1 t_2 = 1 - t_1^2$$

It follows that $t_1 t_2 = -1$.

(c) It follows from Example 2(a) that the equations of the tangents at P_1 and P_2 are

$$t_1 y = x + a t_1^2$$

and

$$t_2 y = x + a t_2^2$$

respectively.

Now it follows also from part (a) of Example 2 that the slopes of the tangents at P_1 and P_2 are $1/t_1$ and $1/t_2$, respectively. These tangents are perpendicular if

$$\left(\frac{1}{t_1}\right) \cdot \left(\frac{1}{t_2}\right) = -1$$

and we can rewrite this condition in the form $t_1 t_2 = -1$.

We have already seen in part (b) that $t_1 t_2 = -1$, and so we deduce that the tangents at P_1 and P_2 are indeed perpendicular.

(d) The equations of the tangents at P_1 and P_2 are

$$t_1 y = x + a t_1^2 \quad \textit{and} \quad t_2 y = x + a t_2^2$$

respectively. By subtracting these, we see that at the point (x, y) of intersection,

$$(t_1 - t_2)\, y = a\left(t_1^2 - t_2^2\right)$$

so that

$$y = a\,(t_1 + t_2) \ .$$

It then follows from the equation $t_1 y = x + at_1^2$ that

$$t_1 a\,(t_1 + t_2) = x + at_1^2$$

so that

$$\begin{aligned} x &= at_1 t_2 \\ &= -a \quad (since\ t_1 t_2 = -1) \end{aligned}$$

The point of intersection is therefore $(-a, a\,(t_1 + t_2))$.

Since the first coordinate of the point of intersection is $-a$, it follows that the point of intersection lies on the directrix of the parabola.

(e) Since (by the result of Example 2 (a)) the tangent at Q has slope $1/t$, when $t \neq 0$, it follows that in this case the normal at Q has slope $-t$. When $t = 0$, the point Q is the origin, the vertex of the parabola; so in this case too the slope of the normal is $-t$.

Hence in general the equation of the normal at Q is

$$y - 2at = -t\left(x - at^2\right) \qquad (***)$$

If this normal passes through $F(a, 0)$, then the coordinates of F must satisfy equation $(***)$; that is,

$$-2at = -t\left(a - at^2\right)$$

We can divide through by a and then rearrange the terms in this equation to get

$$0 = t\left(1 + t^2\right)$$

It follows that $t = 0$, and so Q must be the vertex of the parabola.

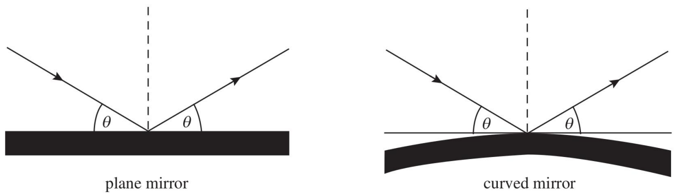
□

1.2.2 Reflections

We use the reflection properties of mirrors all the time. For example, we look in plane mirrors while shaving or combing our hair, and we use electric fires with reflecting rear surfaces to throw radiant heat out into a room.

All reflecting surfaces - mirrors, for example - obey the same Reflection Law. The Reflection Law is often expressed in terms of the angles made with the normal to the surface rather than the surface itself. However in this section we shall state and use it in the following form.

The Reflection Law The angle that incoming light makes with the tangent to a surface is the same as the angle that the reflected light makes with the tangent.



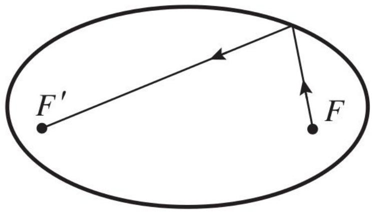
This law applies to all mirrors, no matter whether the reflecting surface is plane or curved. Indeed, in many practical applications the mirror is designed to have a cross-section that is a conic curve - for example, the Lovell radiotelescope at Jodrell Bank in Cheshire, England uses a parabolic reflector to focus parallel radio waves from space onto a receiver.

We now investigate the reflection properties of mirrors in the shape of the non-degenerate conics.

Reflection Property of the Ellipse

We start with the following interesting property of the ellipse.

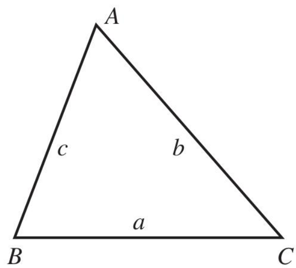
Reflection Property of the Ellipse Light which comes from one focus of an elliptical mirror is reflected at the ellipse to pass through the second focus.



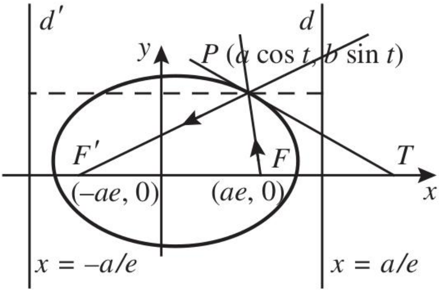
In our proof we use the following trigonometric result for triangles.

Sine Formula In a triangle $\triangle ABC$ with sides a, b, c opposite the vertices A, B, C, respectively,

$$\frac{a}{\sin \angle BAC} = \frac{b}{\sin \angle ABC} = \frac{c}{\sin \angle ACB}$$



Proof of Reflection Property Let E be the ellipse in standard form, and $P(a\cos t, b\sin t)$ an arbitrary point on E; for simplicity, we shall assume that P lies in the first quadrant.



Then, as we saw earlier,

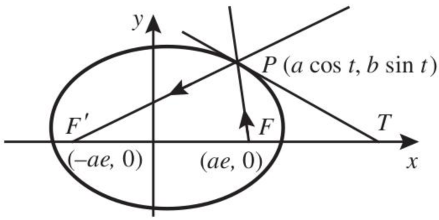
$$\begin{aligned} PF &= e \times \text{ (distance from } P \text{ to corresponding directrix } d) \\ &= e \times \left(\frac{a}{e} - a\cos t\right) = a - ae\cos t \end{aligned}$$

and

$$\begin{aligned} PF' &= e \times (\text{ distance from } P \text{ to } d') \\ &= e \times \left(\frac{a}{e} + a\cos t\right) = a + ae\cos t \end{aligned}$$

Hence,

$$\frac{PF}{PF'} = \frac{a - ae\cos t}{a + ae\cos t} = \frac{1 - e\cos t}{1 + e\cos t}.$$



Next, we saw earlier that the equation of the tangent at P

to the ellipse is

$$\frac{x}{a}\cos t + \frac{y}{b}\sin t = 1$$

hence at the point T where the tangent at P intersects the x -axis, we have

$$\frac{x}{a}\cos t = 1, \qquad \text{or} \qquad x = a/\cos t$$

It follows that

$$\frac{TF}{TF'} = \frac{(a/\cos t) - ae}{(a/\cos t) + ae} = \frac{1 - e\cos t}{1 + e\cos t}.$$

We deduce that

$$\frac{PF}{PF'} = \frac{TF}{TF'}, \qquad \text{or} \qquad \frac{PF}{TF} = \frac{PF'}{TF'}.$$

By applying the Sine Formula to the triangles $\triangle PFT$ and $\triangle PF'T$, we obtain that

$$\frac{PF}{TF} = \frac{\sin\angle PTF}{\sin\angle TPF} \qquad \text{and} \qquad \frac{PF'}{TF'} = \frac{\sin\angle PTF'}{\sin\angle TPF}$$

so that

$$\frac{\sin\angle PTF}{\sin\angle TPF} = \frac{\sin\angle PTF'}{\sin\angle TPF'}$$

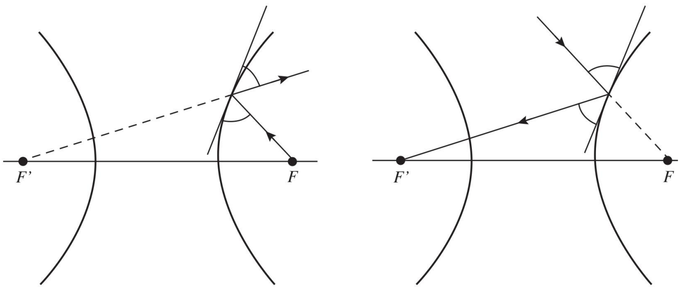
Since $\angle PTF = \angle PTF'$ it follows that $\sin\angle TPF = \sin\angle TPF'$, so that $\angle TPF = \pi - \angle TPF'$ since $\angle TPF \neq \angle TPF'$. Hence $\angle TPF$ equals the angle denoted by the symbol α in the diagram, and this is equal to the angle β (as α and β are vertically opposite).

This completes the proof of the Reflection Property. □

An amusing illustration of the property is as follows. A poor snooker player could appear to be a 'crack shot' if he used a snooker table in the shape of an ellipse: for if he places his snooker ball on the table at one focus and a target ball at the other focus, then no matter what direction he hits his ball, he is certain to reach his target!

Reflection Property of the Hyperbola

The hyperbola has a reflection property similar to that of the ellipse, with an appropriate modification.



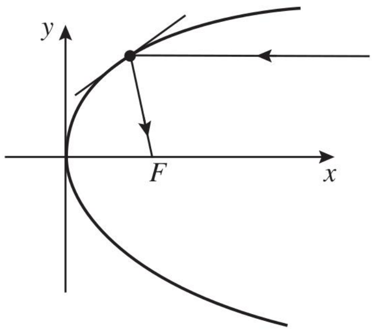
> **Reflection Property of the Hyperbola** Light coming from one focus of a hyperbolic mirror is reflected at the hyperbola in such a way that the light appears to have come from the other focus.

Also, light going towards one focus of a hyperbolic mirror is reflected at the mirror towards the other focus.

We omit a proof of this result, as it is similar to the proof of the Reflection Property of the ellipse.

Reflection Property of the Parabola

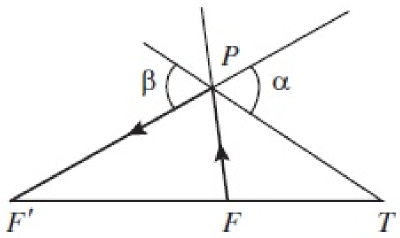
The Reflection Property of the parabola is also similar to the reflection property of the ellipse.



Reflection Property of the Parabola Incoming light parallel to the axis of a parabolic mirror is reflected at the parabola to pass through the focus. Conversely, light coming from the focus of a parabolic mirror is reflected at the parabola to give a beam of light parallel to the axis of the parabola.

Proof: Let E be the parabola in standard form, and let $P\left(at^2, 2at\right)$ be an arbitrary point on E.

We have seen that the equation of the tangent at P to the parabola has equation $ty = x + at^2$. If T is the point where this tangent meets the x -axis, then at T we have $y = 0$ and $t \cdot 0 = x + at^2$, so that $x = -at^2$.



In the triangle $\triangle PTF$ we have

$$TF = TO + OF = at^2 + a$$

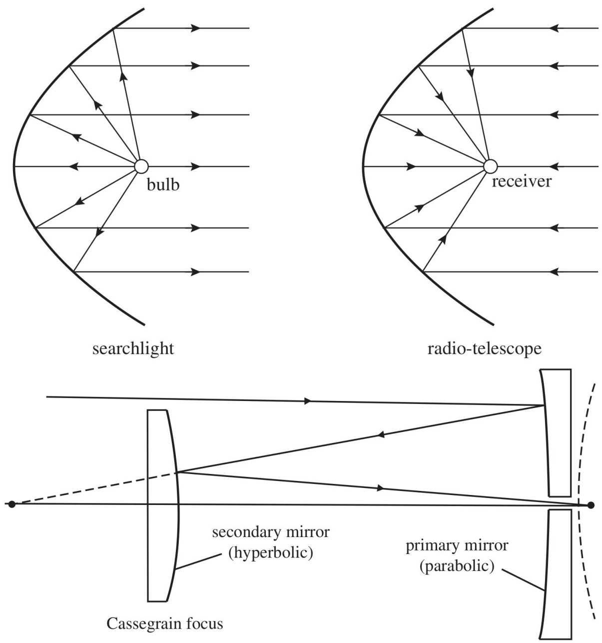
and, by the Distance Formula,

$$\begin{aligned} FP = \sqrt{\left(a - at^2\right)^2 + (2at)^2} &= \sqrt{a^2 + 2a^2t^2 + a^2t^4} \\ &= a + at^2. \end{aligned}$$

Then, since $TF = FP$, the triangle $\triangle PTF$ is isosceles, and so $\angle TPF = \angle FTP$. Now since the horizontal line through P is parallel to the x -axis, the angle between the tangent at P and the horizontal line through P is equal to $\angle FTP$ (as they are corresponding angles), and so also to $\angle TPF$. This completes the proof of the reflection property. $\square$

The reflection property of the parabola is also the principle behind the design of searchlights as well as radio-telescopes. The reflector of a searchlight is a parabolic mirror, with the bulb at its focus. Light from the bulb hits the mirror and is reflected outwards as a parallel beam.

The design of optical telescopes sometimes uses the Reflection Properties of other conics too. For example, the 4.2 metre William Herschel telescope at the Roque de los Muchachos Observatory on the island of La Palma in the Canary Islands, has an arrangement of mirrors known as a Cassegrain focus: a primary parabolic mirror reflects light towards a secondary hyperbolic mirror, which reflects it again to a focus behind the primary mirror.



The secondary mirror is used to focus the light to a much more convenient place than the focus of the primary mirror, and to increase the effective focal length of the telescope (and so its resolution).

We can summarize the above three Reflection Properties concisely as follows. All mirrors in the shape of a non-degenerate conic reflect light coming from or going to one focus towards the other focus.

Problem 8. Let E and H be an ellipse and a hyperbola, both having the same points F and F' as their foci. Use the reflection properties of the ellipse and hyperbola to prove that at each point of intersection, E and H meet at right angles.

Solution: Let $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \beta_1, \beta_2, \beta_3, \beta_4$ be the angles indicated in the above diagram.

Then

$$\begin{aligned} \alpha_1 &= \alpha_2 \quad \text{(vertically opposite angles)} \\ &= \alpha_3 \quad \text{(by the Reflection Property for the ellipse)} \\ &= \alpha_4 \quad \text{(vertically opposite angles),} \end{aligned}$$

and

$$\begin{aligned} \beta_1 &= \beta_2 \quad \text{(vertically opposite angles)} \\ &= \beta_3 \quad \text{(by the Reflection Property for the hyperbola)} \\ &= \beta_4 \quad \text{(vertically opposite angles).} \end{aligned}$$

Since $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \beta_1 + \beta_2 + \beta_3 + \beta_4 = 2\pi$, it follows that

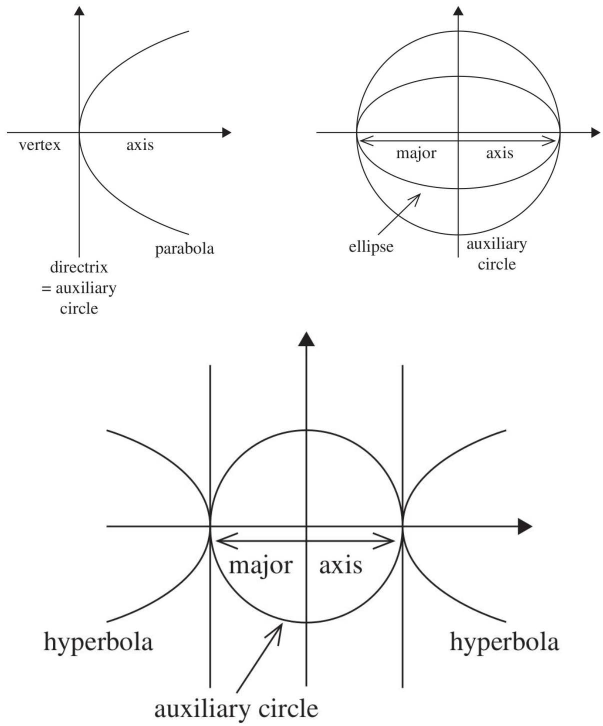
$$\alpha_i + \beta_j = \frac{1}{2}\pi \quad \text{for any } i \text{ and any } j$$

In particular, $\alpha_3 + \beta_3 = \frac{1}{2}\pi$, so that the tangents to E and H are perpendicular. In other words, E and H intersect at right angles. □

1.2.3 Conics as envelopes of tangent families

We now show how we can construct the non-degenerate conics as the envelope of a family of lines that are tangents to the conics. In other words, the conic being constructed is the curve in the plane that has each of the lines in the family as a tangent.

The method depends on the use of a circle associated with each nondegenerate conic, called its auxiliary circle. The auxiliary circle of an ellipse or hyperbola is the circle whose diameter is its major axis; analogously we shall define the tangent to a parabola at its vertex to be the auxiliary circle of the parabola.



The mathematical tool that we use in our construction is the following result.

Theorem 4. A perpendicular from a focus of a non-degenerate conic to a tangent meets the tangent on the auxiliary circle of the conic.

Proof: (for a parabola) Let the point $P\,(at^2, 2at)$ lie on the parabola in standard form with equation $y^2 = 4ax$, and let the perpendicular from the focus $F(a, 0)$ to the tangent at P meet it at T. By Theorem 2 of Subsection 1.2.1, the tangent at P

has equation

$$y \cdot 2at = 2a\left(x + at^2\right)$$

which we may rewrite in the form

$$y = \frac{1}{t}x + at \tag{5}$$

From this we see that the slope of the tangent PT is $1/t$, so that the slope of the perpendicular FT must be $-t$. Since FT also passes through $F(a,0)$, FT must have equation

$$y + tx = 0 + t \cdot a$$

which we may rewrite in the form

$$y = -tx + at \tag{6}$$

The equations (5) for PT and (6) for FT clearly have the solution $x = 0$, $y = at$. This means that the point of intersection T of the lines PT and FT has coordinates $(0, at)$. Hence T lies on the directrix of the parabola, as required. $\square$

Remark

Given a parabola and its axis, we can use Theorem 4 to identify the focus of the parabola. We draw the tangent at any point P on the parabola, and then the perpendicular to the tangent at the point T where the tangent meets the directrix.

This perpendicular crosses the parabola's axis at its focus.

Problem 9. Prove Theorem 4 for an ellipse.

Solution: Let the point $P(a\cos t, b\sin t)$ lie on the ellipse in standard form with equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, a \geq b > 0$, and let the perpendicular from the focus $F(ae, 0)$ to the tangent at P meet that tangent at T.

By Theorem 2 of Subsection 1.2.1, the tangent at P has equation

$$\frac{x \cdot a\cos t}{a^2} + \frac{y \cdot b\sin t}{b^2} = 1$$

which we may rewrite in the form

$$bx\cos t + ay\sin t = ab \qquad (*)$$

From this we see that, if $t \notin \{-\pi/2, 0, \pi/2, \pi\}$ the slope of the tangent PT is $-(b/a)\cot t$, so that the slope of the perpendicular FT must be $(a/b)\tan t$. Since FT also passes through $F(ae, 0), FT$ has equation

$$\begin{aligned} y - \frac{a}{b}\tan t \cdot x &= -\frac{a}{b}\tan t \cdot ae \\ &= -\frac{a^2 e}{b}\tan t \end{aligned}$$

which we may rewrite in the form

$$ax\sin t - by\cos t = a^2 e\sin t. \qquad (**)$$

Then the coordinates of the point $T(x,y)$ of intersection of PT and FT must satisfy both equations (*) and (**). So, squaring each of (*) and (**) and adding, we find that the coordinates of T must satisfy the equation

$$\left(x^2+y^2\right)\left(b^2\cos^2 t + a^2\sin^2 t\right) = a^2\left(b^2 + a^2e^2\sin^2 t\right).$$

We then rewrite this equation in the form

$$\begin{aligned}
x^2+y^2 &= a^2\frac{b^2+a^2e^2\sin^2 t}{b^2\cos^2 t + a^2\sin^2 t}\\
&= a^2\frac{\left(a^2-a^2e^2\right)+a^2e^2\sin^2 t}{\left(a^2-a^2e^2\right)\left(1-\sin^2 t\right)+a^2\sin^2 t}\\
&= a^2\frac{1-e^2+e^2\sin^2 t}{1-\sin^2 t - e^2 + e^2\sin^2 t + \sin^2 t}\\
&= a^2.
\end{aligned}$$

It follows that the point T must lie on the auxiliary circle $x^2+y^2=a^2$, as required.

If $t=0$ or π, the tangent to the ellipse at P is a vertical line perpendicular to FP; so the tangent at P meets the perpendicular to it from F at $P-$ which lies on the auxiliary circle.

Finally, if $t=\pm\pi/2$, the tangent to the ellipse at P is a horizontal line with equation $y=\pm b$. The point T where PT is perpendicular to FT must thus satisfy the equations $x=ae$

and $y = \pm b$; this lies on the auxiliary circle, since

$$\begin{aligned} x^2 + y^2 &= a^2e^2 + (\pm b)^2 \\ &= a^2e^2 + a^2\left(1 - e^2\right) = a^2 \end{aligned}$$

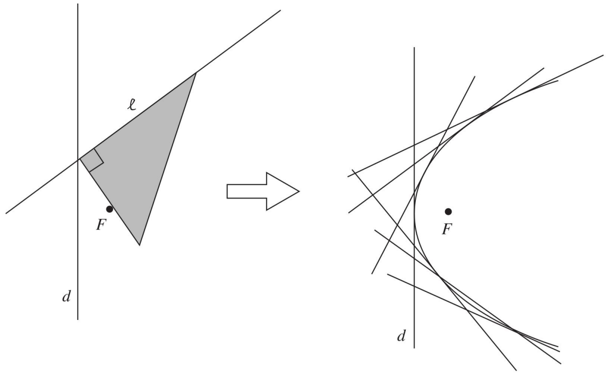
□

To construct the envelopes of the conics, you will need a sheet of paper, a pair of compasses, a set square and a pin.

Parabola

Draw a line d for the directrix of the parabola and a point F (not on d) for its focus. Place a set square so that its right-angled vertex lies at a point of d and one of its adjacent sides passes through F; draw the line ℓ along the other adjacent side of the set square. By Theorem 4, ℓ is a tangent to the parabola with focus F and directrix d.

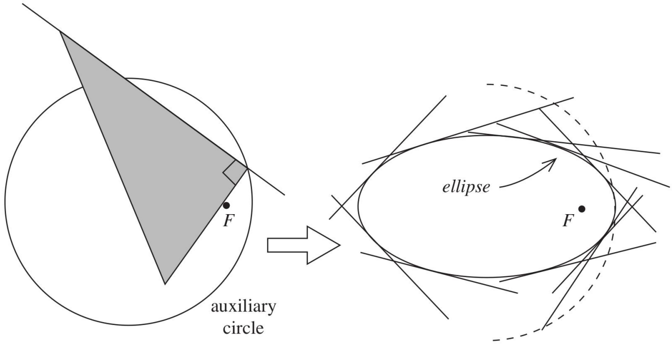
Repeating the process with the vertex of the set square at different points of d gives a family of lines ℓ that is the envelope of tangents to the parabola, as shown below.



Ellipse

Draw a circle C for the auxiliary circle of the ellipse and a point F inside C (but not at its centre) for a focus. Place a set square so that its right-angled vertex lies at a point of C and one of its adjacent sides passes through F; draw Properties of Conics the line ℓ along the other adjacent side of the set square. By Theorem 4, ℓ is a tangent to the ellipse with focus F and auxiliary circle C.

Repeating the process with the vertex of the set square at different points of C gives a family of lines that is the envelope of tangents to the ellipse, as shown below.

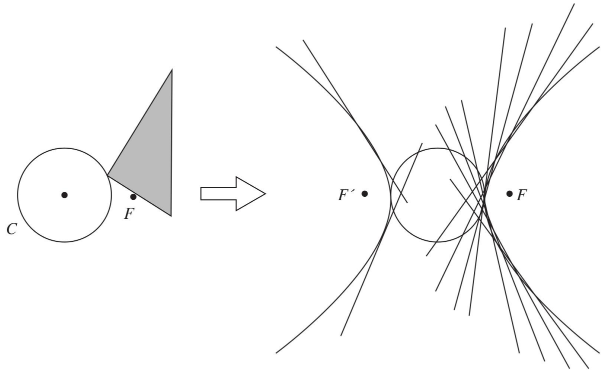


Hyperbola

Draw a circle C for the auxiliary circle of the hyperbola and a point F outside C for a focus. Place a set square so that its right-angled vertex lies at a point of C and one of its adjacent sides passes through F; draw the line ℓ along the other adjacent side of the set square. By Theorem $4, \ell$ is a tangent to the hyperbola with focus F and auxiliary circle C.

Repeating the process with the vertex of the set square at different points of C gives a family of lines that is the envelope of tangents to one branch of the hyperbola, as shown below.

Repeating the construction with the other focus F' (diametrically opposite F with respect to C) gives the other branch of the hyperbola.



1.3 Recognizing Conics

So far, we have considered the equation of a conic largely when it is in 'standard form'; that is, when the centre of the conic (if it has a centre) is at the origin, and the axes of the conic are parallel to the x -and y -axes. However, most of the conics which arise in calculations are not in standard form; thus we need some way of determining from the equation of a conic which type of conic it describes.

First we observe that all the equations of all (non-degenerate) conics in standard form can be expressed in the form

$$Ax^2 + Bxy + Cy^2 + Fx + Gy + H = 0 \qquad (1)$$

where not all of A, B and C are zero. For example, the equation of the circle

$$x^2 + y^2 + 4x + 6y - 23 = 0 \qquad (2)$$

is of the form (1), with $A = C = 1, B = 0, F = 4, G = 6$ and $H = -23$.

Now we can obtain any non-degenerate conic from a conic in standard form by a suitable rotation

$$(x, y) \mapsto (x \cos\theta - y \sin\theta, x \sin\theta + y \cos\theta)$$

followed by a suitable translation

$$(x, y) \mapsto (x - a, y - b)$$

Both of these transformations are linear, so that the equation of the conic at each stage is a second degree equation of the type (1); in other words, any non-degenerate conic has an equation of type (1).

The equations of degenerate conics can also be expressed in the form (1). For example,

$x^2 + y^2 = 0$	represents the single point $(0, 0)$;
$y^2 - 2xy + x^2 = 0$	represents the single line $y = x$, since $y^2 - 2xy + x^2 = (y - x)^2$;
$y^2 - x^2 = 0$	represents the pair of lines $y = \pm x$, since $y^2 - x^2 = (y + x)(y - x)$

However, an equation of the form (1) can also describe the empty set; an example of this is the equation $x^2 + y^2 + 1 = 0$, as there are no points (x, y) in $\mathbb{R}^2$ for which $x^2 + y^2 = -1$. For simplicity in the statement of the theorem below, therefore, we add the empty set to our existing list of degenerate conics.

In the above discussion, we proved one part of the following result.

> **Theorem 1.** Any conic has an equation of the form
>
> $$Ax^2 + Bxy + Cy^2 + Fx + Gy + H = 0 \tag{3}$$
>
> where A, B, C, F, G and H are real numbers, and not all of A, B and C are zero. Conversely, any set of points in $\mathbb{R}^2$ whose coordinates (x, y) satisfy equation (3) is a conic.

In this section we investigate the classification of conics in terms of equation (3). In particular, if we are given the equation of a non-degenerate conic in the form (3) how can we determine whether it is a parabola, an ellipse or a hyperbola? And how can we identify its vertex or centre? And its axis, or its major and minor axes? A key tool in this work is the matrix representation of the equation of a conic.

Introducing Matrices

We can express a general second degree equation in x and y

$$Ax^2 + Bxy + Cy^2 + Fx + Gy + H = 0 \tag{4}$$

where A, B and C are not all zero, in terms of matrices as follows.

Let $\mathbf{A} = \begin{pmatrix} A & \frac{1}{2}B \\ \frac{1}{2}B & C \end{pmatrix}$, $\mathbf{J} = \begin{pmatrix} F \\ G \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$. Then

$$\begin{aligned}\mathbf{x}^T\mathbf{A}\mathbf{x} &= \begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} A & \frac{1}{2}B \\ \frac{1}{2}B & C \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \\ &= \left(Ax + \frac{1}{2}By \frac{1}{2}Bx + Cy\right)\begin{pmatrix} x \\ y \end{pmatrix} \\ &= Ax^2 + Bxy + Cy^2\end{aligned}$$

and

$$\begin{aligned}\mathbf{J}^T\mathbf{x} &= \begin{pmatrix} F & G \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \\ &= Fx + Gy\end{aligned}$$

We may therefore write the equation (4) in the form

$$\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{J}^T\mathbf{x} + H = 0 \tag{5}$$

For example, let E be the conic with equation

$$3x^2 - 10xy + 3y^2 + 14x - 2y + 3 = 0$$

The equation of E is of the form (4) with $A = 3, B = -10, C = 3, F = 14$, $G = -2$ and $H = 3$. It follows from the above discussion that we can express the equation of E in matrix form as $\mathbf{x}^{\mathbf{T}}\mathbf{A}\mathbf{x} + \mathbf{J}^{\mathrm{T}}\mathbf{x} + H = 0$, where

$$\mathbf{A} = \begin{pmatrix} 3 & -5 \\ -5 & 3 \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} 14 \\ -2 \end{pmatrix}, \quad H = 3 \quad \textit{and} \quad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Problem 1. Write the equation of each of the following conics in matrix form.

(a) $11x^2 + 4xy + 14y^2 - 4x - 28y - 16 = 0$

(b) $x^2 - 4xy + 4y^2 - 6x - 8y + 5 = 0$

Solution:

(a) The equation of the conic

$$11x^2 + 4xy + 14y^2 - 4x - 28y - 16 = 0$$

can be written in matrix form $\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{J}^T\mathbf{x} + H = 0$, where

$$\mathbf{A} = \begin{pmatrix} 11 & 2 \\ 2 & 14 \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} -4 \\ -28 \end{pmatrix},$$
$$H = -16 \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

(b) The equation of the conic

$$x^2 - 4xy + 4y^2 - 6x - 8y + 5 = 0$$

can be written in matrix form $\mathbf{x}^T\mathbf{A}\mathbf{x}+\mathbf{J}^T\mathbf{x}+H=0$ where

$$\mathbf{A}=\begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}, \quad \mathbf{J}=\begin{pmatrix} -6 \\ -8 \end{pmatrix},$$

$$H=5 \quad \text{and} \quad \mathbf{x}=\begin{pmatrix} x \\ y \end{pmatrix}.$$

□

A key tool in our use of matrices will be the following result.

Theorem 2. A 2×2 matrix $\mathbf{P}$ represents a rotation of $\mathbb{R}^2$ about the origin if and only if it satisfies the following two conditions:

(a) $\mathbf{P}$ is orthogonal;

(b) $\det \mathbf{P}=1$.

Proof: A matrix $\mathbf{P}$ represents a rotation about the origin (anticlockwise through an angle θ) if and only it is of the form

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{6}$$

It is easy to verify that $\mathbf{P}$ satisfies conditions (a) and (b).

Next, let $\mathbf{P}=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be a matrix that satisfies conditions (a) and (b).

Then, since $\mathbf{P}$ is orthogonal, the vector $\begin{pmatrix} a \\ c \end{pmatrix}$ has length 1 ; that is, $a^2 + c^2 = 1$. Thus there is a number θ for which

$$a = \cos\theta \quad \textit{and} \quad c = \sin\theta.$$

Also, since $\mathbf{P}$ is orthogonal, the vectors $\begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$ and $\begin{pmatrix} b \\ d \end{pmatrix}$ are orthogonal; that is, $(\cos\theta \quad \sin\theta)\begin{pmatrix} b \\ d \end{pmatrix} = 0$ or

$$\cos\theta \cdot b + \sin\theta \cdot d = 0$$

So there exists some number λ, say, such that

$$b = -\lambda \sin\theta \quad \text{and} \quad d = \lambda\cos\theta$$

Then since $\det \mathbf{P} = 1$, we have

$$1 = ad - bc = \lambda\cos^2\theta + \lambda\sin^2\theta$$

so that $\lambda = 1$. It follows that $\mathbf{P}$ must be of the form (6), and so represent a rotation of $\mathbb{R}^2$ about the origin. $\square$

Using Matrices

We now use the methods of Linear Algebra to recognize conics specified by their equations.

Example 1. Prove that the conic E with equation

$$3x^2 - 10xy + 3y^2 + 14x - 2y + 3 = 0$$

is a hyperbola. Determine its centre, and its major and minor axes.

Solution: We saw above that the equation of E can be written in matrix form as $\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{J}^T\mathbf{x} + H = 0$, where

$$\mathbf{A} = \begin{pmatrix} 3 & -5 \\ -5 & 3 \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} 14 \\ -2 \end{pmatrix}, \quad H = 3 \text{ and } \quad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix};$$

that is, as

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 3 & -5 \\ -5 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 14 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + 3 = 0.$$

We start by diagonalizing the matrix $\mathbf{A}$. Its characteristic equation is

$$\begin{aligned} 0 = \det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} 3-\lambda & -5 \\ -5 & 3-\lambda \end{vmatrix} \\ &= \lambda^2 - 6\lambda - 16 \\ &= (\lambda - 8)(\lambda + 2), \end{aligned}$$

so that the eigenvalues of $\mathbf{A}$ are $\lambda = 8$ and $\lambda = -2$. The eigenvector equations of $\mathbf{A}$ are

$$\begin{aligned} (3-\lambda)x - 5y &= 0 \\ -5x + (3-\lambda)y &= 0 \end{aligned}$$

When $\lambda = 8$, these equations both become

$$-5x - 5y = 0$$

so that we may take as a corresponding eigenvector $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, which we normalize to have unit length as $\begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$.

When $\lambda = -2$, the eigenvector equations of $\mathbf{A}$ become

$$5x - 5y = 0$$
$$-5x + 5y = 0$$

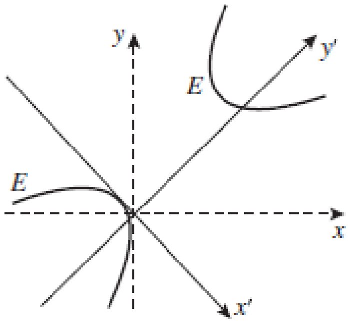
so that we may take as a corresponding eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, which we normalize to have unit length as $\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$.

Now
$$\begin{vmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{vmatrix} = \frac{1}{2} + \frac{1}{2} = 1,$$
so we take as our rotation of the plane the transformation $\mathbf{x} = \mathbf{P}\mathbf{x}'$ where $\mathbf{P} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$. The transformation $\mathbf{x} = \mathbf{P}\mathbf{x}$ changes the equation of the conic to the form

$$(\mathbf{P}\mathbf{x}')^T \mathbf{A} (\mathbf{P}\mathbf{x}') + \mathbf{J}^T (\mathbf{P}\mathbf{x}') + H = 0$$

or

$$(\mathbf{x}')^T \left(\mathbf{P}^T \mathbf{A} \mathbf{P}\right) \mathbf{x}' + \left(\mathbf{J}^T \mathbf{P}\right) \mathbf{x}' + H = 0.$$



Since $\mathbf{P}^T \mathbf{A} \mathbf{P} = \begin{bmatrix} 8 & 0 \\ 0 & -2 \end{bmatrix}$, the equation of the conic is now

$$(x' \quad y') \begin{pmatrix} 8 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

$$+ (14 \quad -2) \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + 3 = 0,$$

which we can rewrite in the form

$$8x'^2 - 2y^2 + 8\sqrt{2}x' + 6\sqrt{2}y' + 3 = 0$$

We may rewrite this equation in the form

$$8\left(x^2 + \sqrt{2}x'\right) - 2\left(y'^2 - 3\sqrt{2}y'\right) + 3 = 0$$

so that, on completing the square, we have

$$8\left(x' + 1/\sqrt{2}\right)^2 - 4 - 2\left(y' - 3/\sqrt{2}\right)^2 + 9 + 3 = 0$$

which we can rewrite in the form

$$8\left(x' + 1/\sqrt{2}\right)^2 - 2\left(y' - 3/\sqrt{2}\right)^2 = -8$$

or

$$\frac{\left(y' - 3/\sqrt{2}\right)^2}{4} - \frac{\left(x' + 1/\sqrt{2}\right)^2}{1} = 1 \tag{7}$$

This is the equation of a hyperbola.

From equation (7) it follows that the centre of the hyperbola E is the point where $x' = -1/\sqrt{2}$ and $y' = 3/\sqrt{2}$. From the equation $\mathbf{x} = \mathbf{P}\mathbf{x}'$, it follows that in terms of the original coordinate system this is the point

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} \\ 3/\sqrt{2} \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \end{aligned}$$

that is, the point $(1, 2)$.

It also follows from equation (7) that the major axis of E has equation $x' + 1/\sqrt{2} = 0$, or $x' = -1/\sqrt{2}$; and the minor axis of E has equation $y' - 3/\sqrt{2} = 0$, or $y' = 3/\sqrt{2}$.

Finally, since the matrix $\mathbf{P}$ is orthogonal we can rewrite the equation $\mathbf{x} = \mathbf{P}\mathbf{x}'$ in the form $\mathbf{x}' = \mathbf{P}^{-1}\mathbf{x} = \mathbf{P}^T\mathbf{x}$, so that

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

or as a pair of equations

$$\begin{aligned} x' &= \frac{1}{\sqrt{2}}x - \frac{1}{\sqrt{2}}y \\ y' &= \frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y \end{aligned}$$

It follows that the equation, $x' = -1/\sqrt{2}$, of the major axis of the hyperbola E can be expressed in terms of the original coordinate system as

$$\frac{1}{\sqrt{2}}x - \frac{1}{\sqrt{2}}y = -\frac{1}{\sqrt{2}}, \qquad \text{or} \qquad x - y = -1$$

Similarly, the equation, $y' = 3/\sqrt{2}$, of the minor axis of the hyperbola can be expressed in terms of the original coordinate system as

$$\frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y = \frac{3}{\sqrt{2}}, \qquad \text{or} \qquad x + y = 3.$$

$\square$

The above problem illustrates a general strategy for identifying conics from their second degree equations.

Strategy. To classify a conic E with equation

$$Ax^2 + Bxy + Cy^2 + Fx + Gy + H = 0:$$

1. Write the equation of E in matrix form $\mathbf{x}^T\mathbf{A}\mathbf{x}+\mathbf{J}^T\mathbf{x}+H=0$.

2. Determine an orthogonal matrix $\mathbf{P}$, with determinant 1 , that diagonalizes $\mathbf{A}$.

3. Make the change of coordinate system $\mathbf{x}=\mathbf{P}\mathbf{x}'$. The equation of E then becomes of the form

$$\lambda_1 x'^2 + \lambda_2 y'^2 + fx' + gy' + h = 0$$

   where λ_1 and λ_2 are the eigenvalues of $\mathbf{A}$.

4. 'Complete the squares', if necessary, to rewrite the equation of E in terms of an (x'', y'') -coordinate system as the equation of a conic in standard form.

5. Use the equation $\mathbf{x}' = \mathbf{P}^T\mathbf{x}$ to determine the centre and axes of E in terms of the original coordinate system.

Problem 2. Classify the conics in $\mathbb{R}^2$ with the following equations. Determine the centre of those that have a centre.

(a) $11x^2 + 4xy + 14y^2 - 4x - 28y - 16 = 0$

(b) $x^2 - 4xy + 4y^2 - 6x - 8y + 5 = 0$

Solution:

(a) We saw in Problem 1 (a) that the matrix form of the equation of this conic is $\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{J}^T\mathbf{x} + H = 0$, where

$$\mathbf{A} = \begin{pmatrix} 11 & 2 \\ 2 & 14 \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} -4 \\ -28 \end{pmatrix},$$
$$H = -16 \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix};$$

that is

$$(xy)\begin{pmatrix} 11 & 2 \\ 2 & 14 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + (-4 - 28)\begin{pmatrix} x \\ y \end{pmatrix} - 16 = 0$$

First we diagonalize A. Its characteristic equation is

$$\begin{aligned} 0 = \det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} 11-\lambda & 2 \\ 2 & 14-\lambda \end{vmatrix} \\ &= \lambda^2 - 25\lambda + 150 \\ &= (\lambda - 15)(\lambda - 10), \end{aligned}$$

so that the eigenvalues of $\mathbf{A}$ are $\lambda = 15$ and $\lambda = 10$. The eigenvector equations of $\mathbf{A}$ are

$$\begin{aligned} (11 - \lambda)x + 2y &= 0 \\ 2x + (14 - \lambda)y &= 0 \end{aligned}$$

When $\lambda = 15$, these equations become

$$\begin{aligned} -4x + 2y &= 0 \\ 2x - y &= 0 \end{aligned}$$

so that we may take as a corresponding eigenvector $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, which we normalize to have unit length as $\begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$.

When $\lambda = 10$, the eigenvector equations of $\mathbf{A}$ become

$$\begin{aligned} x + 2y &= 0 \\ 2x + 4y &= 0 \end{aligned}$$

so that we may take as a corresponding eigenvector $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$ which we normalize to have unit length as $\begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix}$.

Now $\begin{vmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{vmatrix} = -\frac{1}{5} - \frac{4}{5} = -1$, so interchanging the order of the eigenvectors as columns of $\mathbf{P}-$ in order that $\det \mathbf{P} = +1$, so that then $\mathbf{P}$ represents a rotation rather than a rotation composed with a reflection $-$ we take as our rotation of the plane the transformation $\mathbf{x} = \mathbf{P}\mathbf{x}'$, where $\mathbf{P} = \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$

This transformation changes the equation of the conic to the form

$$(\mathbf{Px}')^T \mathbf{A} (\mathbf{Px}') + \mathbf{J}^T (\mathbf{Px}') + H = 0$$

or

$$(\mathbf{x}')^T \left(\mathbf{P}^T\mathbf{AP}\right) \mathbf{x}' + \left(\mathbf{J}^T\mathbf{P}\right) \mathbf{x}' + H = 0$$

Since $\mathbf{P}^T\mathbf{AP} = \begin{pmatrix} 10 & 0 \\ 0 & 15 \end{pmatrix}$, this is the equation $(x' \quad y') \begin{pmatrix} 10 & 0 \\ 0 & 15 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + (-4 - 28) \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} - 16 = 0.$

We may rewrite this equation in the form

$$10x^2 + 15y'^2 + 4\sqrt{5}x' - 12\sqrt{5}y' - 16 = 0$$

or

$$10\left(x^2 + \frac{2}{\sqrt{5}}x'\right) + 15\left(y'^2 - \frac{4}{\sqrt{5}}y'\right) - 16 = 0;$$

so that, on completing the square, we have

$$10\left(x' + \frac{1}{\sqrt{5}}\right)^2 - 2 + 15\left(y' - \frac{2}{\sqrt{5}}\right)^2 - 12 - 16 = 0$$

or

$$10\left(x' + \frac{1}{\sqrt{5}}\right)^2 + 15\left(y' - \frac{2}{\sqrt{5}}\right)^2 - 30 = 0$$

or

$$\frac{\left(x' + \frac{1}{\sqrt{5}}\right)^2}{3} + \frac{\left(y' - \frac{2}{\sqrt{5}}\right)^2}{2} = 1 \qquad (*)$$

This is the equation of an ellipse.

From equation $(*)$ it follows that the centre of the ellipse is the point where $x' + \frac{1}{\sqrt{5}} = 0$ and $y' - \frac{2}{\sqrt{5}} = 0$, that is, where $x' = -\frac{1}{\sqrt{5}}$ and $y' = \frac{2}{\sqrt{5}}$. From the equa- tion $\mathbf{x} = \mathbf{P}\mathbf{x}'$, it follows that in terms of the original coordinate system this is the point

$$\begin{aligned}\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}\end{aligned}$$

that is, the point $(0, 1)$.

Since $3 > 2$, it also follows from equation (2) that the major axis of the ellipse has equation $y' - \frac{2}{\sqrt{5}} = 0$, or $y' = \frac{2}{\sqrt{5}}$; and the minor axis has equation $x' + \frac{1}{\sqrt{5}} = 0$, or $x' = -\frac{1}{\sqrt{5}}$. Finally, since the matrix $\mathbf{P}$ is orthogonal we can rewrite the equation $\mathbf{x} = \mathbf{P}\mathbf{x}'$ in the form $\mathbf{x}' = \mathbf{P}^{-1}\mathbf{x} = \mathbf{P}^T\mathbf{x}$, so that

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

or as a pair of equations

$$x' = \frac{2}{\sqrt{5}}x - \frac{1}{\sqrt{5}}y$$
$$y' = \frac{1}{\sqrt{5}}x + \frac{2}{\sqrt{5}}y$$

It follows that the equation, $y' = \frac{2}{\sqrt{5}}$, of the major axis of the ellipse can be expressed in terms of the original coordinate system as

$$\frac{1}{\sqrt{5}}x + \frac{2}{\sqrt{5}}y = \frac{2}{\sqrt{5}} \quad \text{or} \quad x + 2y = 2$$

Similarly, the equation, $x' = -\frac{1}{\sqrt{5}}$, of the minor axis of the ellipse can be expressed in terms of the original coordinate system as $\frac{2}{\sqrt{5}}x - \frac{1}{\sqrt{5}}y = -\frac{1}{\sqrt{5}} \quad \text{or} \quad 2x - y = -1$

(b) We saw in Problem 1(b) that the matrix form of the equation of this conic is $\mathbf{x}^T\mathbf{A}\mathbf{x}+ \mathbf{J}^T\mathbf{x} + H = 0$, where

$$\mathbf{A} = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} -6 \\ -8 \end{pmatrix}, \quad H = 5 \quad \textit{and} \quad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$$

that is

$$\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + (-6 - 8)\begin{pmatrix} x \\ y \end{pmatrix} + 5 = 0$$

First we diagonalize A. Its characteristic equation is

$$
\begin{aligned}
0 = \det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} 1-\lambda & -2 \\ -2 & 4-\lambda \end{vmatrix} \\
&= \lambda^2 - 5\lambda \\
&= \lambda(\lambda - 5),
\end{aligned}
$$

so that the eigenvalues of $\mathbf{A}$ are $\lambda = 0$ and $\lambda = 5$. The eigenvector equations of $\mathbf{A}$ are

$$(1-\lambda)x - 2y = 0$$

$$-2x + (4-\lambda)y = 0$$

When $\lambda = 0$, these equations become

$$x - 2y = 0$$

$$-2x + 4y = 0$$

so that we may take as a corresponding eigenvector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, which we normalize to have unit length as $\begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}$.

When $\lambda = 5$, the eigenvector equations of A become

$$
\begin{aligned}
-4x - 2y &= 0 \\
-2x - y &= 0
\end{aligned}
$$

so that we may take as a corresponding eigenvector $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$, which we normalize to have unit length as $\begin{pmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{pmatrix}$.

Now $\begin{vmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{vmatrix} = -\frac{4}{5} - \frac{1}{5} = -1$ so interchanging the order of the eigenvectors as columns of $\mathbf{P}-$ in order that det $\mathbf{P} = +1$, so that then $\mathbf{P}$ represents a rotation rather than a rotation composed with a reflection $-$ we take as our rotation of the plane the transformation $\mathbf{x} = \mathbf{P}\mathbf{x}'$, where $\mathbf{P} = \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}$. This transformation changes the equation of the conic to the form

$$(\mathbf{P}\mathbf{x}')^T \mathbf{A} (\mathbf{P}\mathbf{x}') + \mathbf{J}^T (\mathbf{P}\mathbf{x}') + H = 0$$

or

$$(\mathbf{x}')^T \left(\mathbf{P}^T\mathbf{A}\mathbf{P}\right) \mathbf{x}' + \left(\mathbf{J}^T\mathbf{P}\right) \mathbf{x}' + H = 0$$

Since $\mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix}$, this is the equation

$$\begin{pmatrix} x' & y' \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + (-6 - 8) \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

+5=0 (8)

which we can rewrite in the form

$$5x'^2 + 2\sqrt{5}x' - 4\sqrt{5}y' + 5 = 0$$

We may rewrite this equation in the form

$$5\left(x'^2 + \frac{2}{\sqrt{5}}x'\right) - 4\sqrt{5}y' + 5 = 0$$

so that, on completing the square, we have

$$5\left(x' + \frac{1}{\sqrt{5}}\right)^2 - 4\sqrt{5}y' + 4 = 0$$

or

$$5\left(x' + \frac{1}{\sqrt{5}}\right)^2 - 4\sqrt{5}\left(y' - \frac{1}{\sqrt{5}}\right) = 0$$

or

$$\left(x' + \frac{1}{\sqrt{5}}\right)^2 = \frac{4}{\sqrt{5}}\left(y' - \frac{1}{\sqrt{5}}\right) \qquad (**)$$

This is the equation of a parabola. (It is not quite in standard form $(y'')^2 = 4ax''$, but in the similar form $(x'')^2 = 4ay''$; the argument will be similar.)

From equation $(**)$ it follows that the vertex of the parabola is the point where $x' + \frac{1}{\sqrt{5}} = 0$ and $y' - \frac{1}{\sqrt{5}} = 0$, that is, where $x' = -\frac{1}{\sqrt{5}}$ and $y' = \frac{1}{\sqrt{5}}$. From the equation $\mathbf{x} = \mathbf{P}\mathbf{x}'$, it follows that in terms of the original coordinate system this is the point

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix}$$
$$= \begin{pmatrix} \frac{1}{5} \\ \frac{3}{5} \end{pmatrix},$$

that is, the point $\left(\frac{1}{5}, \frac{3}{5}\right)$.

It also follows from equation $(**)$ that the axis of the parabola has equation $x' + \frac{1}{\sqrt{5}} = 0$, or $x' = -\frac{1}{\sqrt{5}}$. Then, since the matrix $\mathbf{P}$ is orthogonal we can rewrite the equation $\mathbf{x} = \mathbf{P}\mathbf{x}'$ in the form $\mathbf{x}' = \mathbf{P}^{-1}\mathbf{x} = \mathbf{P}^T\mathbf{x}$, so that

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

or as a pair of equations

$$x' = \frac{1}{\sqrt{5}}x - \frac{2}{\sqrt{5}}y$$
$$y' = \frac{2}{\sqrt{5}}x + \frac{1}{\sqrt{5}}y$$

It follows that the equation, $x' = -\frac{1}{\sqrt{5}}$, of the axis of the parabola can be expressed in terms of the original coordinate system as

$$\frac{1}{\sqrt{5}}x - \frac{2}{\sqrt{5}}y = -\frac{1}{\sqrt{5}}$$

or

$$x - 2y = -1$$

$\square$

In fact, using the above strategy we can prove the following result.

Theorem 3. A non-degenerate conic with equation

$$Ax^2 + Bxy + Cy^2 + Fx + Gy + H = 0$$

and matrix $\mathbf{A} = \begin{pmatrix} A & \frac{1}{2}B \\ \frac{1}{2}B & C \end{pmatrix}$ can be classified as follows:

(a) If $\det \mathbf{A} < 0, E$ is a hyperbola.

(b) If $\det \mathbf{A} = 0, E$ is a parabola.

(c) If $\det \mathbf{A} > 0, E$ is an ellipse.

Problem 3. Use Theorem 3 to classify the non-degenerate conics in $\mathbb{R}^2$ with the following equations.

(a) $3x^2 - 8xy + 2y^2 - 2x + 4y - 16 = 0$

(b) $x^2 + 8xy + 16y^2 - x + 8y - 12 = 0$

(c) $52x^2 - 72xy + 73y^2 - 32x - 74y + 28 = 0$

Solution: Here we use Theorem 3 .

(a) The matrix of the non-degenerate conic is $\mathbf{A} =$

$\begin{pmatrix} 3 & -4 \\ -4 & 2 \end{pmatrix}$, so that

$$\det \mathbf{A} = 6 - 16 = -10 < 0$$

so that the conic is a hyperbola.

(b) The matrix of the non-degenerate conic is $\mathbf{A} = \begin{pmatrix} 1 & 4 \\ 4 & 16 \end{pmatrix}$, so that

$$\det \mathbf{A} = 16 - 16 = 0,$$

so that the conic is a parabola.

(c) The matrix of the non-degenerate conic is $\mathbf{A} = \begin{pmatrix} 52 & -36 \\ -36 & 73 \end{pmatrix}$, so that

$$\det \mathbf{A} = 52 \cdot 73 - 36^2 = 3796 - 1296 = 2500 > 0$$

so that the conic is an ellipse.

$\square$

1.4 Exercises

Section 1.1

1. Determine the equation of the circle with centre $(2, 1)$ and radius 3.

2. Determine the points of intersection of the line with equation $y = x + 2$ and the circle in Exercise 1 .

3. Determine whether the circles with equations

$$2x^2 + 2y^2 - 3x - 4y + 2 = 0 \quad \textit{and} \quad x^2 + y^2 - 4x + 2y = 0$$

intersect orthogonally. Find the equation of the line through their points of intersection.

4. This question concerns the parabola $y^2 = 4ax (a > 0)$ with parametric equations $x = at^2, y = 2at$ and focus F. Let P and Q be points on the parabola with parameters t_1 and t_2, respectively.

   (a) If PQ subtends a right angle at the vertex O of the parabola, prove that $t_1 \cdot t_2 = -4$.

   (b) If $t_1 = 2$ and PQ is perpendicular to OP, prove that $t_2 = -4$.

5. This question concerns the rectangular hyperbola $xy = c^2 (c > 0)$ with parametric equations $x = ct, y = c/t$. Let

P and Q be points on the hyperbola with parameters t_1 and t_2, respectively.

(a) Determine the equation of the chord PQ.

(b) Determine the coordinates of the point N where PQ meets the x -axis.

(c) Determine the midpoint M of PQ.

(d) Prove that $OM = MN$, where O is the origin.

6. Let P be a point in the plane and C a circle with centre O and radius r. Then we define the power of P with respect to C as

$$\text{power of } P \text{ with respect to } C = OP^2 - r^2$$

(a) Determine the sign of the power of P with respect to C when

(i) P lies inside C;

(ii) P lies on C;

(iii) P lies outside C.

In parts (b) and (c) we regard distances as directed distances; that is, distances along a line in one direction have a positive sign associated with their length and distances in the opposite direction have a negative sign associated with their length.

(b) If P lies inside C and a line through P meets C at two distinct points A and B, prove that

$$\text{power of } P \text{ with respect to } C = PA \cdot PB$$

(c) If P lies outside C, a line through P meets C at two distinct points A and B, and PT is one of the tangents from P to C, prove that

$$\begin{aligned}\text{power of } P \text{ with respect to } C &= PA \cdot PB \\ &= PT^2.\end{aligned}$$

(d) If C has equation $x^2 + y^2 + fx + gy + h = 0$ and P has coordinates (x, y), find the power of P with respect to C in terms of x, y, f, g and h.

7. (a) Let a plane π in $\mathbb{R}^3$ meet both portions of a right circular cone, in two separate portions of a curve E. Let the two spheres inside the cone (on the same side of π as the vertex) that each touch both the cone in a horizontal circle (C_1 and C_2, respectively) and π touch π at F and F', respectively. Let P be any point of E, and the generator of the cone through P meet C_1 and C_2 at A and B, respectively. Prove that $PF' - PF = AB$. Deduce that E is a hyperbola.

(b) Let a plane π in $\mathbb{R}^3$ that is parallel to a generator of a right circular cone meet the cone in a curve E. Let the sphere inside the cone (on the same side of π as the vertex) that touches both the cone in a horizontal circle C and π meet π at F. Let P be any point of E, and the generator of the cone through P

meet C at A. Let N be the foot of the perpendicular from P to the line of intersection of the horizontal plane and π, and let NA meet C again at M. Prove that $PF = PN$. Deduce that E is a parabola.

Section 1.2

1. Determine the slope of the tangent to the cycloid in $\mathbb{R}^2$ with parametric equations

$$x = t - \sin t, \quad y = 1 - \cos t$$

at the point with parameter t, where t is not a multiple of 2π.

2. Determine the equation of the tangent to the curve in $\mathbb{R}^2$ with parametric equations

$$x = 1 + 4t + t^2, \quad y = 1 - t$$

at the point where $t = 1$.

3. Let P be a point on the ellipse with equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, where $a > b > 0$, $b^2 = a^2\left(1 - e^2\right)$, and $0 < e < 1$.

   (a) If P has coordinates $(a\cos t, b\sin t)$, determine the equation of the tangent at P to the ellipse.

   (b) Determine the coordinates of the point T where the tangent in part (a) meets the directrix $x = a/e$.

(c) Let F be the focus with coordinates $(ae, 0)$. Prove that PF is perpendicular to TF.

4. The perpendicular from a point P on the hyperbola H with parametric equations $x = 2\sec t, y = 3\tan t$, to the x -axis meets the x -axis at the point N The tangent at P to H meets the x -axis at the point T.

   (a) Write down the coordinates of N.

   (b) Find the coordinates of T.

   (c) Prove that $ON \cdot OT = 4$, where O is the origin.

5. Let P be a point on the ellipse with equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, where $a > b > 0$, $b^2 = a^2\left(1 - e^2\right)$, and $0 < e < 1$.

   (a) If P has coordinates $(a\cos t, b\sin t)$, determine the equation of the normal at P to the ellipse.

   (b) Determine the coordinates of the point Q where the normal in part (a) meets the axis $y = 0$

   (c) Let F be the focus with coordinates $(ae, 0)$. Prove that $QF = e \cdot PF$.

6. Let $\mathscr{F}$ denote the family of parabolas $\{(x, y) : y^2 = 4a(x + a)\}$ as a takes all positive values, and $\mathscr{G}$ denote the family of parabolas $\{(x, y) : y^2 = 4a(-x + a)\}$ as a takes all positive values. Use the reflection property of the parabola to prove that, if $F \in \mathscr{F}$ and $G \in \mathscr{G}$, then, at each point of intersection, F and G cross at right angles.

7. Prove that a perpendicular from the focus nearer to a point P on an ellipse meets the tangent at P on the auxiliary circle of the ellipse, in the following geometric way.

   It is sufficient to prove the result for the ellipse $E : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, a > b > 0$, and points P of E in the first quadrant. Let T be the foot of the perpendicular from $F(ae, 0)$ to the tangent at P, let T' be the foot of the perpendicular from $F'(-ae, 0)$ to the tangent at P, and let FT meet $F'P$ at X.

   (a) Prove that the triangles ΔFPT and ΔXPT are congruent.

   (b) Using the sum of focal distances property for E, prove that $F'X = 2a$.

   (c) Prove that OT is parallel to $F'X$, where O is the centre of E.

   (d) Prove that $OT = a$, so that T lies on the auxiliary circle of E.

   Remark: A similar argument to that in parts (a)-(d) shows that $O\bar{T} = a$, so that T' also lies on the auxiliary circle of E.

8. (a) Let E be an ellipse with major axis AB and minor axis CD, and let the tangents to E at A and B meet the tangent at D at the points T and T', respectively. Prove that the circle with diameter TT' cuts the major axis of E at its foci.

(b) Let H be a hyperbola with major axis AB, whose midpoint is O, and let the perpendicular at A to the major axis meet an asymptote at a point T. Prove that the circle with centre O and radius OT cuts the major axis of H at its foci.

Section 1.3

1. Classify the conics in $\mathbb{R}^2$ with the following equations. Determine the centre/vertex and axis of each.

   (a) $x^2 - 4xy - 2y^2 + 6x + 12y + 21 = 0$

   (b) $5x^2 + 4xy + 5y^2 + 20x + 8y - 1 = 0$

   (c) $x^2 - 4xy + 4y^2 - 6x - 8y + 5 = 0$

   (d) $21x^2 - 24xy + 31y^2 + 6x + 4y - 25 = 0$

   (e) $3x^2 - 10xy + 3y^2 + 14x - 2y + 3 = 0$

2. Determine the eccentricities of the conics in parts (a), (b) and (c) of Exercise 1 .

MODULE 2

AFFINE GEOMETRY

2.1 Geometry and Transformations

Before embarking on a study of various other geometries, it is useful first to look back at our familiar Euclidean geometry.

2.1.1 What is Euclidean Geometry?

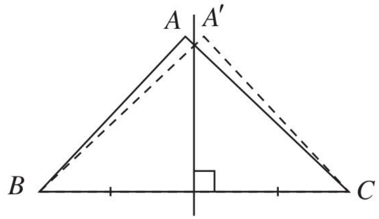
To help us answer this question, we begin by considering the following wellknown result.

Example 1. Let $\triangle ABC$ be a triangle in which $\angle ABC = \angle ACB$. Prove that $AB = AC$

Solution: First, reflect the triangle in the perpendicular bisector of BC, so that the points B and C change places and the point A moves to some point A', say. Since reflection preserves angles, it follows that $\angle A'BC = \angle ACB$.

Also, we are given that $\angle ACB = \angle ABC$, so

$$\angle A'BC = \angle ABC$$



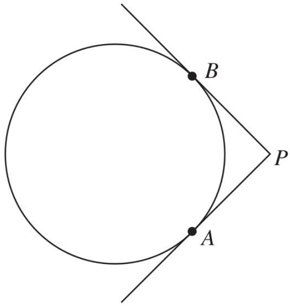
But this can happen only if A' lies on the line through A and B. Similarly,

$$\angle A'CB = \angle ABC = \angle ACB$$

so A' must also lie on the line through A and C. This means that A' and A must coincide. Hence the line segment AB reflects to the line segment AC, and vice versa. Since reflection preserves lengths, it follows that $AB = AC$. □

Problem 1. Let A and B be two points on a circle, and let the tangents to the circle at A and B meet at P. Prove that $AP = BP$.

Hint: Consider a reflection in the line which passes through P and the centre of the circle.



Solution: First, reflect the figure in the line through O, the centre of the circle, and P. Under this reflection, P remains fixed, and the circle maps onto itself. In particular, the point A maps to a point A' on the circle, and so the tangent PA maps onto the line PA'.

Now the tangent PA meets the circle at a single point A, so the image of the tangent must meet the circle at a single point. But the only way in which that can happen is if A' coincides with B. Hence the line segment PA is reflected onto the line segment PB. Since reflection preserves lengths, it follows that $PA = PB$. $\square$

The result in Example 1 is concerned with the properties of length and angle associated with the triangle $\triangle ABC$. To investigate these properties, we introduced a reflection that enabled us to compare various lengths and angles. We were able to do this because reflections leave lengths and angles unchanged.

Of course, reflections are not the only transformations that preserve lengths and angles: other examples include rotations and translations. In general, any transformation that preserves lengths and angles can be used to tackle problems which involve these properties. In fact, we need worry only about leaving distances unchanged, since any transformation from $\mathbb{R}^2$ onto $\mathbb{R}^2$ that changes angles must also change lengths. Transformations that leave distances unchanged are called isometries.

Definition. An isometry of $\mathbb{R}^2$ is a function which maps $\mathbb{R}^2$ onto $\mathbb{R}^2$ and preserves distances.

In fact, every isometry has one of the following forms:

a translation along a line in $\mathbb{R}^2$

a reflection in a line in $\mathbb{R}^2$;

a rotation about a point in $\mathbb{R}^2$;

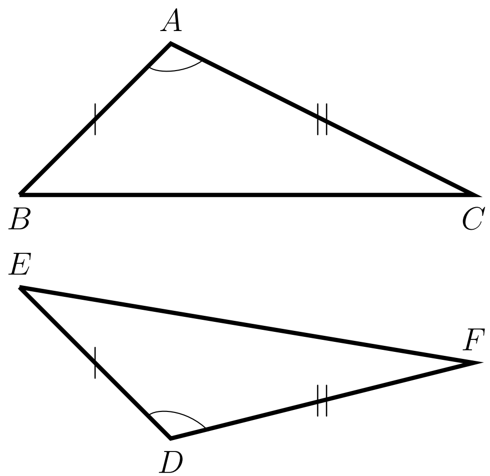
a composite of translations, reflections and rotations in $\mathbb{R}^2$.

The composite of any two isometries is also an isometry, and so it is easy to multiple of 2π. verify that the set $S(\mathbb{R}^2)$ of isometries of $\mathbb{R}^2$ forms a group under composition of functions. These observations can be used to build up the transformations we need in order to prove Euclidean results.

Example 2. Prove that if $\triangle ABC$ and ΔDEF are two triangles such that

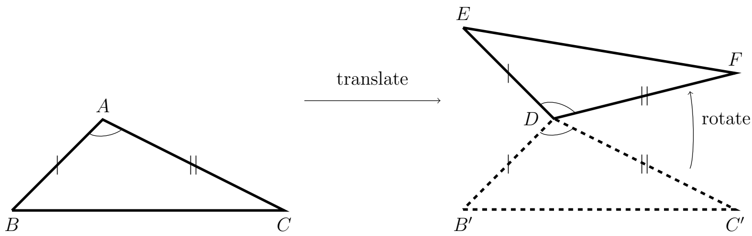
$$AB = DE, \quad AC = DF \text{ and } \angle BAC = \angle EDF$$

$$\text{then } BC = EF, \angle ABC = \angle DEF \text{ and } \angle ACB = \angle DFE.$$



Solution: It is sufficient to show that there is an isometry which maps $\triangle ABC$ onto ΔDEF. We construct this isometry in stages, starting with the translation which maps A to D.

This translation maps $\triangle ABC$ onto $\Delta DB'C'$, where B' and C' are the images of B and C under the translation.



Since we are given that $DF = AC$, and since the translation maps AC onto DC', it follows that $DF = DC'$. We can therefore rotate the point C', about D, until it coincides with the point F. This rotation maps $\Delta DB'C'$ onto $\Delta DB''F$, as shown in the margin, where B'' is the image of B' under the rotation.

Finally, notice that

$$\begin{aligned} \angle FDE &= \angle CAB && \text{(given)} \\ &= \angle C'DB' && \text{(translation)} \\ &= \angle FDB'' && \text{(rotation)} \end{aligned}$$

so either B'' lies on DE or the reflection of B'' in the line FD lies on DE. Also

$$\begin{aligned} DE &= AB && \text{(given)} \\ &= DB' && \text{(translation)} \\ &= DB'' && \text{(rotation)} \end{aligned}$$

It follows that either B'' coincides with E or the reflection of B'' in the line FD coincides with E.

So, composing the translation, the rotation, and (if necessary) a reflection, we obtain the required isometry that maps $\triangle ABC$ onto ΔDEF. Since isometries preserve length and angle, it follows that $BC = EF, \angle ABC = \angle DEF$ and $\angle ACB = \angle DFE$. $\square$

Problem 2. Prove that if $\triangle ABC$ and ΔDEF are two triangles such that

$$AC = DF, \quad \angle BAC = \angle EDF \text{ and } \angle ACB = \angle DFE,$$

then $BC = EF, AB = DE$ and $\angle ABC = \angle DEF$.

Solution: It is sufficient to show that there is an isometry which maps ΔABC onto ΔDEF. To construct this isometry, we start with the translation which maps A to D. This translation maps ΔABC onto $\Delta DB'C'$, where B' and C' are the images of B and C, respectively.

Since $DC' = AC = DF$, we can now rotate the point C' about D until it coincides with the point F. This rotation maps $\Delta DB'C'$ onto the triangle $\Delta DB''F$ shown below, where B'' is the image of B' under the rotation.

Now notice that

$$\begin{aligned}\angle EDF &= \angle BAC \quad \text{(given)}\\ &= \angle B'DC' \quad \text{(translation)}\\ &= \angle B''DF \quad \text{(rotation)},\end{aligned}$$

so either B'' lies on DE or the reflection of B'' in the line DF lies on DE. Also

$$\begin{aligned}\angle DFE &= \angle ACB \quad \text{(given)}\\ &= \angle DC'B' \quad \text{(translation)}\\ &= \angle DFB'' \quad \text{(rotation)},\end{aligned}$$

so either B'' lies on FE or the reflection of B'' in the line DF lies on FE. It follows that either B'' coincides with E or the reflection of B'' in the line DF coincides with E.

So, composing the translation, the rotation and (if necessary) a reflection, we obtain the required isometry which maps $\triangle ABC$ onto ΔDEF. Since isometries preserve length and angle, it follows that $BC = EF, AB = DE$ and $\angle ABC = \angle DEF$. $\square$

We can now answer the question 'What is Euclidean geometry?'. Euclidean geometry is the study of those properties of figures that are unchanged by the group of isometries. We call these properties Euclidean properties. Roughly plane $\mathbb{R}^2$. speaking, a Euclidean property is one that is preserved by a rigid figure as it moves around the plane. Of course, these properties include distance and angle, but they also include other properties such as collinearity of points and concurrence of lines. This idea, that geometry can be thought of in terms of a group of transformations acting on a space, is known as the Kleinian view of geometry. It enables us to generate many geometries, without losing sight of the relationship between them.

When we consider geometries in this way, it is often convenient to have an algebraic representation for the transformations involved. This not only enables us to solve problems in the geometry algebraically, but also provides us with formulas that can be used to compare different geometries.

In the case of Euclidean geometry, perhaps the easiest way to represent isometries algebraically is to use matrices. For example, the function defined by

$$t : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \qquad \left((x, y) \in \mathbb{R}^2\right) \quad (1)$$

is an isometry because it is the composite of an anticlockwise rotation through an angle θ about the origin, followed by a translation through the vector (e, f).

Similarly, the function

$$t : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \quad \left((x, y) \in \mathbb{R}^2\right) \quad (2)$$

is an isometry because it is the composite of a reflection in the line through the origin that makes an angle $\theta/2$ with the x -axis, followed by a translation through the vector (e, f)

Remarkably, we can represent any isometry by one or other of the forms given in (1) and (2). To see this, notice that any isometry t can be written in the form

$$t(\mathbf{x}) = t_0(\mathbf{x}) + (e, f) \quad \left(\mathbf{x} \in \mathbb{R}^2\right) \tag{3}$$

where t_0 is an isometry which fixes the origin. Indeed, if we let $(e, f) = t(\mathbf{0})$, then we can let t_0 be the transformation defined by $t_0(\mathbf{x}) = t(\mathbf{x}) - (e, f)$. This is an isometry because it is the composite of the isometry t and the translation through the vector $-(e, f)$. It fixes the origin since $t_0(\mathbf{0}) = t(\mathbf{0}) - (e, f) = \mathbf{0}$.

Now an isometry that fixes the origin must be either a rotation about the origin, or a reflection in a line through the origin. If t_0 is a rotation about the origin, then (3) can be written in the matrix form given in (1), whereas if t_0 is a reflection in a line through the origin, then (3) can be written in the matrix form given in (2).

So together, equations (1) and (2) provide us with an algebraic representation of all possible isometries of the plane. The next problem indicates how we can obtain a more concise description of this algebraic representation by using orthogonal matrices to combine equations (1) and (2).

Problem 3. Show that both the matrices

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

which appear in (1) and (2), are orthogonal for each real number θ.

Solution: Here we use the fact that a matrix $\mathbf{U}$ is orthogonal if $\mathbf{U}^T\mathbf{U} = \mathbf{I}$. We have

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

So both matrices are orthogonal for all real θ. $\square$

By applying the solution of Problem 3 to equations (1) and (2), we see that every isometry t has an algebraic representation of the form

$$t(\mathbf{x}) = \mathbf{U}\mathbf{x} + \mathbf{a}$$

where $\mathbf{U}$ is an orthogonal 2×2 matrix, and $\mathbf{a}$ is a vector in $\mathbb{R}^2$.

Definition. A Euclidean transformation of $\mathbb{R}^2$ is a function $t : \mathbb{R}^2 \to \mathbb{R}^2$ of the form

$$t(\mathbf{x}) = \mathbf{U}\mathbf{x} + \mathbf{a}$$

where $\mathbf{U}$ is an orthogonal 2×2 matrix and $\mathbf{a} \in \mathbb{R}^2$. The set of all Euclidean 1sformations of $\mathbb{R}^2$ is denoted by $E(2)$.

We may summarize the discussion above by saying that every isometry of the plane is a Euclidean transformation of $\mathbb{R}^2$.

In fact, the converse is also true, for if $\mathbf{U}$ is any orthogonal matrix, then its columns are orthonormal. In particular, its first and second columns have unit length and can therefore be written in the form $\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$ and $\begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}$, respectively, for some real θ, ϕ. For these to be orthonormal, we must have $\cos\theta \cdot \cos\phi + \sin\theta \cdot \sin\phi = 0$, so that $\tan\theta \cdot \tan\phi = -1$ and hence $\phi = \theta \pm \frac{\pi}{2}$. It follows that the second column must be

$$\begin{pmatrix} \cos(\theta + \pi/2) \\ \sin(\theta + \pi/2) \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} \qquad \text{or}$$
$$\begin{pmatrix} \cos(\theta - \pi/2) \\ \sin(\theta - \pi/2) \end{pmatrix} = \begin{pmatrix} \sin\theta \\ -\cos\theta \end{pmatrix}$$

So

$$\mathbf{U} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{or} \quad \mathbf{U} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

It follows that every Euclidean transformation $t(\mathbf{x}) = \mathbf{U}\mathbf{x} + \mathbf{a}$ of $\mathbb{R}^2$ has one of the forms given in equations (1) and (2). Since both of these forms represent isometries of the plane,we have the following theorem.

Theorem 1. Every isometry of $\mathbb{R}^2$ is a Euclidean transformation of $\mathbb{R}^2$ and vice versa.

Now the set of all isometries of $\mathbb{R}^2$ forms a group under composition of functions, so it follows from Theorem 1 that the same must be true of the set of all Euclidean transformations of $\mathbb{R}^2$. We therefore have the following theorem.

Theorem 2. The set of Euclidean transformations of $\mathbb{R}^2$ forms a group under the operation of composition of functions.

It is instructive to check the group axioms algebraically, for in the process of doing so we obtain formulas for the composites and inverses of Euclidean transformations.

We start by considering closure. Suppose that t_1 and t_2 are two Euclidean transformations given by

$$t_1(\mathbf{x}) = \mathbf{U}_1\mathbf{x} + \mathbf{a}_1 \quad \text{and} \quad t_2(\mathbf{x}) = \mathbf{U}_2\mathbf{x} + \mathbf{a}_2$$

where $\mathbf{U}_1$ and $\mathbf{U}_2$ are orthogonal 2×2 matrices. Then the composite $t_1 \circ t_2$ is given by

$$\begin{aligned} t_1 \circ t_2(\mathbf{x}) &= t_1\left(\mathbf{U}_2\mathbf{x} + \mathbf{a}_2\right) \\ &= \mathbf{U}_1\left(\mathbf{U}_2\mathbf{x} + \mathbf{a}_2\right) + \mathbf{a}_1 \\ &= \mathbf{U}_1\mathbf{U}_2\mathbf{x} + \left(\mathbf{U}_1\mathbf{a}_2 + \mathbf{a}_1\right) \end{aligned}$$

This is a Euclidean transformation since $\mathbf{U}_1\mathbf{U}_2$ is orthogonal. Indeed,

$$(\mathbf{U}_1\mathbf{U}_2)^T = \mathbf{U}_2^T\mathbf{U}_1^T = \mathbf{U}_2^{-1}\mathbf{U}_1^{-1} = (\mathbf{U}_1\mathbf{U}_2)^{-1}$$

So the set of Euclidean transformations is closed under composition of functions.

Problem 4. Let the Euclidean transformations t_1 and t_2 of $\mathbb{R}^2$ be given by

$$t_1(\mathbf{x}) = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}\mathbf{x} + \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

and

$$t_2(\mathbf{x}) = \begin{pmatrix} -\frac{4}{5} & \frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{pmatrix}\mathbf{x} + \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

Determine $t_1 \circ t_2$ and $t_2 \circ t_1$.

Solution: First, $t_1 \circ t_2(\mathbf{x})$ is equal to

$$
\begin{aligned}
& t_1\left(\left(\begin{array}{cc} -\frac{4}{5} & \frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{array}\right)\mathbf{x} + \left(\begin{array}{c} -2 \\ 1 \end{array}\right)\right) \\
&= \left(\begin{array}{cc} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{array}\right)\left(\left(\begin{array}{cc} -\frac{4}{5} & \frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{array}\right)\mathbf{x} + \left(\begin{array}{c} -2 \\ 1 \end{array}\right)\right) + \left(\begin{array}{c} 1 \\ -2 \end{array}\right) \\
&= \left(\begin{array}{cc} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{array}\right)\left(\begin{array}{cc} -\frac{4}{5} & \frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{array}\right)\mathbf{x} + \left(\begin{array}{c} -2 \\ -1 \end{array}\right) + \left(\begin{array}{c} 1 \\ -2 \end{array}\right) \\
&= \left(\begin{array}{cc} -\frac{24}{25} & -\frac{7}{25} \\ -\frac{7}{25} & \frac{24}{25} \end{array}\right)\mathbf{x} + \left(\begin{array}{c} -1 \\ -3 \end{array}\right).
\end{aligned}
$$

Next, $t_2 \circ t_1(\mathbf{x})$ is equal to

$$
\begin{aligned}
t_2\left(\left(\begin{array}{cc} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{array}\right)\mathbf{x} + \left(\begin{array}{c} 1 \\ -2 \end{array}\right)\right) &= \left(\begin{array}{cc} -\frac{4}{5} & \frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{array}\right)\left(\left(\begin{array}{cc} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{array}\right)\mathbf{x} + \left(\begin{array}{c} 1 \\ -2 \end{array}\right)\right) + \left(\begin{array}{c} -2 \\ 1 \end{array}\right). \\
&= \left(\begin{array}{cc} -\frac{4}{5} & \frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{array}\right)\left(\begin{array}{cc} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{array}\right)\mathbf{x} + \left(\begin{array}{c} -2 \\ -1 \end{array}\right) + \left(\begin{array}{c} -2 \\ 1 \end{array}\right). \\
&= \left(\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right)\mathbf{x} + \left(\begin{array}{c} -4 \\ 0 \end{array}\right) \qquad (6)
\end{aligned}
$$

□

Next recall that under composition of functions the identity is the transformation given by $i(\mathbf{x}) = \mathbf{x}$. This is a Euclidean transformation since it can be written in the form

$$i(\mathbf{x}) = \mathbf{I}\mathbf{x} + \mathbf{0}$$

where $\mathbf{I}$ is the 2×2 identity matrix, which is orthogonal.

The next problem asks you to show that inverses exist.

Problem 5. Prove that if t_1 is a Euclidean transformation of $\mathbb{R}^2$ given by

$$t_1(\mathbf{x}) = \mathbf{U}\mathbf{x} + \mathbf{a} \quad \left(\mathbf{x} \in \mathbb{R}^2\right)$$

then:

(a) the transformation of $\mathbb{R}^2$ given by

$$t_2(\mathbf{x}) = \mathbf{U}^{-1}\mathbf{x} - \mathbf{U}^{-1}\mathbf{a} \quad \left(\mathbf{x} \in \mathbb{R}^2\right)$$

is also a Euclidean transformation;

(b) the transformation t_2 is the inverse of t_1.

The solution of Problem 5 shows that we can calculate the inverse of a Euclidean transformation by using the following result.

The inverse of the Euclidean transformation $t(\mathbf{x}) = \mathbf{U}\mathbf{x} + \mathbf{a}$ is given by

$$t^{-1}(\mathbf{x}) = \mathbf{U}^{-1}\mathbf{x} - \mathbf{U}^{-1}\mathbf{a}.$$

Solution:

(a) Since $\mathbf{U}$ is an orthogonal matrix, it follows that $\mathbf{U}^{-1} = \mathbf{U}^{\mathrm{T}}$. Taking the transpose of both sides, we have

$$\left(\mathbf{U}^{-1}\right)^T = \left(\mathbf{U}^T\right)^T = \mathbf{U} = \left(\mathbf{U}^{-1}\right)^{-1}$$

Thus $\mathbf{U}^{-1}$ is an orthogonal matrix, and so t_2 is a Euclidean transformation.

(b) We have

$$\begin{aligned} t_1 \circ t_2(\mathbf{x}) &= t_1\left(\mathbf{U}^{-1}\mathbf{x} - \mathbf{U}^{-1}\mathbf{a}\right) \\ &= \mathbf{U}\left(\mathbf{U}^{-1}\mathbf{x} - \mathbf{U}^{-1}\mathbf{a}\right) + \mathbf{a} \\ &= (\mathbf{x} - \mathbf{a}) + \mathbf{a} \\ &= \mathbf{x} \end{aligned}$$

and

$$\begin{aligned} t_2 \circ t_1(\mathbf{x}) &= t_2(\mathbf{U}x + a) \\ &= \mathbf{U}^{-1}(\mathbf{U}\mathbf{x} + \mathbf{a}) - \mathbf{U}^{-1}\mathbf{a} \\ &= \left(\mathbf{x} + \mathbf{U}^{-1}\mathbf{a}\right) - \mathbf{U}^{-1}\mathbf{a} \\ &= \mathbf{x} \end{aligned}$$

so t_2 is the inverse of t_1.

□

Problem 6. Determine the inverse of the Euclidean transformation given by

$$t(\mathbf{x}) = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Solution: We have

$$\begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix}$$

and

$$\begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$$

so that

$$t^{-1}(\mathbf{x}) = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

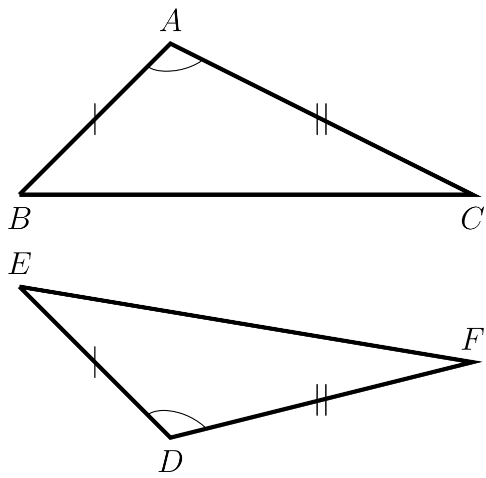
$\square$

Finally, composition of functions is always associative. So all four group properties hold, as we expected.

Earlier, we described Euclidean geometry as the study of those properties of figures that are preserved by isometries. Having identified these isometries with the group of Euclidean transformations, we can now give the equivalent algebraic description of Euclidean geometry. Euclidean geometry is the study of those properties of figures that are preserved by Euclidean transformations of $\mathbb{R}^2$.

2.1.2 Euclidean-Congruence

In the solution to Example 2 we showed that if two triangles $\triangle ABC$ and $\triangle DEF$ are such that $AB = DE, AC = DF$ and $\angle BAC = \angle EDF$, then there is a Euclidean transformation which maps $\triangle ABC$ onto ΔDEF.



The existence of this transformation enabled us to deduce that both triangles have the same Euclidean properties. In particular, we were able to deduce that $BC = EF, \angle ABC = \angle DEF$ and $\angle ACB = \angle DFE$.

In order to formalize this way of relating two figures, we say that two figures are congruent if one can be moved to fill exactly the position of the other by means of a Euclidean transformation. Loosely speaking, two figures are congruent if they have the same size and shape.

Later we consider congruence with respect to other groups of transformations (that is, congruence in other geometries), so if there is any danger of confusion we sometimes say that two figures are Euclidean-congruent.

Definition. A figure F_1 is Euclidean-congruent to a figure F_2 if there is a Euclidean transformation which maps F_1 onto F_2

For example, any two circles of unit radius are Euclidean-congruent to each other because we can map one of the circles onto the other by means of a translation that makes their centres coincide.

Problem 7. Which of the following sets consist of figures that are Euclidean-congruent to each other?

(a) The set of all ellipses

(b) The set of all line segments of length 1

(c) The set of all triangles

(d) The set of all squares that have sides of length 2

Solution:

(a) Not congruent

(b) Congruent

(c) Not congruent

(d) Congruent

□

Earlier, we emphasized that the Euclidean transformations form a group. This is important because it ensures that Euclidean-congruence has the kind of properties that we should expect. For example, we should expect every figure to be congruent to itself. Also, if a figure F_1 is congruent to a figure F_2, then we should expect F_2 to be congruent to F_1. We can, in fact, establish the following result.

Theorem 3. Euclidean-congruence is an equivalence relation.

Proof: We show that the three equivalence relation axioms E1, E2 and E3 hold.

E1 REFLEXIVE: For all figures F in $\mathbb{R}^2$, the identity transformation maps F onto itself; so Euclidean-congruence is reflexive.

E2 SYMMETRIC: Let a figure F_1 in $\mathbb{R}^2$ be congruent to a figure F_2, and let t be a Euclidean transformation which maps F_1 onto F_2. Then the inverse Euclidean transformation t^{-1}

maps F_2 onto F_1, so that F_2 is congruent to F_1. Thus Euclidean-congruence is symmetric.

E3 TRANSITIVE: Let a figure F_1 in $\mathbb{R}^2$ be congruent to a figure F_2, and let F_2 be congruent to a figure F_3. Then there exist Euclidean transformations t_1 mapping F_1 onto F_2 and t_2 mapping F_2 onto F_3. Thus the Euclidean transformation $t_2 \circ t_1$ maps F_1 onto F_3, so that F_1 is congruent to F_3. Hence Euclidean-congruence is transitive.

It follows that Euclidean-congruence is an equivalence relation, because it satisfies the axioms E1, E2 and E3. $\square$

Problem 8. Prove that if two figures in $\mathbb{R}^2$ are each Euclidean-congruent to a third figure, then they are Euclidean-congruent to each other.

Solution: Suppose that we are given three plane figures F_1, F_2 and F_3 such that

$$F_1 \text{ is congruent to } F_3 \qquad (*)$$

and

$$F_2 \text{ is congruent to } F_3. \qquad (**)$$

It follows from $(**)$ and the symmetric property of congruence that

$$F_3 \text{ is congruent to } F_2. \qquad (***)$$

Hence from $(*)$ and $(* * *)$ and the transitive property of congruence, F_1 is congruent to F_2, as required. $\square$

Since Euclidean-congruence is an equivalence relation, it partitions the set of all figures into disjoint equivalence classes. Each class consists of figures which are Euclidean-congruent to each other, and hence share the same Euclidean properties (for example, one class consists of all circles of unit radius, another class consists of all equilateral triangles with sides of length 3 , and so on). If we wish to show that two figures have the same Euclidean properties, then it is sufficient to show that they are Euclidean-congruent.

Now Euclidean geometry is just one of several different geometries. Each geometry is defined by a group G of transformations that act on a space. In general, we say that two figures are G -congruent if there is a transformation in G which maps one of the figures onto the other. Since the only properties used in the proof of Theorem 3 are the group properties of Euclidean transformations, the theorem holds also with ' G -congruent' in place of 'Euclidean-congruent'. Thus, like Euclidean-congruence, G -congruence is an equivalence relation that partitions the set of all figures into disjoint equivalence classes.

This idea of partitioning figures into equivalence classes is central to geometry. It enables us to distinguish between figures in different equivalence classes, without having to worry about the differences between figures in the same equivalence class. For example, if we are interested in whether a conic is an ellipse rather than a hyperbola or a parabola, but do not care about its shape (that is, the ratio of the lengths of its axes), we might choose to work with some geometry whose group of transformations makes all ellipses congruent to each other - but not congruent to any hyperbola or parabola. We describe a group of transformations which defines such a geometry in Section 2.2.

2.2 Affine Transformations and Parallel Projections

2.2.1 Affine Transformations

In Section 2.1 you met a new approach to Euclidean geometry in $\mathbb{R}^2-$ namely, the idea that Euclidean geometry of $\mathbb{R}^2$ can be interpreted as a space, $\mathbb{R}^2$, together with the group of Euclidean transformations which act on that space. Recall that a Euclidean transformation is a function $t : \mathbb{R}^2 \to \mathbb{R}^2$ of the form

$$t(\mathbf{x}) = \mathbf{U}\mathbf{x} + \mathbf{a} \quad \left(\mathbf{x} \in \mathbb{R}^2\right)$$

where $\mathbf{U}$ is an orthogonal 2×2 matrix. Euclidean properties of figures are those, like distance and angle, that are preserved by these transformations.

In this section we meet the first of our new geometries in $\mathbb{R}^2-$ affine geometry. This geometry consists of the space $\mathbb{R}^2$ together with a group of transformations, the affine transformations, acting on $\mathbb{R}^2$.

Definition. An affine transformation of $\mathbb{R}^2$ is a function $t : \mathbb{R}^2 \to \mathbb{R}^2$ of the form

$$t(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b},$$

where $\mathbf{A}$ is an invertible 2×2 matrix and $b \in \mathbb{R}^2$. The set of all affine transformations of $\mathbb{R}^2$ is denoted by $A(2)$.

Remark

Note that every Euclidean transformation of $\mathbb{R}^2$ is an affine transformation of $\mathbb{R}^2$ since every orthogonal matrix is invertible. (In terms of groups, the group of Euclidean transformations of $\mathbb{R}^2$ is a proper subgroup of the group of affine transformations of $\mathbb{R}^2$.) This means that all properties of figures that are preserved by affine transformations must be preserved also by Euclidean transformations.

Problem 1. Determine whether or not each of the following transformations of $\mathbb{R}^2$ is an affine transformation.

(a) $t_1(\mathbf{x}) = \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 4 \\ -2 \end{pmatrix}$ (b) $t_2(\mathbf{x}) = \begin{pmatrix} -6 & 5 \\ 3 & 2 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

(c) $t_3(\mathbf{x}) = \begin{pmatrix} -2 & -1 \\ 8 & 4 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ (d) $t_4(\mathbf{x}) = \begin{pmatrix} 5 & -3 \\ -2 & 2 \end{pmatrix}\mathbf{x}$

Solution: We use the fact that a 2×2 matrix is invertible if and only if its determinant is non-zero. Each transformation is of the form

$$\mathbf{x} \mapsto \mathbf{A}\mathbf{x} + \mathbf{b}$$

where $\mathbf{A}$ is a 2×2 matrix, and so it is an affine transformation if and only if the determinant of the matrix $\mathbf{A}$ is non-zero.

(a) Here,

$$\begin{vmatrix} 1 & 3 \\ 1 & 2 \end{vmatrix} = 2 - 3 = -1$$

which is non-zero; hence t_1 is an affine transformation.

(b) Here,

$$\begin{vmatrix} -6 & 5 \\ 3 & 2 \end{vmatrix} = -12 - 15 = -27$$

which is non-zero; hence t_2 is an affine transformation.

(c) Here,

$$\begin{vmatrix} -2 & -1 \\ 8 & 4 \end{vmatrix} = -8 + 8 = 0;$$

hence t_3 is not an affine transformation.

(d) Here, $\mathbf{b} = \mathbf{0}$ and

$$\begin{vmatrix} 5 & -3 \\ -2 & 2 \end{vmatrix} = 10 - 6 = 4$$

which is non-zero; hence t_4 is an affine transformation.

$\square$

The algebra required to compose affine transformations is similar to the algebra that we used to compose Euclidean transformations.

Problem 2. For the transformations of $\mathbb{R}^2$ given in Problem 1, determine formulas for the following composites. In each case, state whether or not the composite is an affine transformation.

(a) $t_1 \circ t_2$ (b) $t_2 \circ t_4$

Solution:

(a) Here, $t_1 \circ t_2(\mathbf{x})$ is equal to

$$\begin{aligned} t_1\left(\begin{pmatrix} -6 & 5 \\ 3 & 2 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right) &= \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix}\left(\begin{pmatrix} -6 & 5 \\ 3 & 2 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right) + \begin{pmatrix} 4 \\ -2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} -6 & 5 \\ 3 & 2 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 5 \\ 4 \end{pmatrix} + \begin{pmatrix} 4 \\ -2 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 11 \\ 0 & 9 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 9 \\ 2 \end{pmatrix}. \end{aligned}$$

Since

$$\begin{vmatrix} 3 & 11 \\ 0 & 9 \end{vmatrix} = 27 - 0 = 27 \neq 0$$

it follows that $t_1 \circ t_2$ is an affine transformation.

(b) Here, $t_2 \circ t_4(\mathbf{x})$ is equal to

$$t_2\left(\begin{pmatrix} 5 & -3 \\ -2 & 2 \end{pmatrix}\mathbf{x}\right) = \begin{pmatrix} -6 & 5 \\ 3 & 2 \end{pmatrix}\left(\begin{pmatrix} 5 & -3 \\ -2 & 2 \end{pmatrix}\mathbf{x}\right) + \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$
$$= \begin{pmatrix} -40 & 28 \\ 11 & -5 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Since

$$\begin{vmatrix} -40 & 28 \\ 11 & -5 \end{vmatrix} = 200 - 308 = -108 \neq 0$$

it follows that $t_2 \circ t_4$ is an affine transformation.

$\square$

We now verify our assertion above that the set of affine transformations forms a group.

Theorem 1. The set of affine transformations $A(2)$ forms a group under the operation of composition of functions.

Proof: We check that the four group axioms hold.

G1 CLOSURE: Let t_1 and t_2 be affine transformations given by

$$t_1(\mathbf{x}) = \mathbf{A}_1\mathbf{x} + \mathbf{b}_1 \text{ and } t_2(\mathbf{x}) = \mathbf{A}_2\mathbf{x} + \mathbf{b}_2$$

where $\mathbf{A}_1$ and $\mathbf{A}_2$ are invertible 2×2 matrices. Then, for

each $\mathbf{x} \in \mathbb{R}^2$,

$$\begin{aligned}(t_1 \circ t_2)(\mathbf{x}) &= t_1(\mathbf{A}_2\mathbf{x} + \mathbf{b}_2) \\ &= \mathbf{A}_1(\mathbf{A}_2\mathbf{x} + \mathbf{b}_2) + \mathbf{b}_1 \\ &= (\mathbf{A}_1\mathbf{A}_2)\mathbf{x} + (\mathbf{A}_1\mathbf{b}_2 + \mathbf{b}_1).\end{aligned}$$

Since $\mathbf{A}_1$ and $\mathbf{A}_2$ are invertible, it follows that $\mathbf{A}_1\mathbf{A}_2$ is also invertible. So by definition $t_1 \circ t_2$ is an affine transformation.

G2 IDENTITY: Let i be the affine transformation given by

$$i(\mathbf{x}) = \mathbf{I}\mathbf{x} + \mathbf{0} \quad \left(\mathbf{x} \in \mathbb{R}^2\right)$$

where $\mathbf{I}$ is the 2×2 identity matrix. If t is an affine transformation given by

$$t(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b} \quad \left(\mathbf{x} \in \mathbb{R}^2\right),$$

then, for each $\mathbf{x} \in \mathbb{R}^2$,

$$(t \circ i)(\mathbf{x}) = \mathbf{A}(\mathbf{I}\mathbf{x} + \mathbf{0}) + \mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{b} = t(\mathbf{x})$$

and

$$(i \circ t)(\mathbf{x}) = \mathbf{I}(\mathbf{A}\mathbf{x} + \mathbf{b}) + \mathbf{0} = \mathbf{A}\mathbf{x} + \mathbf{b} = t(\mathbf{x})$$

Thus $t \circ i = i \circ t = t$. Hence i is the identity transformation.

G3 INVERSES: If t is an arbitrary affine transformation given by

then we can define another affine transformation t' by

$$t'(\mathbf{x}) = \mathbf{A}^{-1}\mathbf{x} - \mathbf{A}^{-1}\mathbf{b}$$

$$t(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b} \quad \left(\mathbf{x} \in \mathbb{R}^2\right)$$

Now for each $\mathbf{x} \in \mathbb{R}^2$, we have

$$\begin{aligned}
(t \circ t')(\mathbf{x}) &= t\left(\mathbf{A}^{-1}\mathbf{x} - \mathbf{A}^{-1}\mathbf{b}\right) \\
&= \mathbf{A}\left(\mathbf{A}^{-1}\mathbf{x} - \mathbf{A}^{-1}\mathbf{b}\right) + \mathbf{b} \\
&= \left(\mathbf{A}\mathbf{A}^{-1}\mathbf{x} - \mathbf{A}\mathbf{A}^{-1}\mathbf{b}\right) + \mathbf{b} \\
&= (\mathbf{x} - \mathbf{b}) + \mathbf{b}
\end{aligned}$$

$= \mathbf{x}$ Also,

$$\begin{aligned}
(t' \circ t)(\mathbf{x}) &= t'(\mathbf{A}\mathbf{x} + \mathbf{b}) \\
&= \mathbf{A}^{-1}(\mathbf{A}\mathbf{x} + \mathbf{b}) - \mathbf{A}^{-1}\mathbf{b} \\
&= \left(\mathbf{A}^{-1}\mathbf{A}\mathbf{x} + \mathbf{A}^{-1}\mathbf{b}\right) - \mathbf{A}^{-1}\mathbf{b} \\
&= \left(\mathbf{x} + \mathbf{A}^{-1}\mathbf{b}\right) - \mathbf{A}^{-1}\mathbf{b} \\
&= \mathbf{x}.
\end{aligned}$$

Thus $t \circ t' = t' \circ t = i$. Hence t' is an inverse for t.

G4 ASSOCIATIVITY: Composition of functions is always associative.

It follows that the set of affine transformations $A(2)$ forms a group under composition of functions. $\square$

The above proof shows that we can calculate the inverse of an affine transformation by using the following result.

The inverse of the affine transformation $t(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ is given by

$$t^{-1}(\mathbf{x}) = \mathbf{A}^{-1}\mathbf{x} - \mathbf{A}^{-1}\mathbf{b}.$$

Problem 3. Find the inverse of the affine transformation

$$t(\mathbf{x}) = \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 4 \\ -2 \end{pmatrix}$$

Solution: The inverse of a 2×2 matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Hence

$$\begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}$$

and

$$\begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 4 \\ -2 \end{pmatrix} = \begin{pmatrix} -14 \\ 6 \end{pmatrix},$$

so that

$$t^{-1}(\mathbf{x}) = \begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 14 \\ -6 \end{pmatrix}.$$

□

Having shown that the set of affine transformations forms a group under composition of functions, we now define affine geometry to be the study of those properties of figures in the plane $\mathbb{R}^2$ that are preserved by affine transformations. These are the so-called affine properties of figures. We begin our investigation of affine geometry by considering the three affine properties listed below.

Basic Properties of Affine Transformations

Affine transformations:

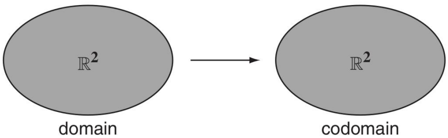
1. map straight lines to straight lines;
2. map parallel straight lines to parallel straight lines;
3. preserve ratios of lengths along a given straight line.

There are two approaches that we shall use to investigate these properties. One approach is to use the definition of an affine transformation to investigate the properties algebraically;

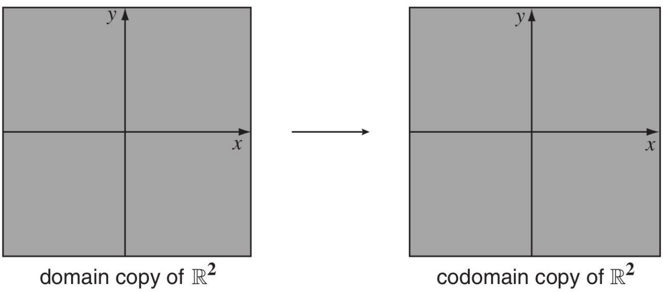
we do this in Section 2.3. First, however, we investigate the properties geometrically. We begin to do this in the next subsection by introducing a special type of affine transformation for which there is a simple geometric interpretation.

2.2.2 Parallel Projections

A parallel projection is a one-one mapping from $\mathbb{R}^2$ onto itself, defined in the following way. First, we think of its domain and codomain as two separate copies of $\mathbb{R}^2$.

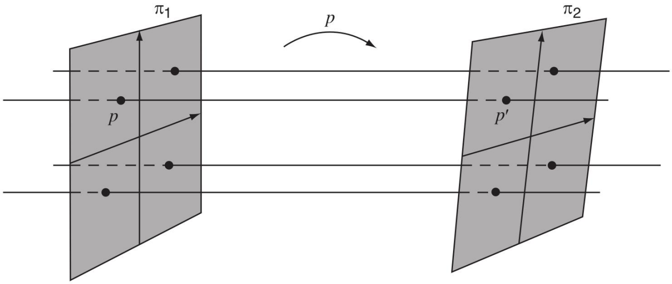


Geometrically, we can represent these copies of $\mathbb{R}^2$ by two separate planes, each equipped with a pair of rectangular axes.



Next we place these planes into three-dimensional space; we denote the domain plane by π_1 and the codomain plane by π_2.

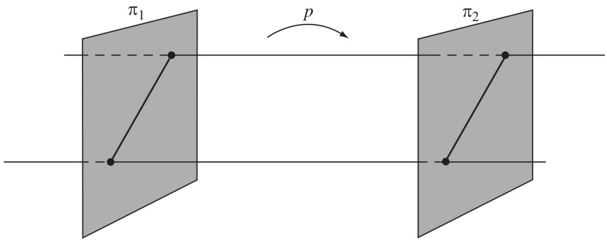
Now imagine parallel rays of light shining through π_1 and π_2. Each point P in the plane π_1 has a (unique) ray passing through it, that also passes through a point P', say, in the plane π_2. This provides us with a one-one correspondence between points in the two planes π_1 and π_2. We call the function p which maps each point P in π_1 to the corresponding point P' in π_2 a parallel projection from π_1 onto π_2.



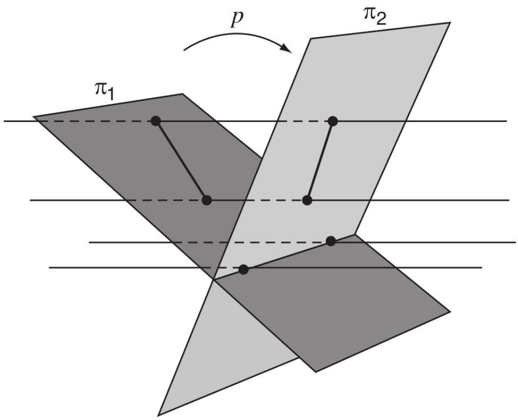
If the roles of the planes π_1 and π_2 are reversed, so that π_2 becomes the domain plane and π_1 becomes the codomain plane, then we obtain the inverse function p^{-1} which maps points P' in π_2 back to the corresponding points P in π_1. Clearly, p^{-1} is a parallel projection of π_2 onto π_1.

Each choice of location for the domain plane π_1, and the codomain plane π_2. and each choice of direction for the rays of light, yields a parallel projection. The only constraint is that the rays of light must not be parallel to either plane.

If the planes π_1 and π_2 are parallel to each other, then any parallel projection p from π_1 onto π_2 is an isometry, since the distance between any two points is unaltered.



On the other hand, if the planes are not parallel to each other, then some distances are changed under the projection, and so the parallel projection is not an isometry; notice, however, that distances along the line of intersection of the planes π_1 and π_2 do remain unchanged by the parallel projection.



Although distances are not always preserved by a parallel projection, there are some basic properties that are preserved; three of these are listed below. As you will see, these are the same as the basic affine properties that we mentioned at the end of Subsection 2.2.1.

Basic Properties of Parallel Projections

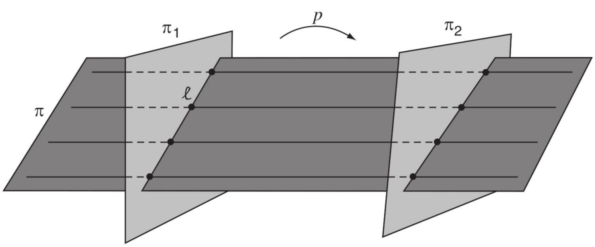
Parallel projections:

1. map straight lines to straight lines;
2. map parallel straight lines to parallel straight lines:
3. preserve ratios of lengths along a given straight line.

Later, we will show that each basic affine property follows directly from the corresponding property for parallel projections. In anticipation of this, we first show that the properties hold for parallel projections.

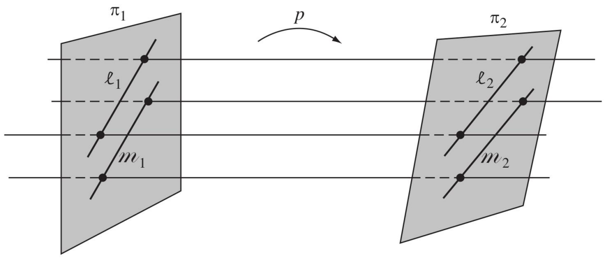
Property 1 A parallel projection maps straight lines to straight lines.

Proof: Let ℓ be a line in the plane π_1, and let p be a parallel projection mapping π_1 onto the plane π_2. Now consider all the rays associated with p that pass through ℓ. Since these rays are parallel, they must fill a plane. Call this plane π.



The image of ℓ under p consists of those points where the rays that pass through ℓ meet π_2. But these points are simply the points of intersection of π with π_2. Since any two intersecting planes in R^3 meet in a line, it follows that the image of ℓ under p is a straight line. $\square$

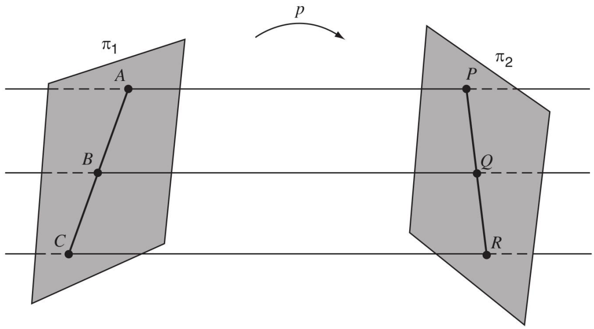
Property 2 A parallel projection maps parallel straight lines to parallel straight lines.



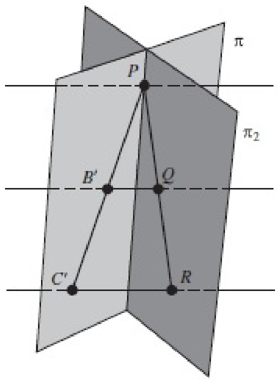
Proof: Let ℓ_1 and m_1 be parallel lines in the plane π_1, and let p be a parallel projection mapping π_1 onto the plane π_2. Let ℓ_2 and m_2 be the lines in π_2 that are the images under p of ℓ_1 and m_1.

If ℓ_2 and m_2 are not parallel, they meet at some point, P_2 say. Let P_1 be the point of π_1 which maps to P_2. Then P_1 must lie on both ℓ_1 and m_1. Since ℓ_1 and m_1 are parallel, no such point of intersection can exist, which is a contradiction. It follows that ℓ_2 and m_2 must indeed be parallel. $\square$

Property 3 A parallel projection preserves ratios of lengths along a given straight line.



Proof: Let A, B, C be three points on a line in the plane π_1, and let p be a parallel projection mapping π_1 onto the plane π_2. Let P, Q, R be the points in π_2 that are the images under p of A, B, C. We know from Property 1 that P, Q, R lie on a line; we have to show that the ratio $AB : AC$ is equal to the ratio $PQ : PR$ If the planes π_1 and π_2 are parallel, then the parallel projection p is an isometry, and so the ratios $AB : AC$ and $PQ : PR$ are equal, as required.



On the other hand, if π_1 and π_2 are not parallel, then we can construct a plane π through the point P which is parallel to π_1, as shown in the margin. This plane intersects the ray through B and Q at some point B', and the ray through C and R at some point C'. So in this case the ratios $AB : AC$ and $PB' : PC'$ are equal. Now consider $\triangle PC'R$. The lines $B'Q$ and $C'R$ are parallel, since they are rays from the parallel projection. Hence $B'Q$ meets the sides PR and PC' in equal ratios. Thus $PQ : PR = PB' : PC'$. It follows that $PQ : PR = AB : AC$, as required. □

Notice, in particular, that if a point is the midpoint of a line segment, then under a parallel projection the image of the point is the midpoint of the image of the line segment.

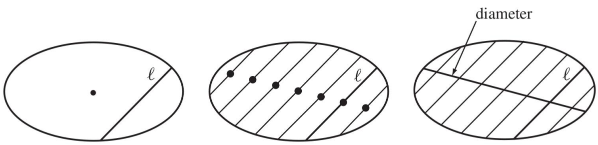
In Subsection 2.2.3 you will see why the basic properties of affine transformations and of parallel projections are the same, and you will meet some further properties of each.

2.2.3 Affine Geometry

In this subsection we explore further the ideas of affine geometry and of parallel projection in order to prove two attractive and unexpected results about ellipses. Also, we examine the relationship between affine transformations and parallel projections.

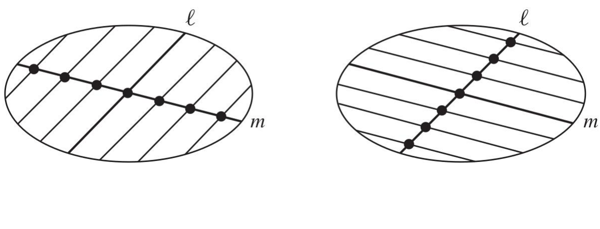
Two Results about Ellipses

First, starting with any chord ℓ of an ellipse, draw all the chords parallel to ℓ and construct their midpoints. We claim that these midpoints lie on a chord through the centre of the ellipse - that is, on a diameter of the ellipse.



Theorem 2. (Midpoint Theorem) Let ℓ be a chord of an ellipse. Then the midpoints of the chords parallel to ℓ lie on a diameter of the ellipse.

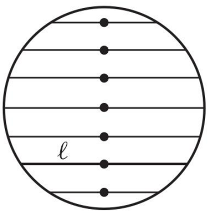
Next, start with any diameter ℓ of an ellipse and construct a second diameter m by following the construction used in Theorem 2, as shown below. Then repeat the construction starting this time with the diameter m; this might reasonably be expected to give us a third diameter of the ellipse - but, surprisingly, it gives us the diameter ℓ with which we started.

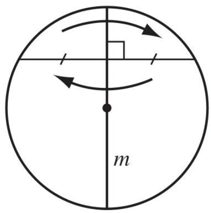


Theorem 3. (Conjugate Diameters Theorem) Let ℓ be a diameter of an ellipse. Then there is another diameter m of the ellipse such that:
(a) the midpoints of all chords parallel to ℓ lie on m;
(b) the midpoints of all chords parallel to m lie on ℓ.

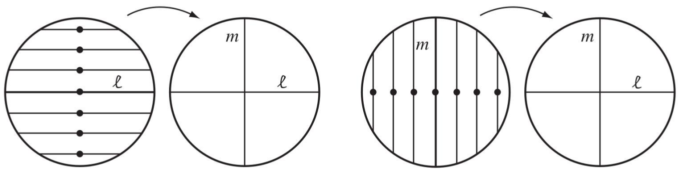
Proofs for the Special Case of a Circle

We now investigate these theorems for the special case when the ellipse is a circle. To prove the Midpoint Theorem in this case, start with a chord ℓ. If necessary, rotate the circle to ensure that ℓ is horizontal. It is then sufficient to prove that every horizontal chord is bisected by the vertical diameter, m.





To do this note that the circle is symmetrical about m; so, reflection in m maps that part of every horizontal chord to the left of m exactly onto the part to the right of m. Since reflection preserves length, these two parts must be the same length; in other words, m bisects each horizontal chord, as required. What about the Conjugate Diameters Theorem for the special case of the circle?



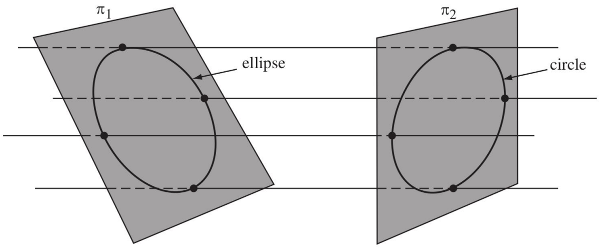
Start with the horizontal diameter ℓ, and carry out the construction of another diameter as in Theorem 2; this yields the vertical diameter m. If we then start with the vertical diameter m and repeat the construction, we obtain ℓ, the horizontal diameter of the circle. So Theorem 3 certainly holds when the ellipse is a circle.

Generalizing the Proof

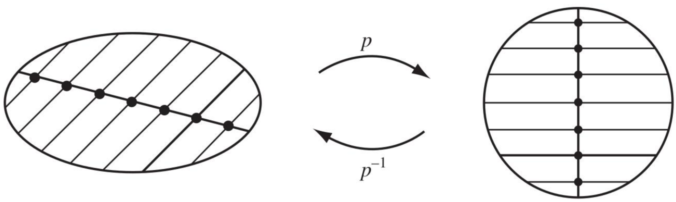
We now investigate how the proofs of Theorems 2 and 3 for the circle can be turned into proofs for any kind of ellipse. The crucial fact is as follows.

Theorem 4. Given any ellipse, there is a parallel projection which maps the ellipse onto a circle.

A suitable parallel projection is illustrated below. Here the plane π_1 (initially parallel to π_2) has been tilted about the minor axis of the ellipse. Under the projection distances which are parallel to the minor axis remain unchanged, but distances parallel to the major axis are scaled by a factor which depends on. the 'angle of tilt'. By choosing just the right amount of tilt we can ensure that the image of the major axis is equal in length to the image of the minor axis, thereby ensuring that the image of the ellipse is a circle.



Both Theorems 2 and 3 may now be proved using the following technique. First, map the given ellipse onto a circle, using a suitable parallel projection p. Since we have seen that the theorems hold in the case of the circle, we then map the circle back to the ellipse, using the inverse parallel projection p^{-1}. Now collinearity and parallelism are preserved under a parallel projection, as is the property of being the midpoint of a line segment, so the above two theorems, which hold for a circle, must hold also for the ellipse.



Notice that certain properties of figures, such as length and angle, are not preserved under a parallel projection. This is one difference between Euclidean geometry and affine geometry. The difference arises because the group of affine transformations is larger than the group of Euclidean transformations. In general, the larger the group that is used to define a geometry, the fewer properties the geometry has.

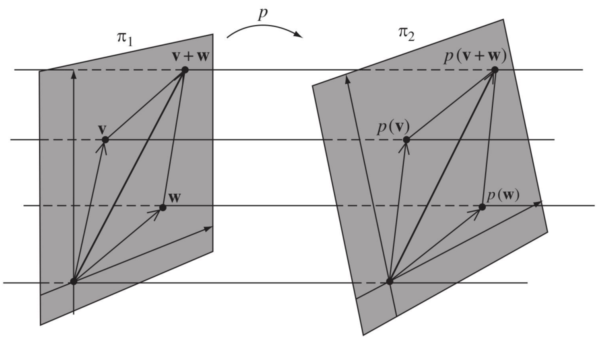
Affine Transformations and Parallel Projections

Earlier we mentioned that a parallel projection is a special type of affine transformation. We now show why this is indeed the case.

First, consider a parallel projection p of a plane π_1 onto a plane π_2. For the moment, suppose that the planes are aligned so that the origin in π_1 is mapped to the origin in π_2. Since ratios of lengths are preserved along a straight line, we must have, for any vector $\mathbf{v} \in \mathbb{R}^2$ and any $\lambda \in \mathbb{R}$,

$$p(\lambda \mathbf{v}) = \lambda p(\mathbf{v}) \tag{1}$$

Next, let $\mathbf{v}$ and $\mathbf{w}$ be two position vectors in π_1. Their sum, $\mathbf{v} + \mathbf{w}$, is found from the Parallelogram Law for addition of vectors, as shown in the diagram below. The images under p in π_2 are $p(\mathbf{v})$ and $p(\mathbf{w})$, and the sum of these two vectors is $p(\mathbf{v}) + p(\mathbf{w})$.



But a parallel projection maps parallel lines onto parallel lines, so it must map parallelograms onto parallelograms. Hence it must map the parallelogram in π_1 onto the parallelogram in π_2, and, in particular, it must map $\mathbf{v}+\mathbf{w}$ to $p(\mathbf{v})+p(\mathbf{w})$. We may write this as

$$p(\mathbf{v}+\mathbf{w}) = p(\mathbf{v}) + p(\mathbf{w}) \tag{2}$$

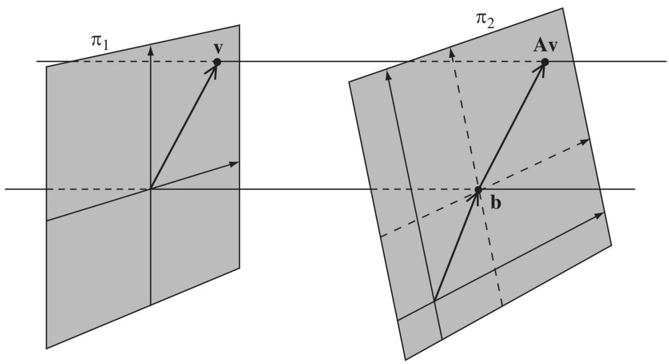
It follows from equations (1) and (2) that p must be a linear transformation of $\mathbb{R}^2$ onto itself. Hence there exists some matrix $\mathbf{A}$ such that for each $\mathbf{v} \in \mathbb{R}^2$

$$p(\mathbf{v}) = \mathbf{A}\mathbf{v}. \tag{3}$$

Since the linear transformation p is invertible, it follows that A is invertible. Now suppose that the parallel projection maps the origin in π_1 to some point B with position vector $\mathbf{b}$ in π_2, as shown below. If we temporarily construct a new set of axes in π_2 that are parallel to the original axes, but which intersect at the point B, then with respect to these new axes $p(\mathbf{v}) = \mathbf{Av}$ for some invertible matrix $\mathbf{A}$, as before. To express $p(\mathbf{v})$ with respect to the original axes, we simply add on the vector $\mathbf{b}$ to obtain

$$p(\mathbf{v}) = \mathbf{Av} + \mathbf{b} \tag{4}$$

for some invertible 2×2 matrix $\mathbf{A}$.



It follows from equation (4) that p must be an affine transformation.

> **Theorem 5.** Each parallel projection is an affine transformation.

The converse is false, for it is not true that every affine transformation can be represented as a parallel projection.

For example, consider the so-called 'doubling map' of $\mathbb{R}^2$ to itself given by

$$t(\mathbf{v}) = 2\mathbf{v} \quad \left(\mathbf{v} \in \mathbb{R}^2\right) \tag{5}$$

This is an affine transformation, since it can be written in the form $t(\mathbf{x}) = \mathbf{Ax}+\mathbf{b}$ with $\mathbf{A} = 2\mathbf{I}$ and $\mathbf{b} = \mathbf{0}$. However, a parallel projection is either between two parallel planes, in which case all lengths are unchanged, or between two intersecting planes, in which case distances along the line of intersection are unchanged. The doubling map has neither of these properties and so is not a parallel projection.

Observation An affine transformation is not necessarily a paralle projection.

Although the doubling map is not a parallel projection, it is possible to double lengths in $\mathbb{R}^2$ by following one parallel projection by another: the first doubles all horizontal lengths, and the second doubles all vertical lengths. Thus the doubling map (5) can be represented as the composition of two parallel projections.

We end this subsection by showing that every affine transformation can be expressed as a composition of two parallel projections. Recall that any affine transformation $t : \mathbb{R}^2 \to \mathbb{R}^2$ has the form

$$t(\mathbf{x}) = \mathbf{Ax} + \mathbf{b} \quad \left(\mathbf{x} \in \mathbb{R}^2\right) \tag{6}$$

where $\mathbf{A}$ is an invertible 2×2 matrix. Now, t is not a linear transformation unless $\mathbf{b} = \mathbf{0}$, but we can use methods similar to those for linear transformations to determine $\mathbf{A}$ and $\mathbf{b}$.

First, it follows from equation (6) that $t(0) = \mathbf{b}$; so $\mathbf{b}$ is the image of the origin under t. If we let e and f be the coordinates of $t(0)$, then we can write

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} e \\ f \end{pmatrix}$$

where a, b, c, d are real numbers that have yet to be found. It follows from equation (6) that the images under t of the points $(1,0)$ and $(0,1)$ are given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}$$

and

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}.$$

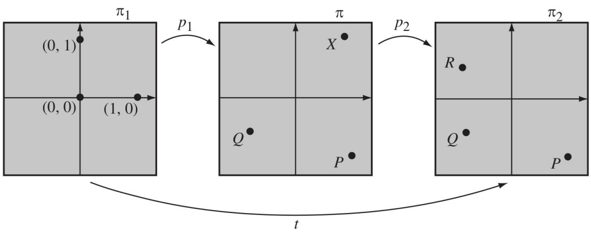
So if, in addition to $t(0) = (e, f)$, we know the points onto which $(1,0)$ and $(0,1)$ are mapped by t, then we can determine the values of a, b, c and d. Indeed, we have

$$(a, c) = t(1,0) - (e, f) \quad \text{and} \quad (b, d) = t(0,1) - (e, f)$$

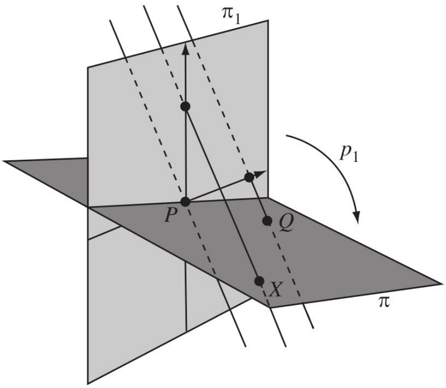
It follows that an affine transformation is uniquely determined by its effect on the three non-collinear points $(0,0), (1,0)$ and $(0,1)$. We shall return to this method of determining affine transformations in Section 2.3.

So suppose that a given affine transformation t maps the points $(0,0), (1,0)$ and $(0,1)$ to three non-collinear points P, Q

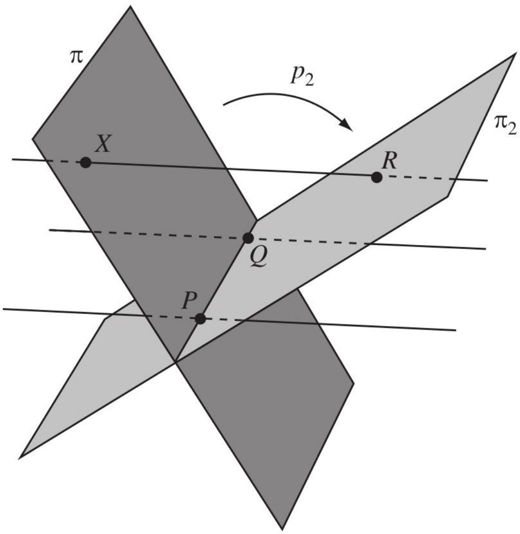
and R, respectively. In order to express t as the composition of two parallel projections p_1 and p_2, we need to define p_1 and p_2 in such a way that $p_2 \circ p_1$ has the same effect as t on $(0,0)$, $(1,0)$ and $(0,1)$. To do this, we first define p_1 so that it maps $(0,0)$ to $P, (1,0)$ to Q, and $(0,1)$ to some point X, say, and then define p_2 so that it maps X to R while leaving P and Q fixed.



To construct p_1 we embed its domain plane π_1, and its codomain plane π, into $\mathbb{R}^3$ so that the point $(0,0)$ in π_1 coincides with the point P in π, as shown below. It does not matter how this is done, provided that $(1,0)$ does not lie in π. We then define p_1 by the family of rays that are parallel to the ray through the point $(1,0)$ in π_1 and the point Q in π. When defined in this way, p_1 maps $(0,0)$ to $P, (1,0)$ to Q, and $(0,1)$ to some point X, as required.



To construct p_2 we embed its domain plane π, and its codomain plane π_2, into $\mathbb{R}^3$ so that the points P and Q in π coincide with the points P and Q in π_2, as shown below. Again it does not matter how this is done, provided that X does not lie in π_2. We then define p_2 by the family of rays that are parallel to the ray through the point X in π and the point R in π_2. Then p_2 leaves P and Q fixed and maps X to R.



Overall, the composite $p_2 \circ p_1$ of the two parallel projections maps $(0,0)$, $(1,0)$ and $(0,1)$ to P, Q and R, respectively. Now p_1 and p_2 are affine transformations, so $p_2 \circ p_1$ is also an affine transformation. Furthermore, $p_2 \circ p_1$ maps $(0,0), (1,0)$ and $(0,1)$ to the same points as does t. Since such affine transformations are unique, it follows that $t = p_2 \circ p_1$. We have therefore demonstrated the following result.

Theorem 6. An affine transformation can be expressed as the composite of two parallel projections.

An important consequence of this theorem is that all properties of figures that are unchanged by parallel projections must also be unchanged by affine transformations. In particular, the three properties of parallel projections that we met in Subsection 2.2.2 must, in fact, be affine properties.

2.3 Properties of Affine Transformations

In the previous section you saw how parallel projections can be used to explore affine geometry from a visual point of view. In this section we explore some of the same ideas from an algebraic point of view.

2.3.1 Images of Sets Under Affine Transformations

We begin by describing how to find the image of a line under an affine transformation. To do this, recall that an affine transformation is a mapping $t : \mathbb{R}^2 \to \mathbb{R}^2$ given by a formula of the form

$$t(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b} \tag{1}$$

where $\mathbf{A}$ is an invertible 2×2 matrix. The set of such transformations forms a group, in which the transformation inverse to t is given by

$$t^{-1}(\mathbf{x}) = \mathbf{A}^{-1}\mathbf{x} - \mathbf{A}^{-1}\mathbf{b} \tag{2}$$

When equations (1) and (2) are used to find images under t, it is easy to confuse points in the domain plane with points in the codomain plane, as both planes are copies of $\mathbb{R}^2$. To avoid such confusion, we often reserve the symbol $\mathbf{x}$ and the coordinates (x, y) for points in the domain of t, and use the symbol $\mathbf{x}'$ and the coordinates (x', y') to denote the image of $\mathbf{x}$ under t. With this notation, we may rewrite equations (1) and (2) in the form

$$\begin{aligned} \mathbf{x}' &= \mathbf{A}\mathbf{x} + \mathbf{b} && (3) \\ \mathbf{x} &= \mathbf{A}^{-1}\mathbf{x}' - \mathbf{A}^{-1}\mathbf{b} && (4) \end{aligned}$$

The next example illustrates how these equations can be used to find the image of a line under an affine transformation.

Example 1. Determine the image of the line $y = 2x$ under the affine transformation

$$t(\mathbf{x}) = \begin{pmatrix} 4 & 1 \\ 2 & 1 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 2 \\ -1 \end{pmatrix} \quad \left(\mathbf{x} \in \mathbb{R}^2\right) \tag{5}$$

Solution: Let (x, y) be an arbitrary point on the line $y = 2x$, and let (x', y') be the image of (x, y) under t. Then

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Next we use equation (4) to express (x, y) in terms of (x', y').

We have

$$\begin{pmatrix} 4 & 1 \\ 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -1 & 2 \end{pmatrix} \text{ and } \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ -4 \end{pmatrix}$$

so

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -1 & 2 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} -\frac{3}{2} \\ 4 \end{pmatrix}.$$

It follows that under the inverse mapping t^{-1} we have

$$x = \frac{1}{2}x' - \frac{1}{2}y' - \frac{3}{2} \quad \text{and} \quad y = -x' + 2y' + 4$$

Since x and y are related by the equation $y = 2x$, it follows that x' and y' are related by the equation

$$-x' + 2y' + 4 = 2\left(\frac{1}{2}x' - \frac{1}{2}y' - \frac{3}{2}\right)$$

which simplifies to

$$2x' - 3y' = 7$$

Dropping the dashes, we see that the image of the line $y = 2x$ under t is the line

$$2x - 3y = 7$$

$\square$

Problem 1. Determine the image of the line $3x - y + 1 = 0$

under the affine transformation

$$t(\mathbf{x}) = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -1 & 2 \end{pmatrix} \mathbf{x} + \begin{pmatrix} -\frac{3}{2} \\ 4 \end{pmatrix} \quad \left(\mathbf{x} \in \mathbb{R}^2\right).$$

Solution: Let (x, y) be an arbitrary point on the line $3x - y + 1 = 0$, and let (x', y') be the image of (x, y) under t. Then

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -\frac{3}{2} \\ 4 \end{pmatrix}.$$

Since the inverse of the inverse of any invertible transformation is the original transformation, it follows from Example 1 that under t^{-1}, we have

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Thus

$$x = 4x' + y' + 2 \quad \textit{and} \quad y = 2x' + y' - 1.$$

Hence the image under t of the line $3x - y + 1 = 0$ has equation

$$3\,(4x' + y' + 2) - (2x' + y' - 1) + 1 = 0.$$

Dropping the dashes and simplifying, we obtain

$$5x + y + 4 = 0$$

$\square$

Problem 2. Determine the image of the circle $x^2 + y^2 = 1$ under the affine transformation

$$t(\mathbf{x}) = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -1 & 2 \end{pmatrix} \mathbf{x} + \begin{pmatrix} -\frac{3}{2} \\ 4 \end{pmatrix} \quad \left(\mathbf{x} \in \mathbb{R}^2\right).$$

Solution: The argument here is similar to that of Problem 1. For, if (x, y) is an arbitrary point on the circle $x^2 + y^2 = 1$ and (x', y') is the image of (x, y) under t, then under t^{-1} we have

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Thus

$$x = 4x' + y' + 2 \quad \textit{and} \quad y = 2x' + y' - 1.$$

Hence the image under t of the circle $x^2 + y^2 = \mathrm{I}$ has equation

$$(4x' + y' + 2)^2 + (2x' + y' - 1)^2 = 1$$

Dropping the dashes and simplifying. we obtain

$$10x^2 + 6xy + y^2 + 6x + y + 2 = 0$$

$\square$

The same technique can be used to find the images of other types of figures, such as other conics. You will meet some examples of this in Section 2.5.

2.3.2 The Fundamental Theorem of Affine Geometry

The algebraic approach can also be used to investigate whether there is an affine transformation which maps one given figure onto another. Recall that if there is such a transformation, then the two figures are said to be affinecongruent. This concept of congruence is important because, as we explained in Section 2.1, figures that are affine-congruent to each other share the same affine properties.

In this subsection we prove the remarkable result that all triangles are affinecongruent and therefore share the same affine properties. In fact, since a triangle is completely determined by its three vertices, the congruence of triangles follows from the so-called Fundamental Theorem of Affine Geometry which states that any three non-collinear points can be mapped to any other three non-collinear points by an affine transformation.

First, recall that in Subsection 2.2.3 we described how the points $(0, 0)$, $(1, 0)$ and $(0, 1)$ in $\mathbb{R}^2$ can be mapped to any three non-collinear points P, Q and R by an affine transformation. This transformation is unique in the sense that it is completely determined by the choice of P, Q and R. The following example should remind you of how such transformations are constructed.

Example 2. Determine the affine transformation which maps the points $(0,0)$, $(1,0)$ and $(0,1)$ to the points $(3,2),(5,8)$ and $(7,3)$, respectively.

Solution: Let t be the affine transformation given by

$$t : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}. \tag{6}$$

Since $t(0,0) = (3,2)$, it follows from (6) that $e = 3$ and $f = 2$. Next, $t(1,0) = (5,8)$, so it follows from (6) that

$$\begin{pmatrix} 5 \\ 8 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix} + \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

The first column of the matrix for t is therefore

$$\begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} 5 \\ 8 \end{pmatrix} - \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \end{pmatrix}.$$

Finally, $t(0,1) = (7,3)$, so that

$$\begin{pmatrix} 7 \\ 3 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix} + \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

The second column of the matrix for t is therefore

$$\begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix} - \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}.$$

Hence the desired affine transformation is given by

$$t: \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2 & 4 \\ 6 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

$\square$ In general, if we want to find an affine transformation t of the form

$$t: \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \tag{7}$$

which maps $(0,0)$ to $\mathbf{p}$, $(1,0)$ to $\mathbf{q}$ and $(0,1)$ to $\mathbf{r}$, then we must choose $a, b, c,\ d, e$ and f so that

$$\begin{aligned} \mathbf{p} &= t(0,0) = (e,f), & &\text{so } (e,f) = \mathbf{p}; \\ \mathbf{q} &= t(1,0) = (a,c) + (e,f), & &\text{so } (a,c) = \mathbf{q} - \mathbf{p}; \\ \mathbf{r} &= t(0,1) = (b,d) + (e,f), & &\text{so } (b,d) = \mathbf{r} - \mathbf{p} \end{aligned}$$

Notice that any three points $\mathbf{p}, \mathbf{q}$ and $\mathbf{r}$ uniquely determine a transformation t of the form (7), but t is affine only if the matrix

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is invertible. Since the columns of $\mathbf{A}$ correspond to the vectors $\mathbf{q}-\mathbf{p}$ and $\mathbf{r}-\mathbf{p}$. it follows that $\mathbf{A}$ is invertible only if the vectors $\mathbf{q} - \mathbf{p}$ and $\mathbf{r} - \mathbf{p}$ are linearly independent. That is, provided that $\mathbf{p}, \mathbf{q}$ and $\mathbf{r}$ are not collinear.

So if $\mathbf{p}, \mathbf{q}$ and $\mathbf{r}$ are not collinear, then we can use the following strategy to find an affine transformation which maps $(0,0)$ to $\mathbf{p}, (1,0)$ to $\mathbf{q}$ and $(0,1)$ to $\mathbf{r}$.

Strategy. To determine the unique affine transformation $t(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}$ which maps $(0,0), (1,0)$ and $(0,1)$ to the three non-collinear points $\mathbf{p}, \mathbf{q}$ and $\mathbf{r}$, respectively:
1. take $\mathbf{b} = \mathbf{p}$:
2. take $\mathbf{A}$ to be the matrix with columns given by $\mathbf{q} - \mathbf{p}$ and $\mathbf{r} - \mathbf{p}$.

Problem 3. Use the strategy to determine the affine transformation which maps the points $(0,0), (1,0)$ and $(0,1)$ to the points $(2,3)$, $(1,6)$ and $(3,-1)$, respectively.

Solution: First, we take $\mathbf{b} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$. Next, we construct the matrix A whose first column is

$$\begin{pmatrix} 1 \\ 6 \end{pmatrix} - \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix},$$

and whose second column is

$$\begin{pmatrix} 3 \\ -1 \end{pmatrix} - \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ -4 \end{pmatrix}$$

thus

$$\mathbf{A} = \begin{pmatrix} -1 & 1 \\ 3 & -4 \end{pmatrix}$$

The required affine transformation t is therefore

$$t(\mathbf{x}) = \begin{pmatrix} -1 & 1 \\ 3 & -4 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \left(\mathbf{x} \in \mathbb{R}^2\right).$$

$\square$

Problem 4. Use the strategy to determine the affine transformation which maps the points $(0,0),(1,0)$ and $(0,1)$ to the points $(1,-2)$, $(2,1)$ and $(-3,5)$, respectively.

Solution: First, we take $\mathbf{b} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$. Next, we construct the matrix A whose first column is

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

and whose second column is

$$\begin{pmatrix} -3 \\ 5 \end{pmatrix} - \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -4 \\ 7 \end{pmatrix};$$

thus

$$\mathbf{A} = \begin{pmatrix} 1 & -4 \\ 3 & 7 \end{pmatrix}.$$

The required affine transformation t is therefore

$$t(\mathbf{x}) = \begin{pmatrix} 1 & -4 \\ 3 & 7 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ -2 \end{pmatrix} \quad \left(\mathbf{x} \in \mathbb{R}^2\right).$$

$\square$

Notice that the inverse of the transformation in Problem 3 is an affine transformation which maps the points $(2, 3), (1, 6)$ and $(3, -1)$ to the points $(0, 0)$, $(1, 0)$ and $(0, 1)$, respectively. So if, after applying this inverse, we apply the affine transformation in Problem 4 , then the overall effect is that of a composite affine transformation which sends the points $(2, 3), (1, 6)$ and $(3, -1)$ to the points $(1, -2), (2, 1)$ and $(-3, 5)$, respectively.

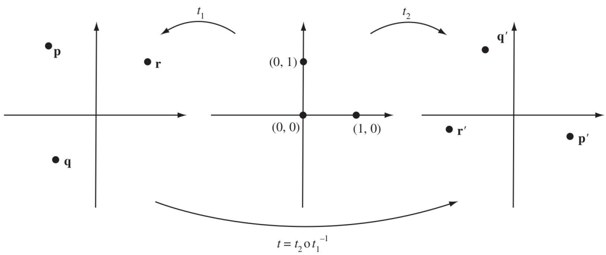
In a similar way, we can find an affine transformation which sends any three non-collinear points to any other three non-collinear points.

Theorem 1. (Fundamental Theorem of Affine Geometry) Let $\mathbf{p}, \mathbf{q}, \mathbf{r}$ and $\mathbf{p}', \mathbf{q}', \mathbf{r}'$ be two sets of three non-collinear points in $\mathbb{R}^2$. Then

(a) there is an affine transformation t which maps $\mathbf{p}, \mathbf{q}$ and $\mathbf{r}$ to $\mathbf{p}', \mathbf{q}'$ and $\mathbf{r}'$, respectively:
(b) the affine transformation t is unique.

Proof:

(a) Let t_1 be the affine transformation which maps $(0, 0), (1, 0)$ and $(0, 1)t$ the points $\mathbf{p}, \mathbf{q}$ and $\mathbf{r}$, respectively, and let t_2 be the affine transformation which maps $(0, 0), (1, 0)$ and $(0, 1)$ to the points $\mathbf{p}', \mathbf{q}'$ and $\mathbf{r}'$, respectively Then the composite $t = t_2 \circ t_1^{-1}$ is an affine transformation, and it maps $\mathbf{p}$ $\mathbf{q}$ and $\mathbf{r}$ to $\mathbf{p}', \mathbf{q}'$ and $\mathbf{r}'$, respectively.



(b) Suppose that t and s are both affine transformations which map p, $\mathbf{q}$ and $\mathbf{r}$ to p$'$, $\mathbf{q}'$ and $\mathbf{r}'$, respectively, and let t_1 be the affine transformation defined in part (a). Then the composites $t \circ t_1$ and $s \circ t_1$ are both affine transformations which map $(0,0),(1,0)$ and $(0,1)$ to $\mathbf{p}', \mathbf{q}'$ and $\mathbf{r}'$, respectively. Since an affine transformation is uniquely determined by its effect on the points $(0,0),(1,0)$ and $(0,1)$, it follows that $t \circ t_1 = s \circ t_1$

If we then compose both $t \circ t_1$ and $s \circ t_1$ on the right with t_1^{-1}, it follows that $t = s$. Thus the mapping t constructed in part (a) is unique. $\square$

Now suppose that we are given two arbitrary triangles $\triangle ABC$ and $\triangle DEF$. By the Fundamental Theorem there is an affine transformation which maps the vertices A, B, C to the vertices D, E, F, respectively. Since this transformation maps straight lines to straight lines, it must map the sides of $\triangle ABC$ to the sides of ΔDEF, so we have the following important corollary. This will be used extensively in Section 2.4.

Corollary. All triangles are affine-congruent.

In order to find the affine transformation which maps one triangle, vertex to vertex, onto another triangle, we follow the strategy used in part (a) of the proof of the Fundamental Theorem.

> **Strategy.** To determine the affine transformation t which maps three noncollinear points $\mathbf{p}, \mathbf{q}$ and $\mathbf{r}$ to another three non-collinear points $\mathbf{p}', \mathbf{q}'$ and $\mathbf{r}'$, respectively:
>
> 1. determine the affine transformation t_1 which maps $(0,0), (1,0)$ and $(0,1)$ to the points $\mathbf{p}, \mathbf{q}$ and $\mathbf{r}$, respectively;
>
> 2. determine the affine transformation t_2 which maps $(0,0), (1,0)$ and $(0,1)$ to the points $\mathbf{p}', \mathbf{q}'$ and $\mathbf{r}'$, respectively;
>
> 3. calculate the composite $t = t_2 \circ t_1^{-1}$.

Example 3. Determine the affine transformation which maps the points $(2,3)$, $(1,6)$ and $(3,-1)$ to the points $(1,-2), (2,1)$ and $(-3,5)$, respectively.

Solution: You have already seen in Problem 3 that the affine transformation t_1 which maps the points $(0,0), (1,0)$ and $(0,1)$ to the points $(2,3), (1,6)$ and $(3, -1)$, respectively, is given by

$$t_1(\mathbf{x}) = \begin{pmatrix} -1 & 1 \\ 3 & -4 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Also, in Problem 4 you saw that the affine transformation t_2 which maps the points $(0,0),(1,0)$ and $(0,1)$ to the points $(1,-2),(2,1)$ and $(-3,5)$, respectively, is given by

$$t_2(\mathbf{x}) = \begin{pmatrix} 1 & -4 \\ 3 & 7 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Following the strategy, we need to find the inverse of t_1. We have

$$\begin{pmatrix} -1 & 1 \\ 3 & -4 \end{pmatrix}^{-1} = \begin{pmatrix} -4 & -1 \\ -3 & -1 \end{pmatrix}$$

and

$$\begin{pmatrix} -4 & -1 \\ -3 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -11 \\ -9 \end{pmatrix},$$

so that the inverse of t_1 is given by

$$t_1^{-1}(\mathbf{x}) = \begin{pmatrix} -4 & -1 \\ -3 & -1 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 11 \\ 9 \end{pmatrix}.$$

Thus the affine transformation which maps the points $(2,3),(1,6)$ and $(3,-1)$ to the points $(1,-2),(2,1)$ and $(-3,5)$, respectively, is given by

$$\begin{aligned}
t(\mathbf{x}) &= t_2 \circ t_1^{-1}(\mathbf{x}) \\
&= t_2\left(\begin{pmatrix} -4 & -1 \\ -3 & -1 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 11 \\ 9 \end{pmatrix}\right) \\
&= \begin{pmatrix} 1 & -4 \\ 3 & 7 \end{pmatrix}\left(\begin{pmatrix} -4 & -1 \\ -3 & -1 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 11 \\ 9 \end{pmatrix}\right) + \begin{pmatrix} 1 \\ -2 \end{pmatrix} \\
&= \left(\begin{pmatrix} 8 & 3 \\ -33 & -10 \end{pmatrix}\mathrm{x} + \begin{pmatrix} -25 \\ 96 \end{pmatrix}\right) + \begin{pmatrix} 1 \\ -2 \end{pmatrix} \\
&= \begin{pmatrix} 8 & 3 \\ -33 & -10 \end{pmatrix}\mathbf{x} + \begin{pmatrix} -24 \\ 94 \end{pmatrix}.
\end{aligned}$$

$\square$

Problem 5. Determine the affine transformation which maps the points $(1,-1),(2,-2)$ and $(3,-4)$ to the points $(8,13),(3,4)$ and $(0,-1)$, respectively.

Solution: First, we find the affine transformation t_1 which maps $(0,0),(1,0)$ and $(0,1)$ to $(1,-1),(2,-2)$ and $(3,-4)$, respectively. This transformation has the form $t_1(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}$, where $\mathbf{b} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and

$$\mathbf{A} = \begin{pmatrix} 2-1 & 3-1 \\ -2+1 & -4+1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix};$$

that is,

$$t_1(\mathbf{x}) = \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Next, we find the affine transformation t_2 which maps $(0,0), (1,0)$ and $(0,1)$ to $(8,13)$, $(3,4)$ and $(0,-1)$, respectively. This transformation has the form $t_2(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}$, where $\mathbf{b} = \begin{pmatrix} 8 \\ 13 \end{pmatrix}$ and

$$\mathbf{A} = \begin{pmatrix} 3-8 & 0-8 \\ 4-13 & -1-13 \end{pmatrix} = \begin{pmatrix} -5 & -8 \\ -9 & -14 \end{pmatrix};$$

that is,

$$t_2(\mathbf{x}) = \begin{pmatrix} -5 & -8 \\ -9 & -14 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 8 \\ 13 \end{pmatrix}.$$

We now require the formula for the inverse transformation t_1^{-1}. Since

$$\begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix}^{-1} = -\begin{pmatrix} -3 & -2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix}$$

and

$$\begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

it follows that

$$t_1^{-1}(\mathbf{x}) = \begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix} \mathbf{x} - \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

The required affine transformation t is therefore $t(\mathbf{x}) = t_2 \circ t_1^{-1}(\mathbf{x})$, where $t_2 \circ t_1^{-1}(\mathbf{x})$ is equal to

$$\begin{aligned} t_2\left(\begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix}\mathbf{x} - \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) &= \begin{pmatrix} -5 & -8 \\ -9 & -14 \end{pmatrix}\left(\begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix}\mathbf{x} - \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) + \begin{pmatrix} 8 \\ 13 \end{pmatrix} \\ &= \begin{pmatrix} -7 & -2 \\ -13 & -4 \end{pmatrix}\mathbf{x} - \begin{pmatrix} -5 \\ -9 \end{pmatrix} + \begin{pmatrix} 8 \\ 13 \end{pmatrix} \\ &= \begin{pmatrix} -7 & -2 \\ -13 & -4 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 13 \\ 22 \end{pmatrix}. \end{aligned}$$

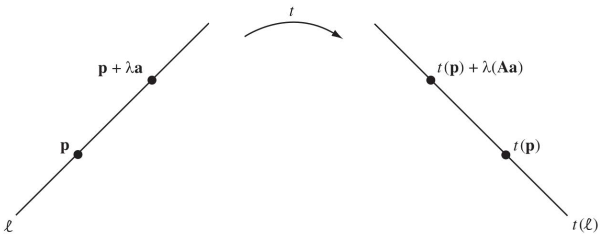
□

2.3.3 Proofs of the Basic Properties of Affine Transformations

In Subsection 2.2.2 we used parallel projections to demonstrate that affine transformations have the following basic properties: they map straight lines to straight lines, they map parallel lines to parallel lines, and they preserve ratios of lengths along a given straight line. We now give algebraic proofs of these assertions.

Theorem 2. An affine transformation maps straight lines to straight lines.

Proof:



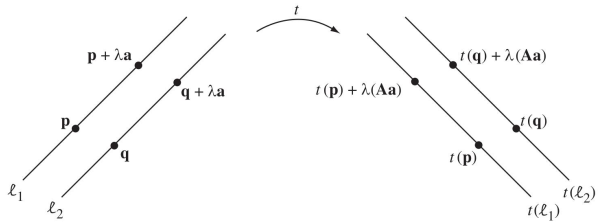
Let ℓ be a line through a point with position vector $\mathbf{p}$, and let the direction of ℓ be that of some vector a. Then $\ell = \{\mathbf{p} + \lambda\mathbf{a} : \lambda \in \mathbb{R}\}$ Now let $t : \mathbb{R}^2 \to \mathbb{R}^2$ be an affine transformation given by $t(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}$ We can find the image under t of an arbitrary point $\mathbf{p} + \lambda\mathbf{a}$ on ℓ as follows:

$$\begin{aligned} t(\mathbf{p} + \lambda\mathbf{a}) &= \mathbf{A}(\mathbf{p} + \lambda\mathbf{a}) + \mathbf{b} \\ &= (\mathbf{Ap} + \mathbf{b}) + \lambda\mathbf{Aa} \\ &= t(\mathbf{p}) + \lambda\mathbf{Aa}. \end{aligned}$$

So the image of ℓ is the set $t(\ell) = \{t(\mathrm{p}) + \lambda\mathrm{Aa} : \lambda \in \mathbb{R}\}$ which is a line through $t(\mathbf{p})$ in the direction of the vector $\mathbf{Aa}$. $\square$

Theorem 3. An affine transformation maps parallel straight lines to parallel straight lines.

Proof:



Let ℓ_1 and ℓ_2 be parallel lines through the points with position vectors $\mathbf{p}$ and q, respectively, and let the direction of the lines be that of the vector a. Then $\ell_1 = \{\mathbf{p} + \lambda\mathbf{a} : \lambda \in \mathbb{R}\}$ and $\ell_2 = \{\mathbf{q} + \lambda\mathbf{a} : \lambda \in \mathbb{R}\}$

As in the proof of Theorem 2, the images of ℓ_1 and ℓ_2 under the affine transformation $t(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}$ are the sets

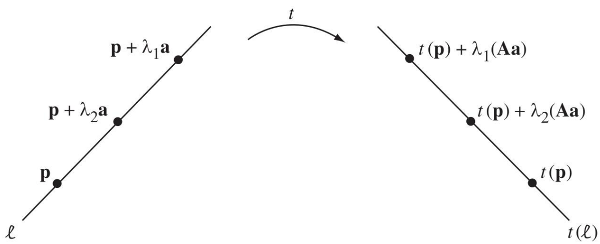
$$t(\ell_1) = \{t(\mathbf{p}) + \lambda\mathbf{Aa} : \lambda \in \mathbb{R}\} \quad \text{and} \quad t(\ell_2) = \{t(\mathbf{q}) + \lambda\mathbf{Aa} : \lambda \in \mathbb{R}\}$$

These sets are straight lines which pass through the image points $t(\mathbf{p})$ and $t(\mathbf{q})$, both in the same direction as that of the vector Aa. Hence the two image lines under t are parallel, as claimed. $\square$

Rather than prove that affine transformations preserve ratios of lengths along a given straight line, as in Property 3 of Subsection 2.2.2, we prove the following more general result illustrated in the margin. The original result follows because any line is parallel to itself.

Theorem 4. An affine transformation preserves ratios of lengths along parallel straight lines.

Proof: We begin by examining what happens to the length of a line segment under an affine transformation.



Let ℓ be a line through a point with position vector $\mathbf{p}$, and let the direction of ℓ be that of some unit vector a. Then

$$\ell = \{\mathbf{p} + \lambda\mathbf{a} : \lambda \in \mathbb{R}\}$$

As in the proof of Theorem 2, the image of ℓ under the affine transformation $t(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}$ is the line

$$t(\ell) = \{t(\mathbf{p}) + \lambda\mathbf{Aa} : \lambda \in \mathbb{R}\}$$

Now consider a segment of ℓ with endpoints $\mathbf{p}+\lambda_1\mathbf{a}$ and $\mathbf{p}+\lambda_2\mathbf{a}$. Since a is a unit vector, the length of the segment is

$$\|(\mathbf{p} + \lambda_2\mathbf{a}) - (\mathbf{p} + \lambda_1\mathbf{a})\| = |\lambda_2 - \lambda_1| \cdot \|\mathbf{a}\| = |\lambda_2 - \lambda_1|$$

The image of the segment has endpoints $t(\mathbf{p}) + \lambda_1$ Aa and

$t(\mathbf{p}) + \lambda_2 \mathbf{Aa}$, so the image of the segment has length

$$\|(t(\mathbf{p}) + \lambda_2 \mathbf{Aa}) - (t(\mathbf{p}) + \lambda_1 \mathbf{Aa})\| = |\lambda_2 - \lambda_1| \cdot \|\mathbf{Aa}\|.$$

So, in the process of mapping segments along ℓ to segments along $t(\ell)$, lengths are stretched by the factor $||Aa||$. Since this factor is the same for all segments which lie along lines parallel to $\mathbf{a}$, it follows that the ratios of lengths along parallel lines are unchanged by t. □

2.4 Use of Fundamental Theorem of Affine Geometry

In this section we explain how the Fundamental Theorem of Affine Geometry can be used to deduce the fact that the medians of any triangle are concurrent from the special case that the medians of an equilateral triangle are concurrent. We then use similar methods to prove the classical theorems of Ceva and Menelaus.

2.4.1 The Median Theorem

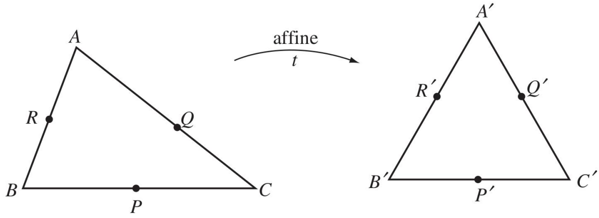
Let $\triangle ABC$ be an arbitrary triangle in the plane. If you join the midpoint of each side of the triangle to the opposite vertex (these lines are called the medians of the triangle), these three lines appear to pass through a single point. In fact, no matter what triangle you choose, you find that its medians meet in a single point.

Theorem 1. (Median Theorem) The medians of any triangle are concurrent.

We can get some evidence that this theorem holds in general by looking first at a special case where a proof of the theorem is straight-forward - namely, when the triangle is an equilateral triangle.

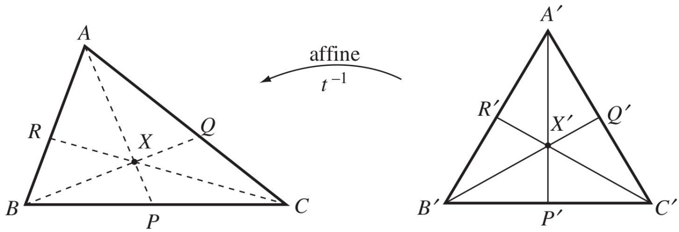
To do this, consider an equilateral triangle $\triangle ABC$, with medians AP, BQ and CR. Since $\triangle ABC$ has sides of equal length, it must be symmetric about the line AP. Thus the point at which BQ meets CR must be symmetrically placed with respect to this line - that is, it must actually lie on the line AP. In other words, the lines AP, BQ and CR are concurrent if the triangle is equilateral.

In order to show that the medians of an arbitrary triangle meet at a point, consider an arbitrary triangle $\triangle ABC$, and let P, Q and R be the midpoints of the sides BC, CA and AB, respectively. Next, choose a particular equilateral triangle $\Delta A'B'C'$, and let P', Q' and R' be the midpoints of the sides $B'C'$, $\boldsymbol{C'}A'$ and $A'B'$, respectively.



According to the Fundamental Theorem of Affine Geometry there is an affine transformation t which maps $\triangle ABC$ onto $\Delta A'B'C'$. Moreover, since affine transformations preserve ratios of lengths along lines it follows that t maps the mid-points P, Q and R to the mid-points P', Q' and R', respectively.

From the above discussion we know that the medians of any equilateral triangle meet at a point, so in particular we know that $A'P', B'Q'$ and $C'R'$ meet at some point X', say, as shown on the right below.



The trick now is to observe that t has an inverse t^{-1} which is also an affine transformation. This inverse maps the medians $A'P', B'Q'$ and $C'R'$ back to the medians AP, BQ and CR of the original triangle $\triangle ABC$. Since X' lies on all three of the lines $A'P', B'Q'$ and $C'R'$ it follows that t^{-1} maps X' to some point X which lies on all three of the lines AP, BQ and CR. In other words, the medians of $\triangle ABC$ are concurrent.

Since $\triangle ABC$ is an arbitrary triangle we have proved the Median Theorem. The essence of the above proof is the fact that all triangles are affinecongruent. That powerful result enables us to prove theorems concerning the affine properties of triangles (such as collinearity, lines being parallel, and ratios of lengths along a given line) following a standard pattern. First, we choose a particular type of triangle for which it is easy to prove the result. Then, by asserting the existence of an affine transformation from that triangle to an arbitrary triangle, we deduce that the result holds for all triangles.

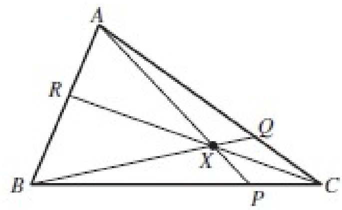
This is the approach we shall use to prove the theorems of Ceva and Menelaus later in the section.

2.4.2 Ceva's Theorem

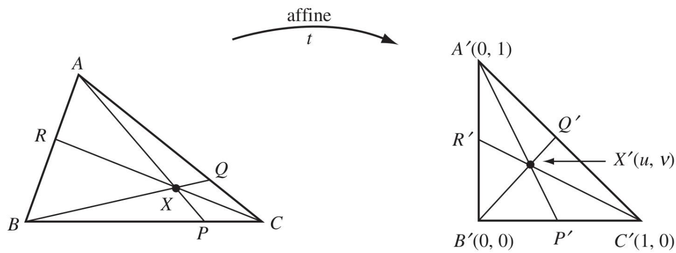
We now prove the following theorem due to Ceva.

Theorem 2. (Ceva's Theorem) Let $\triangle ABC$ be a triangle, and let X be a point which does not lie on any of its (extended) sides. If AX meets BC at P, BX meets CA at Q and CX meets BA at R, then

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = 1$$



Proof: According to the Fundamental Theorem of Affine Geometry there is an affine transformation t which maps the points A, B, C to the points $A' = (0,1)$, $B' = (0,0), C' = (1,0)$, respectively. This transformation maps the triangle $\triangle ABC$ onto the right-angled triangle $\Delta A'B'C'$, and it maps the point X to some point $X' = (u, v)$.



Using coordinate geometry we can calculate the equations of the lines $A'X'$, $B'X', C'X'$ and hence find the coordinates of the point P' where $A'X'$ meets $B'C'$, of the point Q' where $B'X'$ meets $A'C'$, and of the point R' where $C'X'$ meets $A'B'$. Starting with the point P', we note that the line $B'C'$ has equation $y = 0$. Also, the line $A'X'$ has slope $\frac{1-v}{0-u}$, so its equation is $y - 1 = \frac{1-v}{0-u}(x - 0)$. Hence, at the point P' where the two lines meet, we must have $y = 0$ and $y - 1 = \frac{1-v}{0-u}(x - 0)$, so

$$P' = \left(\frac{u}{1-v}, 0 \right).$$

Similarly, at the point R' we have $x = 0$, and $y - 0 = \frac{0-v}{1-u}(x - 1)$, so

$$R' = \left(0, \frac{v}{1-u} \right).$$

Finally, at Q' we have $x + y = 1$ and $y = \frac{v}{u}x$, so $x = \frac{u}{u+v}$ and $y = \frac{v}{u+v}$. Hence

$$Q' = \left(\frac{u}{u+v}, \frac{v}{u+v} \right)$$

Thus, using the coordinate formulas for calculating ratios we obtain

$$\frac{A'R'}{R'B'} = \frac{y_{R'} - y_{A'}}{y_{B'} - y_{R'}} = \frac{\frac{v}{1-u} - 1}{0 - \frac{v}{1-u}} = \frac{u + v - 1}{-v}$$

$$\frac{B'P'}{P'C'} = \frac{x_{P'} - x_{B'}}{x_{C'} - x_{P'}} = \frac{\frac{u}{1-v} - 0}{1 - \frac{u}{1-v}} = \frac{u}{1 - u - v}$$

and

$$\frac{C'Q'}{Q'A'} = \frac{y_{Q'} - y_{C'}}{y_{A'} - y_{Q'}} = \frac{\frac{v}{u+v} - 0}{1 - \frac{v}{U+v}} = \frac{v}{u}$$

Hence

$$\frac{A'R'}{R'B'} \cdot \frac{B'P'}{P'C'} \cdot \frac{C'Q'}{Q'A'} = 1$$

Since t^{-1} is an affine transformation, it preserves ratios along a line. It must therefore map P', Q', R' back to the points P, Q, R

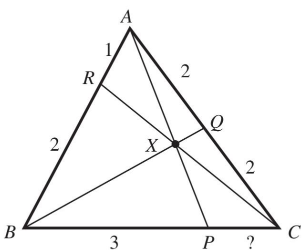
in such a way that

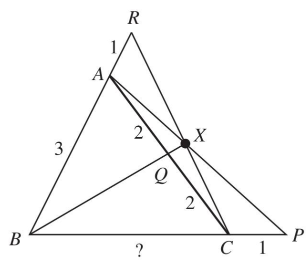
$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = 1$$

as required. $\square$

The next example illustrates how we can use Ceva's Theorem to calculate certain unknown distances along the sides of a triangle. For the method to work correctly, it is important to remember that all the ratios in Ceva's Theorem are signed ratios. Thus, if X lies inside the triangle, as in part (a) of the example, then all the ratios are positive. But if X lies outside the triangle, as in part (b), then two of the ratios will be negative.

Example 1. (a) In the figure on the left below, $AR = 1, RB = 2, BP = 3, CQ = 2$ and $QA = 2$. Calculate the distance PC.
(b) For the figure on the right, $AR = 1, AB = 3, PC = 1, CQ = 2$ and $QA = 2$. Calculate the distance BC.





Solution: (a) By Ceva's Theorem, we have

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = 1$$

so,

$$\frac{1}{2} \cdot \frac{3}{PC} \cdot \frac{2}{2} = 1$$

It follows that $PC = \frac{3}{2}$ (b) By Ceva's Theorem, we have

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = 1$$

so,

$$-\frac{1}{4} \cdot \left(-\frac{BC+1}{1}\right) \cdot \frac{2}{2} = 1$$

It follows that $BC = 3$. □

Problem 1. (a) Determine the ratio $\frac{BP}{PC}$ in the left diagram below, given that

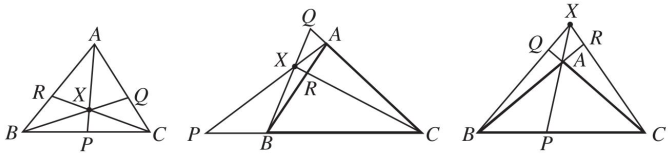
$$\frac{AR}{RB} = \frac{AQ}{QC} = \frac{3}{2}$$

(b) Determine the ratio $\frac{CQ}{QA}$ in the middle diagram below, given that

$$\frac{AR}{RB} = \frac{1}{2} \quad and \quad \frac{BP}{PC} = -\frac{2}{7}$$

(c) Determine the ratio $\frac{AR}{RB}$ in the right diagram below, given that

$$\frac{BP}{PC} = \frac{5}{7} \quad and \quad \frac{CQ}{QA} = -7$$



Solution:

(a) By Ceva's Theorem, we have

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = 1.$$

First, $\frac{AR}{RB} = \frac{3}{2}$. Next, $\frac{AQ}{QC} = \frac{3}{2}$ and so $\frac{CQ}{QA} = \frac{2}{3}$. It follows that

$$\frac{3}{2} \cdot \frac{BP}{PC} \cdot \frac{2}{3} = 1$$

so $\frac{BP}{PC} = 1$.

(b) By Ceva's Theorem, we have

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = 1$$

Since $\frac{AR}{RB} = \frac{1}{2}$ and $\frac{BP}{PC} = -\frac{2}{7}$, it follows that

$$\frac{1}{2} \cdot \left(-\frac{2}{7}\right) \cdot \frac{CQ}{QA} = 1$$

so $\frac{CQ}{QA} = -7$

(c) By Ceva's Theorem, we have

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = 1$$

We are given that $\frac{BP}{PC} = \frac{5}{7}$ and $\frac{CQ}{QA} = -7$, so

$$\frac{AR}{RB} \cdot \frac{5}{7} \cdot (-7) = 1$$

Hence $\frac{AR}{RB} = -\frac{1}{5}$

$\square$

Ceva's Theorem has the following converse, which can be regarded as a generalization of the Median Theorem to configurations where P, Q, R are not all midpoints of sides.

Theorem 3. (Converse to Ceva's Theorem) Let P, Q and R be points, other than vertices, on the (possibly extended) sides BC, CA and AB of a triangle $\triangle ABC$, such that

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = 1 \tag{1}$$

Then the lines AP, BQ and CR are concurrent.

Proof: Let the lines BQ and CR intersect at a point X, and let the line AX meet BC at some point P'. It is sufficient to prove that $P = P'$. It follows from Ceva's Theorem that

$$\frac{AR}{RB} \cdot \frac{BP'}{P'C} \cdot \frac{CQ}{QA} = 1 \tag{2}$$

Hence, from equations (1) and (2), we have

$$\frac{BP}{PC} = \frac{BP}{P'C}$$

so that P and P' must indeed be the same point. $\square$

Example 2. The triangle $\triangle ABC$ has vertices $A(1,3), B(-1,0)$ and $C(4,0)$, and the points $P(0,0), Q\left(\frac{8}{3}, \frac{4}{3}\right)$ and $R\left(-\frac{2}{3}, \frac{1}{2}\right)$ lie on BC, CA and AB, respectively.
(a) Determine the ratios in which P, Q and R divide the sides of the triangle.
(b) Determine whether the lines AP, BQ and CR are concurrent.

Solution: (a) Using the coordinate formulas for calculating ratios, we obtain

$$\frac{AR}{RB} = \frac{x_{R} - x_{A}}{x_{B} - x_{R}} = \frac{-\frac{2}{3} - 1}{-1 + \frac{2}{3}} = 5, \quad \frac{BP}{PC} = \frac{x_{P} - x_{B}}{x_{C} - x_{P}} = \frac{0+1}{4-0} = \frac{1}{4}$$

$$\frac{CQ}{QA} = \frac{x_{Q} - x_{C}}{x_{A} - x_{Q}} = \frac{\frac{8}{3} - 4}{1 - \frac{8}{3}} = \frac{4}{5}$$

so that P divides BC in the ratio $1 : 4, Q$ divides CA in the ratio $4 : 5$ and R divides AB in the ratio $5 : 1$.
(b) It follows from (3) that the product

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = 5 \cdot \frac{1}{4} \cdot \frac{4}{5} = 1$$

so by the converse to Ceva's Theorem the lines AP, BQ and CR must be concurrent. $\square$

Problem 2. The triangle $\triangle ABC$ has vertices $A(-1,1), B(2,-1)$ and $C(3,2)$, and the points $P\left(\frac{8}{3},1\right), Q\left(2,\frac{7}{4}\right)$ and $R\left(\frac{4}{3},-\frac{1}{5}\right)$ lie on BC, CA and AB, respectively.

(a) Determine the ratios in which P, Q and R divide the sides of the triangle.

(b) Determine whether the lines AP, BQ and CR are concurrent.

Solution:

(a) Here we use the formula for calculating ratios given at the beginning of Appendix 2 , just above the Section Formula. This gives

$$\frac{BP}{PC} = \frac{x_P - x_B}{x_C - x_P} = \frac{\frac{8}{3} - 2}{3 - \frac{8}{3}} = 2$$

$$\frac{CQ}{QA} = \frac{x_Q - x_C}{x_A - x_Q} = \frac{2-3}{-1-2} = \frac{1}{3}$$

$$\frac{AR}{RB} = \frac{x_R - x_A}{x_B - x_R} = \frac{\frac{4}{5} + 1}{2 - \frac{4}{5}} = \frac{3}{2}$$

Thus

P divides BC in the ratio $2:1$,

Q divides CA in the ratio $1:3$,
R divides AB in the ratio $3:2$.

(b) It follows from part (a) that

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = \frac{3}{2} \cdot 2 \cdot \frac{1}{3} = 1$$

Thus by the converse to Ceva's Theorem, the lines AP, BQ and CR are concurrent.

$\square$

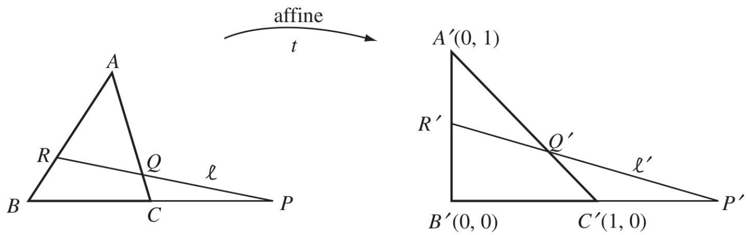
2.4.3 Menelaus' Theorem

Ceva's theorem is concerned with lines through the vertices of a triangle that meet at a point. We now use the Fundamental Theorem of Affine Geometry to prove an analogous theorem due to Menelaus which is concerned with points on the sides of a triangle that are collinear.

Theorem 4. (Menelaus' Theorem) Let $\triangle ABC$ be a triangle, and let ℓ be a line that crosses the sides BC, CA, AB at three distinct points P, Q, R, respectively. Then

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = -1$$

Proof: According to the Fundamental Theorem of Affine Geometry there is an affine transformation t which maps the points A, B, C to the points $A'(0,1)$, $B'(0,0), C'(1,0)$, respectively. This transformation maps the triangle $\triangle ABC$ onto the right-angled triangle $\Delta A'B'C'$, and it maps the line ℓ to some line ℓ'. Let the equation of ℓ' be $y = mx + c$.



We now calculate the coordinates of the points P', Q' and R' where ℓ' meets the sides $B'C', C'A'$ and $A'B'$, respectively.

At P' we have $y = 0$ and $y = mx + c$. This implies that $x = -\frac{c}{m}$, and hence $P' = \left(-\frac{c}{m}, 0\right)$. At R' we have $x = 0$ and $y = mx + c$. This implies that $y = c$, and hence $R' = (0, c)$. At Q' we have $x + y = 1$ and $y = mx + c$. This implies that $1 - x = mx + c$ so that $x = \frac{1-c}{m+1}$; also $y = m(1-y) + c$, so that $y = \frac{m+c}{m+1}$; and hence $Q' = \left(\frac{1-c}{m+1}, \frac{m+c}{m+1}\right)$.

Using the coordinate formulas for calculating ratios we obtain

$$\frac{A'R'}{R'B'} = \frac{y_{R'} - y_{A'}}{y_{B'} - y_{R'}} = \frac{c-1}{0-c} = \frac{c-1}{-c},$$

$$\frac{B'P'}{P'C'} = \frac{x_{P'} - x_{B'}}{x_{C'} - x_{P'}} = \frac{-\frac{c}{m} - 0}{1 + \frac{c}{m}} = \frac{-c}{m+c},$$

and
$$\frac{C'Q'}{Q'A'} = \frac{x_{Q'} - x_{C'}}{x_{A'} - x_{Q'}} = \frac{\frac{1-c}{m+1} - 1}{0 - \frac{1-c}{m+1}} = \frac{-(m+c)}{c-1}$$

Hence,
$$\frac{A'R'}{R'B'} \cdot \frac{B'P'}{P'C'} \cdot \frac{C'Q'}{Q'A'} = -1$$

Since t^{-1} is an affine transformation, it preserves ratios along a line. It must therefore map P', Q', R' back to the points P, Q, R in such a way that

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = -1$$

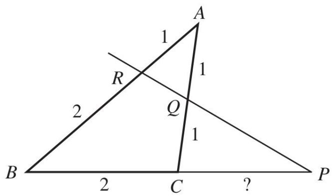
as required. □

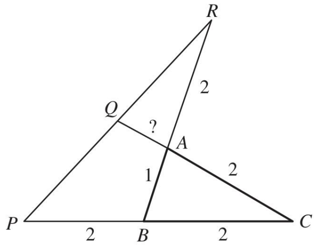
Remark

As for Ceva's Theorem, it is important to remember that all the ratios in Menelaus' Theorem are signed ratios. In fact if ℓ passes through the interior of the triangle, then precisely one of the ratios is negative; otherwise all the ratios are negative.

Example 3. (a) In the figure on the left below: $AR = 1, RB = 2, BC = 2, CQ = 1$ and $QA = 1$. Calculate the distance PC.

(b) In the figure on the right below: $AR = 2, AB = 1, BC = 2, CA = 2$ and $BP = 2$. Calculate the distance QA.





Solution:

(a) By Menelaus' Theorem, we have So $\frac{1}{2} \cdot \left(-\frac{2+PC}{PC}\right) \cdot \frac{1}{1} = -1$ It follows that $2 + PC = 2PC$, and hence $PC = 2$.

(b) By Menelaus' Theorem, we have $\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = -1$ So $\left(-\frac{2}{3}\right) \cdot \left(-\frac{2}{4}\right) \cdot \left(-\frac{2+QA}{QA}\right) = -1$ It follows that $2 + QA = 3QA$, and hence $QA = 1$.

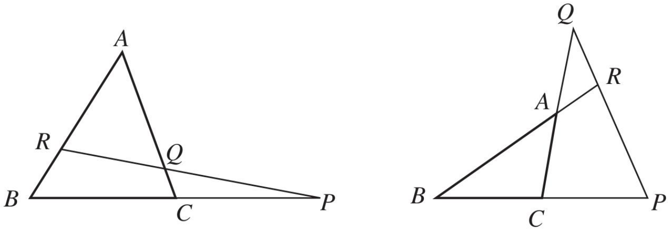
$\square$

Problem 3. (a) Determine the ratio $\frac{CQ}{QA}$ in the left diagram below, given that

$$\frac{AR}{RB} = 2 \quad and \quad \frac{BP}{PC} = -2.$$

b) Determine the ratio $\frac{CQ}{QA}$ in the right diagram below, given that

$$\frac{AR}{RB} = -\frac{1}{4} \quad and \quad \frac{BP}{PC} = -2.$$



Solution:

(a) By Menelaus' Theorem, we have

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = -1$$

We are given that $\frac{AR}{RB} = 2$ and $\frac{BP}{PC} = -2$, so

$$2 \cdot (-2) \cdot \frac{CQ}{QA} = -1.$$

Hence $\frac{CQ}{QA} = \frac{1}{4}$

(b) By Menelaus' Theorem, we have

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = -1$$

We are given that $\frac{AR}{RB} = -\frac{1}{4}$ and $\frac{BP}{PC} = -2$, so

$$\left(-\frac{1}{4}\right) \cdot (-2) \cdot \frac{CQ}{QA} = -1$$

Hence $\frac{CQ}{QA} = -2$.

$\square$

Menelaus' Theorem has a converse that enables us to check whether points on the three sides of a triangle are collinear.

Theorem 5. (Converse to Menelaus' Theorem) Let P, Q and R be points other than vertices on the (possibly extended) sides BC, CA and AB of a triangle $\triangle ABC$, such that

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = -1 \tag{3}$$

Then the points P, Q and R are collinear.

Proof: Let the line ℓ that passes through Q and R meet BC at some point P'. It is sufficient to prove that $P = P'$. It follows from Menelaus" Theorem that

$$\frac{AR}{RB} \cdot \frac{BP}{P'C} \cdot \frac{CQ}{QA} = -1$$

Hence, from equations (3) and (4) we deduce that

$$\frac{BP}{PC} = \frac{BP'}{P'C}$$

It follows that P and P' must indeed be the same point. $\square$

Problem 4. The triangle $\triangle ABC$ has vertices $A(2,4), B(-2,0)$ and $C(1,0)$, and the points $P\left(\frac{5}{2}, 0\right), Q\left(\frac{3}{2}, 2\right)$ and $R(1,3)$ lie on BC, CA and AB, respectively.

(a) Determine the ratios in which P, Q and R divide the sides of the triangle.

(b) Hence determine whether the points P, Q and R are collinear.

Solution:

(a) Here we have

$$\begin{aligned}
\frac{BP}{PC} &= \frac{x_P - x_B}{x_C - x_P} = \frac{\frac{5}{2}+2}{1-\frac{5}{2}} = -3 \\
\frac{CQ}{QA} &= \frac{x_Q - x_C}{x_A - x_Q} = \frac{\frac{3}{2}-1}{2-\frac{3}{2}} = 1 \\
\frac{AR}{RB} &= \frac{x_R - x_A}{x_B - x_R} = \frac{1-2}{-2-1} = \frac{1}{3}
\end{aligned}$$

Thus

$$\begin{aligned}
&P \text{ divides } BC \text{ in the ratio } -3:1, \\
&Q \text{ divides } CA \text{ in the ratio } 1:1, \\
&R \text{ divides } AB \text{ in the ratio } 1:3.
\end{aligned}$$

(b) It follows from part (a) that

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = \frac{1}{3} \cdot (-3) \cdot 1 = -1$$

Thus by the converse to Menelaus' Theorem, the points P, Q and R are collinear.

$\square$

We end this subsection with two revision problems.

Problem 5. Let $\triangle ABC$ be a triangle, and let X be a point which does not lie on any of its (extended) sides. Also, let AX meet BC at P, BX meet CA at Q and CX meet BA at R; and let QR and BC meet at T. Given that $\frac{BP}{PC} = k$, determine $\frac{BT}{TC}$ in terms of k.

Solution: By Ceva's Theorem, we have

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = 1. \qquad (*)$$

Also, by Menelaus' Theorem, we have

$$\frac{AR}{RB} \cdot \frac{BT}{TC} \cdot \frac{CQ}{QA} = -1. \qquad (**)$$

Comparing $(*)$ and $(**)$, we deduce that

$$\frac{BT}{TC} = -\frac{BP}{PC} = -k$$

□

Problem 6. Suppose that P and Q are the midpoints of the sides AB and BC of a parallelogram $ABCD$, and that the lines DP and AQ meet at R.

(a) Determine the image of B under the affine transformation t which maps A, D and C to $(0, 1), (0, 0)$ and $(1, 0)$, respectively.

(b) By considering the image of $ABCD$ under t, determine the ratios $PR : RD$ and $AR : RQ$

Solution: By the Fundamental Theorem of Affine Geometry (Theorem 1 , Subsection 2.3.2), there exists a unique affine transformation t which maps A, D and C to $A'(0,1), D'(0,0)$ and $C'(1,0)$, respectively.

(a) Since t maps AD onto the vertical line $A'D'$, and BC is parallel to AD, it follows that the image of BC under t must be a vertical line. Also, since t maps DC onto the horizontal line $D'C'$, and AB is parallel to DC, the image of AB under t must be a horizontal line. It follows that B', the image of B under t, must be the point with coordinates $(1,1)$.

(b) Since P is the midpoint of AB, its image $P' = t(P)$ must be the midpoint of $A'B'$ since ratios along a line are preserved by t. Hence $P' = \left(\frac{1}{2}, 1\right)$.

Since the slope of $D'P'$ is 2 and the line passes through the origin, the equation of the line $D'P'$ must be

$$y = 2x \qquad (*)$$

Next, since Q is the midpoint of BC, its image $Q' = t(Q)$ must be the midpoint of $B'C'$. Hence $Q' = \left(1, \frac{1}{2}\right)$. Then the slope of the line $A'Q'$ must be

$$\frac{1-\frac{1}{2}}{0-1} = -\frac{1}{2}$$

Since the line passes through the point $A'(0, 1)$, the equation of the line $A'Q'$ must be

$$y - 1 = -\frac{1}{2}(x - 0)$$

that is,

$$y = -\frac{1}{2}x + 1 \qquad (**)$$

Now, R', the image of R under t, must lie on the lines $D'P'$ and $A'Q'$, so that its coordinates must satisfy both $(*)$ and $(*)$. Substituting for y from $(*)$ into $(**)$, we obtain

$$2x = -\frac{1}{2}x + 1$$

so that $x = \frac{2}{5}$. It follows from $(*)$ that $y = \frac{4}{5}$. Thus R' is the point $\left(\frac{2}{5}, \frac{4}{5}\right)$.

Comparing the y -coordinates $1, \frac{4}{5}$ and 0 of P', R' and D', we obtain $P'R' : R'D' = 1 : 4$.

Since ratios along a line are preserved by the affine transformation t^{-1}, it follows that

$$PR : RD = 1 : 4$$

Finally, comparing the x -coordinates $0, \frac{2}{5}$ and 1 of A', R' and Q', we obtain $A'R' : R'Q' = 2 : 3$. Since ratios along a line are preserved under the affine transformation t^{-1}, it follows that

$$AR : RQ = 2 : 3.$$

$\square$

2.5 Affine Transformations and Conics

2.5.1 Classifying Non-Degenerate Conics in Affine Geometry

In Section 2.2 you saw that under an affine transformation a straight line maps to a straight line. Indeed, it follows from the Fundamental Theorem of Affine Geometry that any straight line can be mapped to any other straight line by some affine transformation. We now explore the corresponding situation for conics. Recall that a conic is a set in R^2 given by an equation of the form

$$Ax^2 + Bxy + Cy^2 + Fx + Gy + H = 0 \tag{1}$$

where A, B, C, F, G and H are real numbers, and A, B and C are not all zero. The three types of non-degenerate conic are ellipses, parabolas and hyperbolas. A non-degenerate conic is an ellipse if $B^2 - 4AC < 0$, a parabola if $B^2 - 4AC = 0$, and a hyperbola if $B^2 - 4AC > 0$.

First, consider the case where equation (1) represents an ellipse, as illustrated on the left of the figure below. We can apply a translation to move the centre of the ellipse to the origin, and then a rotation to align its major and minor axes with the directions of the x -axis and y -axis, respectively. After we have applied these two Euclidean transformations, the equation of the ellipse becomes

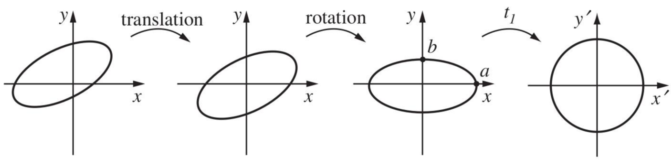
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad a \geq b > 0 \tag{2}$$

If we now apply the affine transformation $t_1 : (x, y) \mapsto (x', y')$, where

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1/a & 0 \\ 0 & 1/b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

then $x' = x/a$ and $y' = y/b$, so equation (2) becomes

$$(x')^2 + (y')^2 = 1.$$



Since the translation, the rotation and the transformation t_1 are all affine, their composite must also be affine. Overall, this shows that each ellipse can be mapped onto the unit circle by an affine transformation. We therefore have the following theorem.

Theorem 1. Every ellipse is affine-congruent to the unit circle with equation $x^2 + y^2 = 1$

Secondly, consider the case where equation (1) represents a hyperbola, as illustrated on the left of the figure below. Again, we can apply a translation to move the centre of the hyperbola to the origin, and then a rotation to align its major and minor axes with the directions of the x -axis and y -axis, respectively. After we have applied these two transformations, the equation of the hyperbola becomes

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{3}$$

Under the affine transformation t_1 defined above, equation (3) becomes

$$(x')^2 - (y')^2 = 1$$

that is,

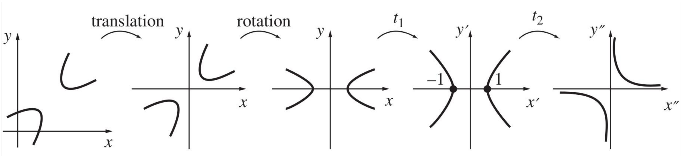
$$(x' - y')(x' + y') = 1 \tag{4}$$

Finally, if we apply the affine transformation $t_2 : (x', y') \mapsto (x'', y'')$, where

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

then equation (4) becomes

$$x''y'' = 1$$



Dropping the dashes from the equation $x''y'' = 1$, we obtain the following theorem.

Theorem 2. Every hyperbola is affine-congruent to the rectangular hyperbola with equation $xy = 1$.

Finally, consider the case where equation (1) represents a parabola, as illustrated on the left of the figure below. We can apply a translation to move the vertex of the parabola to the origin, and then a rotation to align its axis with the (positive) x -axis. After we have applied these two Euclidean transformations, the equation of the parabola becomes

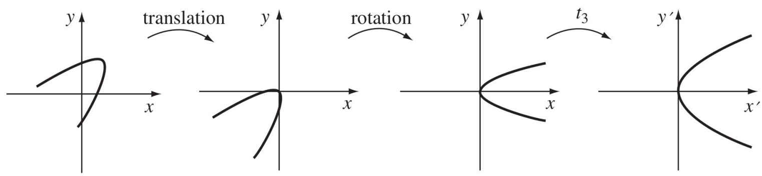
$$y^2 = ax \tag{5}$$

where a is some positive number which depends on the coefficients in equation (1). Next, if we apply the affine transformation $t_3 : (x, y) \mapsto (x', y')$, where

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1/a & 0 \\ 0 & 1/a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

then $x' = x/a$ and $y' = y/a$, so equation (5) becomes $(y'a)^2 = a\,(x'a)$, or

$$(y')^2 = x'$$



Dropping the dashes, we obtain the following theorem.

Theorem 3. Every parabola is affine-congruent to the parabola with equation $y^2 = x$.

Since all parabolas are affine-congruent to $y^2 = x$, they must be affine-congruent to each other. Similarly, by Theorem 1, all ellipses must be affine-congruent to each other; and, by Theorem 2, all hyperbolas must be affine-congruent to each other.

This raises the question as to whether it is possible for one type of conic (such as an ellipse) to be affine-congruent to another type of conic (such as a hyperbola). The next theorem shows that this cannot happen. In fact, since an affine transformation can be expressed as the composite of two parallel projections, this should not surprise you. After all, no parallel projection can change a bounded curve (such as an ellipse) into an unbounded one (such as a parabola or a hyperbola); nor can it change a curve with two branches (a hyperbola) into a curve with just one branch (an ellipse or a parabola).

Theorem 4. Affine transformations map ellipses to ellipses, parabolas to parabolas, and hyperbolas to hyperbolas.

Proof: Consider the non-degenerate conic with equation

$$Ax^2 + Bxy + Cy^2 + Fx + Gy + H = 0 \tag{6}$$

and its image under an affine transformation $t : \mathbf{x} \mapsto \mathbf{x}'$ given by

$$\mathbf{x}' = \mathbf{A}\mathbf{x} + \mathbf{b}$$

where $\mathbf{A}$ is an invertible 2×2 matrix. The inverse affine transformation $t^{-1} : \mathbf{x}' \mapsto \mathbf{x}$ is given by

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{x}' - \mathbf{A}^{-1}\mathbf{b}$$

which we may write in the form

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix}$$

for some real numbers p, q, r, s, u and v. It follows that

$$x = px' + qy' + u \quad \text{and} \quad y = rx' + sy' + v \tag{7}$$

If we now substitute these expressions for x and y into equation

(6), then the resulting equation is a second-degree equation in x' and y', so the image of the conic under the affine transformation t must be another conic.

Next we show that this image conic cannot be degenerate, A degenerate image would consist of a pair of lines, a single line, a point, or the empty set. Since the affine transformation t^{-1} maps lines to lines, it would map the degenerate image to another degenerate conic. But this cannot happen since t^{-1} maps the image back to the original non-degenerate conic (6). It follows that the image of (6) cannot be degenerate.

Finally, if we substitute for x and y from equations (7) into equation (6), and keep careful track of the algebra involved, it turns out that the discriminant of the image conic is just

$$(ps - nq)^2 \left(B^2 - 4AC\right)$$

Here $B^2 - 4AC$ is the discriminant of the original conic. Since $(ps - nq)^2 > 0$, the sign of the discriminant is not changed by an affine transformation of a conic. Hence the type of the conic is also unchanged. $\square$

We can combine the results of Theorems $1-4$ to obtain the following corollary.

Corollary. In affine geometry:

(a) all ellipses are congruent to each other:

(b) all hyperbolas are congruent to each other,

(c) all parabolas are congruent to each other.

Non-degenerate conics are congruent only to non-degenerate conics of the same type.

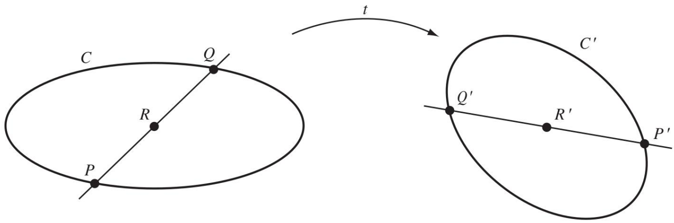
The corollary shows that affine-congruence partitions the set of nondegenerate conics into three disjoint equivalence classes. One class consists of all the ellipses, another class consists of all the hyperbolas, and the third consists of all the parabolas. Each class contains one of the so-called standard conics $x^2 + y^2 = 1, xy = 1$ and $y^2 = x$.

Just as the Fundamental Theorem of Affine Geometry enables us to deduce a given result about an arbitrary triangle by showing that the result holds for an equilateral triangle, so the corollary enables us to deduce a given result about an arbitrary ellipse, hyperbola or parabola by showing that the result holds for the corresponding standard conic. Of course, this works only if the result is concerned with the affine properties of the conic, so we need to be able to recognize such properties.

The following theorem shows that one such property is the property of being the centre of an ellipse or hyperbola.

Theorem 5. Let t be an affine transformation, and let C be an ellipse or hyperbola with centre R. Then $f(C)$ has centre $t(R)$.

Proof: Let C' and R' be the images of C and R under t. If P' is any point on C', then it must be the image of some point P on C. Since R is the centre of C, we can rotate P about R through an angle π to a point Q which must also lie on C. Hence $Q' = t(Q)$ is a point on C'.

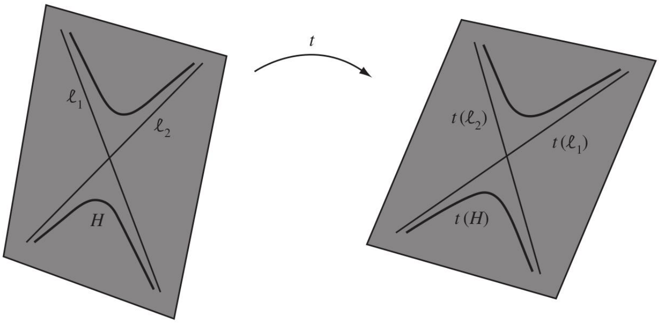


Now t preserves ratios of lengths along lines, so the line segment PRQ maps onto the line segment $P'R'Q'$ with $P'R' = R'Q'$. Thus if we rotate P' about R' through an angle π, it must go to Q' on C'. Now, as our choice for P' as a point on C' varies, so do $P = t^{-1}(P')$ and Q. but the point R is always the same point. It follows that the midpoint of $P'Q'$ is always the same point $R' = t(R)$. Hence $R' = t(R)$ is the centre of C', as required. $\square$

Another affine property is the property of being an asymptote of a hyperbola.

Theorem 6. Let t be an affine transformation, and let H be a hyperbola with asymptotes ℓ_1 and ℓ_2. Then $t(H)$ has asymptotes $t(\ell_1)$ and $t(\ell_2)$

The figure below illustrates that this theorem is plausible for parallel projections.



Proof: The hyperbola H possesses exactly two (distinct) families of parallel lines each of which fills the plane, with each member of each family meeting H exactly once - apart from one line in each family that is an asymptote of H, and so does not meet H.

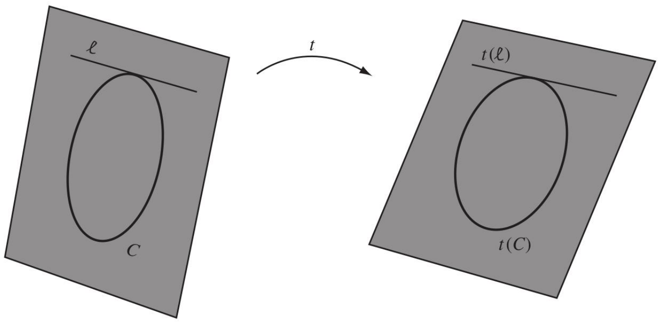
The image of H under the affine transformation t is also a hyperbola, $t(H)$. The images under t of the two families of parallel lines are also (distinct) families of parallel lines; within each family, a line that meets H once is mapped onto a line that meets $t(H)$ once, and the single line that does not meet H maps onto a line that does not meet $t(H)$. So the two exceptional lines in the image families must be the asymptotes of the hyperbola $t(H)$.

It follows that the asymptotes of H are mapped by t to the asymptotes of $t(H)$, as required. $\square$ Many of the problems concerning conics which are particularly amenable to solution using the methods of affine geometry involve tangents.

This is due to the following theorem, which asserts that tangency is an affine property.

Theorem 7. Let t be an affine transformation, and let ℓ be a tangent to a conic C. Then $t(\ell)$ is a tangent to the conic $t(C)$.

The figure below illustrates the theorem for parallel projections.



Solution: We shall use the fact that a tangent to a conic (whether it is an ellipse, a hyperbola or a parabola) intersects the conic at exactly one point.

First, the image of an ellipse E under an affine transformation t is an ellipse. A tangent to E is a line that intersects E in exactly one point. These properties remain unchanged under an affine projection; hence the image of a tangent to E under an affine transformation t must be a tangent to $t(E)$.

Next, the image of a hyperbola H under an affine transformation t is a hyperbola. A tangent to H is a member of a family of parallel lines that fill the plane such that there are lines in the family that meet H twice, once and not at all; there are exactly two lines in the family that meet H exactly once, and these are tangents to H. The image of the family of lines under t is again a family of parallel lines that fill the plane; it contains lines that meet the parabola $t(H)$ twice and not at all, and exactly two lines that meet H exactly once. These lines are the images of the original tangents to H, and must themselves be tangents to $t(H)$. Hence, the image of a tangent to H under an affine transformation t must be a tangent to $t(H)$.

Finally, the image of a parabola P under an affine transformation t is a parabola. A tangent to P is a member of a family of parallel lines that fill the plane such that there are lines in the family that meet P twice, once and not at all; the tangent is the unique member of the family that meets P exactly once. The image of the family of lines under t is again a family of parallel lines that fill the plane; it contains lines that meet the parabola $t(P)$ twice and not at all, and a single line that meets P exactly once. This line is the image of the original tangent to P, and must itself be a tangent to $t(P)$. Hence, the image of a tangent to P under an affine transformation t must be a tangent to $t(P)$.

This completes the proof. $\square$

In applications we often use the following facts that you met earlier

Tangents to Conics in Standard Form The equation of the tangent to a standard conic at the point (x_1, y_1) is as follows.

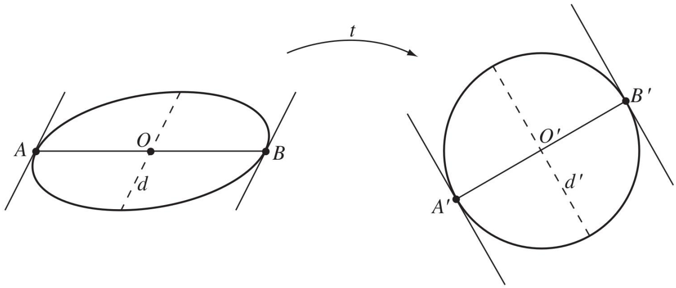
Conic	Tangent
Unit circle $x^2 + y^2 = 1$	x $x_1 + yy_1 = 1$
Rectangular hyperbola $xy = 1$	$xy_1 + yx_1 = 2$
Parabola $y^2 = x$	$2yy_1 = x + x_1$

2.5.2 Applying Affine Geometry to Conics

We are now in a position to apply the methods of affine geometry to th solution of problems involving conics. Of course, affine geometry can be helpful in this task only if the property being investigated is one which preserved under affine transformations. The underlying idea is that we use an affine transformation to map the original conic onto one of our standard conics, tackle the problem in hand there, and then map back to the origin: conic. Affine Transformations and Conics

Example 1. AB is a diameter of an ellipse. Prove that the tangents to the ellipse at A and B are parallel to the diameter conjugate to AB.

Solution: First, map the ellipse onto the unit circle, by an affine transformation t. Since the centre O of the ellipse maps to the centre O' of the circle, the image of the diameter AB is a diameter $A'B'$ of the unit circle.

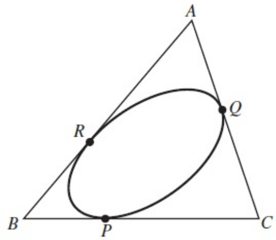


All chords of the circle that are parallel to the tangents at A' and B' are bisected by $A'B'$, and so the diameter through O' is the diameter conjugate to $A'B'$. Since parallel lines map to parallel lines and ratios along parallel lines are preserved under the inverse affine transformation t^{-1}, it follows that all chords of the ellipse that are parallel to the tangents at A and B are bisected by AB, and so the diameter through O that is parallel to the tangents at A and B is the diameter conjugate to AB. $\square$

Problem 1. An ellipse touches the sides BC, CA and AB of $\triangle ABC$ at the points P, Q and R, respectively. Prove that

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = 1$$

and deduce that the lines AP, BQ and CR are concurrent.



Solution: First, map the ellipse onto the unit circle, by some affine transformation t. Since tangency is preserved by affine transformations, the image under t of the triangle $\triangle ABC$ is another triangle $\Delta A'B'C'$, whose sides are tangential to the unit circle.

These sides touch the unit circle at $P' = t(P), Q' = t(Q)$ and $R' = t(R)$.

By Problem 1 of Section 2.1, the two tangents from a point to a circle are of equal length, and so (ignoring the directions of line segments)

$$A'Q' = A'R', \quad B'P' = B'R' \quad and \quad C'P' = C'Q'$$

In terms of signed distances, it follows that

$$\frac{A'R'}{R'B'} \cdot \frac{B'P'}{P'C'} \cdot \frac{C'Q'}{Q'A'} = \pm 1;$$

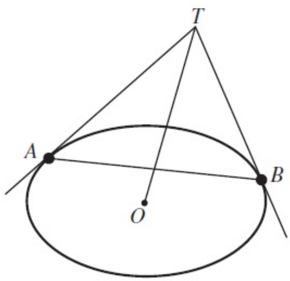
in fact the product must equal 1 since P, Q and R are internal points of the sides of the triangle and therefore each of the above three fractions is positive.

Since ratios of lengths along a line are not changed by the inverse affine transformation t^{-1}, we deduce that

$$\frac{AR}{RB} \cdot \frac{BP}{PC} \cdot \frac{CQ}{QA} = 1,$$

It follows from the converse to Ceva's Theorem that the lines AP, BQ and CR are concurrent. $\square$

Problem 2. The tangents to an ellipse at two points A and B meet at a point T. Prove that the line joining T to the centre O of the ellipse bisects the chord AB.

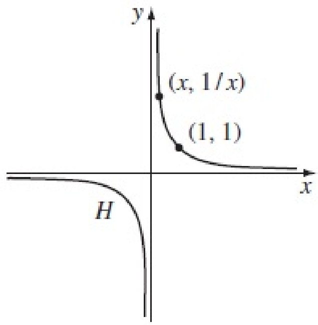


Solution: First, map the ellipse onto the unit circle, by some affine transformation t. Since tangency is preserved by affine transformations, the images under t of the tangents TA and TB are tangents $T'A'$ and $T'B'$ to the unit circle.

Let P' be the point of intersection of the chord $A'B'$ with the line joining T' to O', the centre of the unit circle. By symmetry, the triangles $\Delta T'A'P'$ and $\Delta T'B'P'$ are Euclidean- congruent and so $A'P' = B'P'$; in other words, P' is the midpoint of $A'B'$.

Let the line joining T to the centre O of the ellipse meet AB at P. Then, since P' is the midpoint of $A'B'$ and since midpoints of line segments and centres of ellipses are preserved by the inverse transformation $t^{-1}, P = t^{-1}(P')$ is the midpoint of $AB = t^{-1}(A'B')$. Hence OP bisects all chords of H that are parallel to ℓ. □

The rectangular hyperbola $H = \{(x, y) : xy = 1\}$ does not possess as much symmetry as does the unit circle; so the fact that every hyperbola is affinecongruent to H may not be sufficient to simplify a given problem. Fortunately. however, we can also arrange for any given point on the original hyperbola to map to the point $(1, 1)$ on H.



To see this, note that for any non-zero number a, the affine transformation

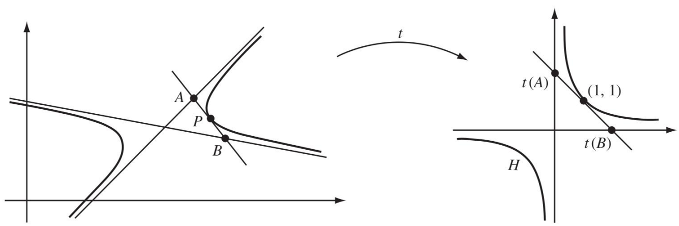
$$t_a : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

maps H to itself. For, an arbitrary point on H has coordinates of the form $(x, 1/x), x \neq 0$, and under t_a this is mapped to the point $(ax, 1/ax)$, which also lies on H. As x varies through $\mathbb{R} - \{0\}$, its image $(ax, 1/ax)$ varies over the whole of H, so the image of H under t_a is the whole of H.

So if we start with a given hyperbola and a point P on it, we can map the hyperbola to H by some affine transformation s. The point $s(P)$ will then have coordinates $(b, 1/b)$ for some number $b \in \mathbb{R} - \{0]$; so if we choose $a = 1/b$, then the affine transformation t_a will map $s(P)$ to $(1, 1)$. Overall, the composite $t = t_a \circ$ s is an affine transformation which maps the given hyperbola to H, and maps P to $(1, 1)$. We now state this as a corollary to Theorem 2 .

Corollary. Given any hyperbola and a point P on it, there is an affine transformation which maps the hyperbola onto the rectangular hyperbola $xy = 1$, and the point P to $(1, 1)$.

Example 2. The tangent at the point P on a hyperbola meets the asymptotes at the points A and B. Prove that $PA = PB$.



Solution: Let t be an affine transformation which maps the hyperbola onto the rectangular hyperbola $H = \{(x, y) : xy =$

$1\}$ in such a way that $t(P) = (1,1)$. Then, by Theorem 6 of Subsection 2.5.1, the asymptotes of the hyperbola map to the asymptotes of H; and, by Theorem 7 of Subsection 2.5.1, the tangent at P maps to the tangent at $(1,1)$.

By symmetry, $(1,1)$ is the midpoint of the line segment from $t(A)$ to $t(B)$. Since midpoints are preserved under the affine transformation t^{-1}, it follows that P is the midpoint of AB. $\square$

Problem 3. P is a point on a hyperbola H with centre O. Prove that there exists a line ℓ through O such that all chords of the hyperbola which are parallel to ℓ are bisected by OP.

Solution: First, map the hyperbola H onto the rectangular hyperbola $H' = \{(x,y) : xy = 1\}$ by some affine transformation t, in such a way that t maps P to the point $(1,1)$. Since the property of being the centre of the hyperbola is preserved under affine transformations, t maps the centre, O, of H to the centre of H', namely the origin.

Let m' be the image of OP under t. Then m' passes through the origin and the point $(1,1)$, so its equation is $y = x$. Clearly, H' is symmetric with respect to m'. Now let ℓ' be the line with equation $y = -x$; this is perpendicular to m'. By symmetry, m' bisects all chords of the rectangular hyperbola which are parallel to ℓ'.

But the properties of parallelism and of ratios along a line are preserved by affine transformations, so if ℓ is the line

$t^{-1}(\ell')$, then OP bisects all chords of H which are parallel to ℓ. $\square$

2.6 Exercises

Section 2.1

1. Let $\triangle ABC$ be a triangle in which $AB = AC$. Prove that

$$\angle ABC = \angle ACB.$$

   Hint: Consider a reflection in the bisector of $\angle BAC$.

2. Determine which of the following transformations $t : \mathbb{R}^2 \to \mathbb{R}^2$ are Euclidean transformations.

   (a) $t(\mathbf{x}) = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \mathbf{x} + \begin{pmatrix} -3 \\ 1 \end{pmatrix}$

   (b) $t(\mathbf{x}) = \begin{pmatrix} -\frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \mathbf{x} + \begin{pmatrix} 3 \\ 2 \end{pmatrix}$

   (c) $t(\mathbf{x}) = \begin{pmatrix} -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{pmatrix} \mathbf{x} + \begin{pmatrix} 2 \\ -3 \end{pmatrix}$

3. The Euclidean transformations t_1 and t_2 are given by

$$t_1(\mathbf{x}) = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{pmatrix} \mathbf{x} + \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

   and

$$t_2(\mathbf{x}) = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{3}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \mathbf{x} + \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

   Determine the composites $t_1 \circ t_2$ and $t_2 \circ t_1$.

4. Determine the inverse of each of the following Euclidean transformations.

(a) $t(\mathbf{x}) = \begin{pmatrix} \frac{5}{13} & -\frac{12}{13} \\ \frac{12}{13} & \frac{5}{13} \end{pmatrix} \mathbf{x} + \begin{pmatrix} -4 \\ 5 \end{pmatrix}$

(b) $t(\mathbf{x}) = \begin{pmatrix} -\frac{12}{13} & -\frac{5}{13} \\ -\frac{5}{13} & \frac{12}{13} \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

5. The Euclidean transformations t_1 and t_2 are given by

$$t_1(\mathbf{x}) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

and

$$t_2(\mathbf{x}) = \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Determine the composite $t_2^{-1} \circ t_1$.

Section 2.2

1. Determine whether or not each of the following transformations $t : \mathbb{R}^2 \to \mathbb{R}^2$ is an affine transformation.

(a) $t(\mathbf{x}) = \begin{pmatrix} 2 & -2 \\ -3 & 3 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 2 \\ -1 \end{pmatrix}$

(b) $t(\mathbf{x}) = \begin{pmatrix} 5 & -2 \\ -2 & 5 \end{pmatrix} \mathbf{x} + \begin{pmatrix} -3 \\ -1 \end{pmatrix}$

(c) $t(\mathbf{x}) = \begin{pmatrix} -1 & 1 \\ -1 & -2 \end{pmatrix} \mathbf{x}$

2. Write down an example (if one exists) of each type of transformation $t : \mathbb{R}^2 \to \mathbb{R}^2$ described below. In each case, justify your answer.

   (a) An affine transformation t which is not a Euclidean transformation

   (b) A Euclidean transformation t which is not an affine transformation

   (c) A transformation t which is both Euclidean and affine

   (d) A transformation t which is one-one, but is neither Euclidean nor affine

3. The affine transformations t_1 and t_2 are given by

$$t_1(\mathbf{x}) = \begin{pmatrix} 2 & -3 \\ 1 & -1 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

   and

$$t_2(\mathbf{x}) = \begin{pmatrix} -1 & 2 \\ -1 & 1 \end{pmatrix} \mathbf{x} + \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

   Determine the following composites.

   (a) $t_1 \circ t_2$

   (b) $t_2 \circ t_1$

   (c) $t_1 \circ t_1$

4. Determine the inverse of each of the following affine transformations.

(a) $t(\mathbf{x}) = \begin{pmatrix} 2 & -3 \\ 3 & -5 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 2 \\ 4 \end{pmatrix}$

(b) $t(\mathbf{x}) = \begin{pmatrix} 3 & 2 \\ 4 & 2 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ -2 \end{pmatrix}$

5. Prove that the transformation

$$t(\mathbf{x}) = 3\mathbf{x} \left(\mathbf{x} \in \mathbb{R}^2\right)$$

is an affine transformation, but not a parallel projection.

6. Which of the following are affine properties?

(a) distance

(b) collinearity

(c) circularity

(d) magnitude of angle

(e) midpoint of line segment

Section 2.3

1. The affine transformation $t : \mathbb{R}^2 \to \mathbb{R}^2$ is given by

$$t(\mathbf{x}) = \begin{pmatrix} 1 & -1 \\ 2 & -3 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 2 \\ -4 \end{pmatrix}$$

Determine the image under t of each of the following lines.

(a) $y = -2x$

(b) $2y = 3x - 1$

2. The affine transformation $t : \mathbb{R}^2 \to \mathbb{R}^2$ is given by

$$t(\mathbf{x}) = \begin{pmatrix} 4 & 5 \\ 1 & 1 \end{pmatrix} \mathbf{x} + \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Determine the image under t of each of the following lines.

(a) $2x - 5y + 3 = 0$ (b) $3x + y - 4 = 0$

3. Determine the affine transformation which maps the points $(0, 0), (1, 0)$ and $(0, 1)$ to the points:

(a) $(0, -1), (1, 1)$ and $(-1, 1)$, respectively:

(b) $(-4, -5), (1, 7)$ and $(2, -9)$, respectively.

4. Determine the affine transformation which maps the points $(1, 1), (3, 2)$ and $(4, 1)$ to the points $(0, 1), (1, 2)$ and $(3, 7)$, respectively.

5. Determine the affine transformation which maps the points $(1, -1), (5, -4)$ and $(-2, 1)$ to the points $(1, 1), (4, 0)$ and $(0, 2)$, respectively.

6. Prove that the affine transformation t for which

$$t(\mathbf{x}) = \begin{pmatrix} -1 & 2 \\ 3 & -2 \end{pmatrix} \mathbf{x}$$

maps each point of the line $y = x$ in $\mathbb{R}^2$ onto itself.

7. Determine the matrices $\mathbf{A}$ and $\mathbf{b}$ for the affine transformation

$$t(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$$

where $\mathbf{A}$ and $\mathbf{b}$ are 2×2 and 2×1 matrices, respectively, given that t maps each point of the line $y = 0$ onto itself and $(0, 1)$ onto $(2, 3)$. Prove also that t is a parallel projection of $\mathbb{R}^2$ onto itself.

Section 2.4

1. The points P, Q, R and S lie on a line, in that order, the distances between them are 4 units, 2 units and 3 units, respectively. Determine the ratios $PR : RS$ and $PS : SQ$

2. A point X lies inside a triangle $\triangle ABC$, and the lines AX, BX and CX meet the opposite sides of the triangle at P, Q and R, respectively. The ratios $AR : AB$ and $BP : BC$ are $1 : 5$ and $3 : 7$, respectively. Determine the ratio $AC : AQ$

3. Let ℓ be a line that crosses the sides BC, CA and AB of a triangle $\triangle ABC$ at three distinct points P, Q and R, respectively. The ratios $BC : CP$ and $CQ : QA$ are $3 : 2$ and $1 : 3$, respectively. Determine the ratio $AR : RB$.

4. $ABCD$ is a parallelogram, and the point P divides AB in the ratio $2 : 1$; the lines AC and DP meet at Q, and the lines BQ and AD meet at R.

(a) Determine the images of P, Q and R under the affine transformation t which maps A, D and C to $(0,1), (0,0)$ and $(1,0)$, respectively.

(b) By considering the image of $ABCD$ under t, determine the ratios $BQ : QR$ and $AR : RD$.

5. The triangle $\triangle ABC$ has vertices $A(-1,2), B(-3,-1)$ and $C(3,1)$, and the points $P\left(1,\frac{1}{3}\right), Q\left(1,\frac{3}{2}\right)$ and $R\left(-\frac{5}{3},1\right)$ lie on BC, CA and AB. respectively.

   (a) Determine the ratios in which P, Q and R divide the sides of the triangle.

   (b) Determine whether or not the lines AP, BQ and CR are concurrent.

6. The triangle $\triangle ABC$ has vertices $A(2,0), B(-3,0)$ and $C(3,-3)$, and the points $P(-1,-1), Q(1,3)$ and $R\left(-\frac{1}{4},0\right)$ lie on BC, CA and AB, respectively.

   (a) Determine the ratios in which P, Q and R divide the sides of the triangle.

   (b) Determine whether or not the points P, Q and R are collinear.

7. $\triangle ABC$ is a triangle, and X a point which does not lie on any of its (extended) sides. Also, AX meets BC at P, BX meets CA at Q and CX meets BA at R. Prove that
$$\frac{AX}{XP} = \frac{AR}{RB} + \frac{AQ}{QC}$$

(This result is often known as van Aubel's Theorem.)

8. $\triangle ABC$ is a triangle, and X a point which does not lie on any of its (extended) sides. Next, AX meets BC at P, BX meets CA at Q and CX meets BA at R. Also, RQ meets BC at L, PR meets CA at M and PQ meets BA at N. Prove that L, M and N are collinear.
   Hint: Apply the result of Problem 5 in Subsection 2.4.3 to $\triangle ABC$ and points L, M and N in turn. Then evaluate the product $\frac{BL}{LC} \cdot \frac{CM}{MA} \cdot \frac{AN}{NB}$.

9. Three disjoint circles of unequal radii lie in the plane, their centres being non-collinear. Pairs of tangents are drawn to each pair of circles such that the point of intersection of the two tangents to each pair of circles lies beyond the two circles. Prove that the three intersection points are collinear.

Section 2.5

1. An ellipse touches the sides AB, BC, CD, DA of a parallelogram $ABCD$ at the points P, Q, R, S, respectively. Prove that the lengths CQ, QB, BP and CR satisfy the equation
$$\frac{CQ}{QB} = \frac{CR}{BP}.$$

2. Determine the equation of the image of the parabola P with equation $y = x^2$ under the affine transformation

$t : \mathbb{R}^2 \mapsto \mathbb{R}^2$ given by

$$t(\mathbf{x}) = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \mathbf{x}$$

Show that the image of the vertex of P is not the vertex of $t(P)$.

3. Prove that for any triangle $\triangle ABC$ there exists an ellipse that touches the sides AB, BC and CA at their midpoints.

4. Let $P(a\cos\theta, b\sin\theta)$, where θ is not a multiple of $\pi/2$, be a point on the ellipse $C : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, where $a \geq b > 0$; and $P'(a\cos\theta, a\sin\theta)$ the corresponding point on the 'auxiliary circle' $C' : x^2 + y^2 = a^2$. Prove that the tangents at P to C and at P' to C' meet on the x -axis. Hint: Write down an affine transformation that maps C to C' and P to P', and that maps each point of the x -axis to itself.

5. Given any two points P and P' on ellipses E and E', respectively, show that there exists an affine transformation that maps E to E' and P to P'.

6. Find the endpoints of the chord AB of the hyperbola H with equation $xy = 1$ that is bisected by the point $P(2, 1)$.

7. E is the ellipse with equation $\frac{x^2}{9} + \frac{y^2}{4} = 1$, and $P\left(\frac{3}{\sqrt{3}}, \frac{2}{\sqrt{3}}\right)$ is a point inside $E.AB$ is a chord of E through P, and O

is the centre of E. Find the maximum value of $\frac{AP}{PB}$ as A varies on E.

MODULE 3

PROJECTIVE GEOMETRY: LINES

Geometry is one branch of mathematics that has an obvious relevance to the 'real world'. Earlier, we studied some results in Euclidean geometry and we described the group of Euclidean transformations, the isometries. We saw that the Euclidean transformations preserve distances and angles, and have a definite physical significance.

In this chapter we study projective geometry, a very different type of geometry, that has important but less obvious applications. It was discovered through artists' attempts over many centuries to paint realistic-looking pictures of scenes composed of objects situated at differing distances from the eye. How can three-dimensional scenes be represented on a two-dimensional canvas? Projective geometry explains how an eye perceives 'the real world', and so explains how artists can achieve realism in their work.

3.1 Perspective

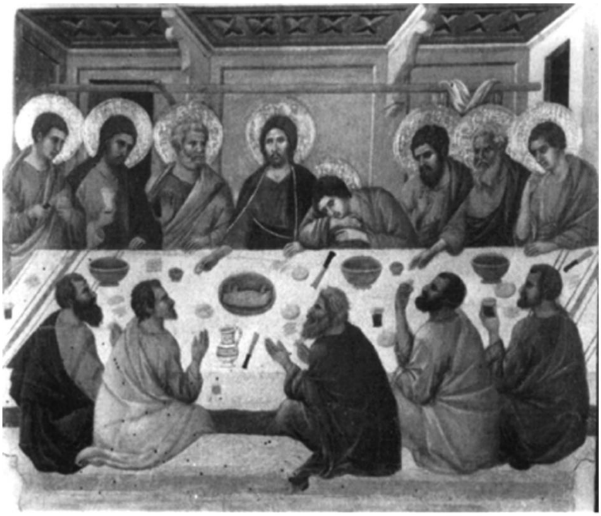
3.1.1 Perspective in Art

The first 'pictures' were probably Cave Art wall paintings: for example, depictions of animals and hunters. Up to the Middle Ages, most pictures were drawn on walls, floors or ceilings of buildings and were intended to convey messages rather than to be accurate illustrations of what an eye might see. For example, Christian religious art portrayed Christ and the Saints, the Bayeux tapestry outlined events such as the Norman Conquest and the Battle of Hastings, and so on.

To the modern eye, the people and animals in these pictures appear to be rather stylized, and the whole scene seems very two-dimensional. The events illustrated do not appear to be properly integrated into the background, even if this is included.

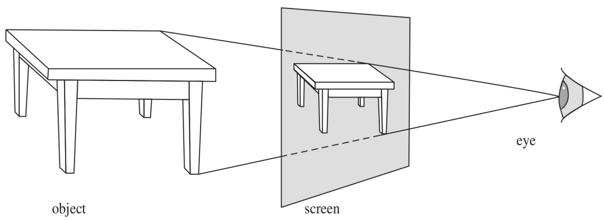
Towards the end of the 13th century, early Renaissance artists began to attempt to portray ‘real’ situations in a realistic way. For example, people at the back of a group would be drawn higher up than those at the front – a technique known as terraced perspective.

As artists struggled to find better techniques to improve the realism of their work, the idea of vertical perspective was developed by the Italian school of artists (including Duccio (1255–1318) and Giotto (1266–1337)). To create an impression of depth in a scene, the artist would represent pairs of parallel lines that are symmetrically placed either side of the scene by lines that meet on the centre line of the picture. The method is not totally realistic, since objects do not appear to recede into the distance in the way that might be expected. The problem of depicting ‘distant objects looking smaller’, with a properly integrated foreground and background, was tackled by many artists, including notably Ambrogio Lorenzetti (c. 1290–1348).

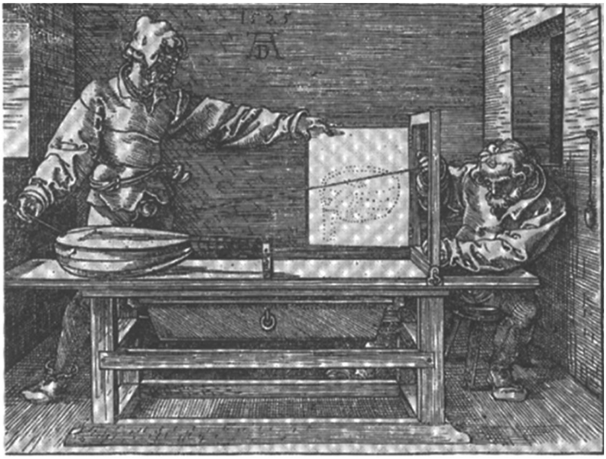


The modern system of focused perspective was discovered around 1425 by the sculptor and architect Brunelleschi (1377–1446), developed by the painter and architect Leone Battista Alberti (1404–1472), and finally perfected by Leonardo da Vinci (1452–1519).

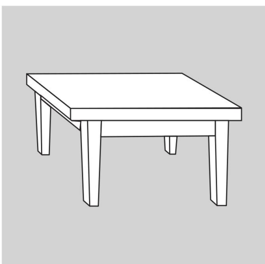
These artists realized that what the eye actually ‘sees’ of a scene are the various rays of light travelling from each point in the scene to the eye. An effective way of deciding how to depict a three-dimensional scene on a two-dimensional canvas so as to create a realistic impression is therefore as follows. Imagine a glass screen placed between the eye and the three-dimensional scene. Each line joining the eye to a point of the scene pierces the glass screen at some point. The set of all such points forms an image on the screen known as a cross-section. Since the eye cannot distinguish between light rays coming from the points of the actual scene and light rays coming from the corresponding points of the cross-section (since these are in exactly the same direction), the cross-section produces the same impression as the original scene. In other words, the cross-section gives a realistic two-dimensional representation of the three-dimensional scene.

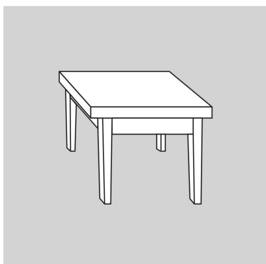


The German artist Albrecht Dürer (1471–1528) introduced the term perspective (from the Latin verb meaning 'to see through') to describe this technique, and illustrated it by a series of well-known woodcuts in his book Underweysung der Messung mit dem Zyrkel und Rychtsscheyed (1525). The Dürer woodcut below shows an artist peering through a grid on a glass screen to study perspective and the effect of foreshortening.



Of course, the picture displayed on the screen is just one representation of the scene. If the screen is placed closer to, or further away from, the eye, the size of the cross-section changes. Also, the screen may be placed at a different angle for a given position of the eye, or the eye itself may be moved to a different position. In each case, a different cross-section is obtained, though they are all related to each other.

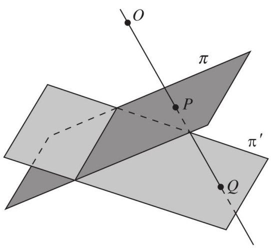


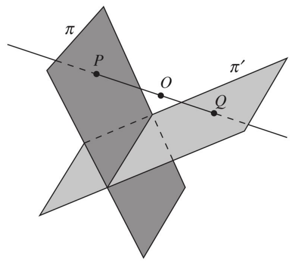


3.1.2 Mathematical Perspective

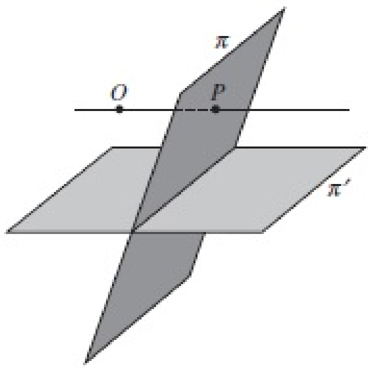
To help us understand the relationship between different representations of a scene, we now look at perspective from a mathematical point of view. In place of an eye and light rays travelling to it, we use the family of all lines in $\mathbb{R}^3$ through a given point. For convenience, this point will often be the origin O. The glass screen is replaced by a plane in $\mathbb{R}^3$ that does not pass through the origin.

In order to compare the cross-sections that appear on different screens, we consider two planes π and π' that do not pass through O. A point P in π and a point Q in π' are said to be in perspective from O if there is a straight line through O, P and Q. A perspectivity from π to π' centred at O is a function that maps a point P of π to a point Q of π' whenever P and Q are in perspective from O. Notice that the planes π and π' may lie on the same side of O as shown on the left below, or they may lie on opposite sides of O as shown on the right.

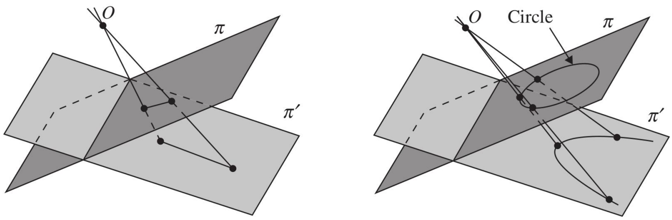




One complication with the above definition of a perspectivity is that the domain of the perspectivity is not necessarily the whole of π. Indeed, if P is any point of π such that OP is parallel to π', as shown in the margin, then P cannot have an image in π', and cannot therefore belong to the domain of the perspectivity. From a mathematical point of view, this need to exclude such exceptional points from the domain of a perspectivity turns out to be rather a nuisance. In Subsection 3.2.3 we shall therefore reformulate the definition of a perspectivity in such a way that these exceptional points can be included in the domain.

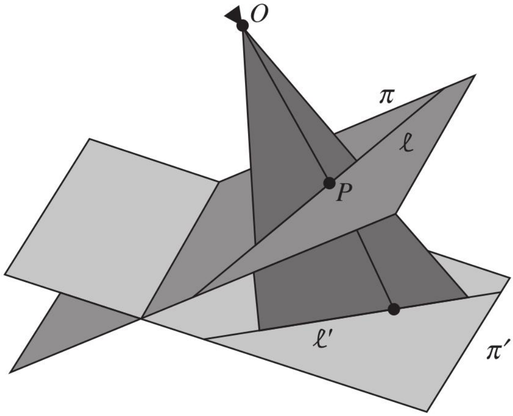


Even with only the preliminary definition of perspectivity given above, it is clear that some features of figures are preserved under a perspectivity, while others are not. For example, the figure on the left below illustrates a particular perspectivity in which a line segment in one plane maps onto a line segment in another plane. This suggests that collinearity is preserved by a perspectivity. On the other hand, the figure on the right illustrates a perspectivity in which a circle in one plane appears to map to a parabolic shape in another plane, which suggests that 'circularity' is not preserved.

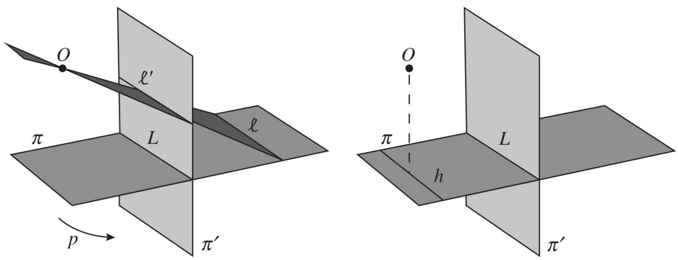


One of our main tasks is to study the images of standard configurations such as lines and conics under perspectivities. This chapter deals with lines; the next chapter deals with conics.

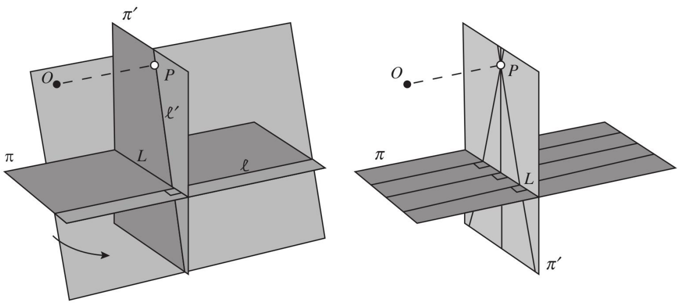
Consider a perspectivity with centre O that maps points in a plane π to points in a plane π'. A convenient way to visualize the image of a line ℓ under the perspectivity is to consider an arbitrary point P on ℓ. As P moves along ℓ, the line OP sweeps out a plane. The line ℓ' where this plane intersects π' is the image of ℓ.



To be specific, consider the perspectivity p with centre O that maps points in a horizontal plane π to points in a vertical plane π', and let L be the line where π and π' intersect. Under p, every line ℓ in π that is parallel to L maps to a horizontal line ℓ' in π'. In particular, L maps to itself. The only exception is the line h that passes through the foot of the perpendicular from O to π. This line does not have an image in π' since the lines joining points of h to O are parallel to π'.

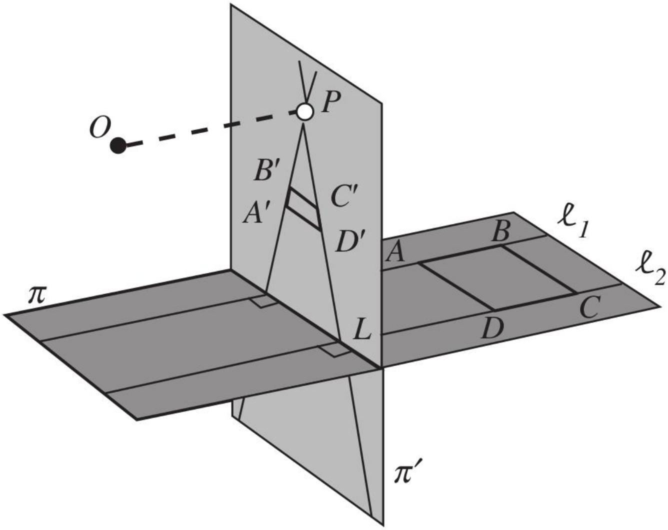


Next, consider the image under the same perspectivity p of a line ℓ in π that is perpendicular to L. To do this, let P denote the foot of the perpendicular from O to the plane π'. Although P is not the image of any point of π, the plane through O and ℓ meets π' in some line ℓ' that passes through P. It follows that the image of ℓ under p is some line ℓ' through P, with the point P itself omitted.



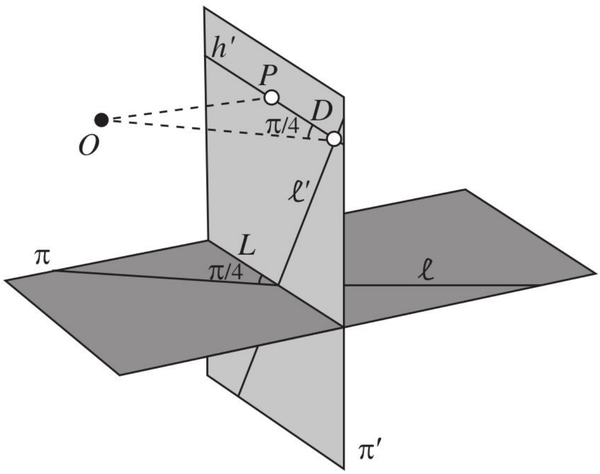
The above argument works for any line in π that is perpendicular to L. All such lines are mapped by the perspectivity p to lines in π' that pass through P, and that omit the point P itself.

We may combine our observations concerning lines in π that are parallel to L or perpendicular to L in the following way. Let $ABCD$ be a rectangle in π on the opposite side of L from O, with sides AB and CD that lie on lines ℓ_1 and ℓ_2, perpendicular to L. Then AD and BC both map onto horizontal lines in π' between L and P. As the side BC recedes from L, its image BC' under the perspectivity p moves further up π' towards P, becoming shorter as it moves.



To an observer whose eye is located at O, the lines ℓ_1 and ℓ_2 appear to meet 'at infinity', and this corresponds to their images under p appearing to meet at P. The point P is called the principal vanishing point of the perspectivity p because the images in π' of all lines in π perpendicular to L appear to vanish there.

In fact, a perspectivity has many vanishing points. For instance, let ℓ be any line in π that intersects L at an angle of $\pi/4$. Now let h' be the horizontal line in π' through P, and let D be the point on h' such that OD is parallel to ℓ. Then the plane through O and ℓ meets π' in some line ℓ' that passes through D. It follows that the image of ℓ under p is a line through D, with the point D itself omitted.

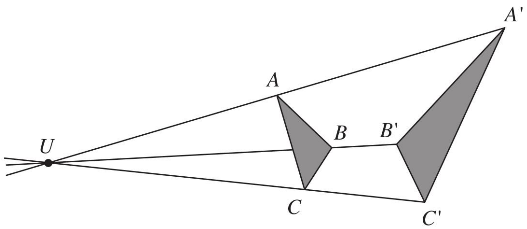


The point D is called a diagonal vanishing point of the perspectivity. All lines in the plane π that are parallel to the given line ℓ have images in π' that lines through D, with the point D itself omitted.

In the same way, each point of the horizontal line h' in π' through P is a vanishing point for the images of all lines in π in some direction; hence the line h' is called the vanishing line. It corresponds to the 'horizon' in the plane -in other words, to the points 'at infinity' towards which an observer's eye is pointing when looking in a horizontal direction.

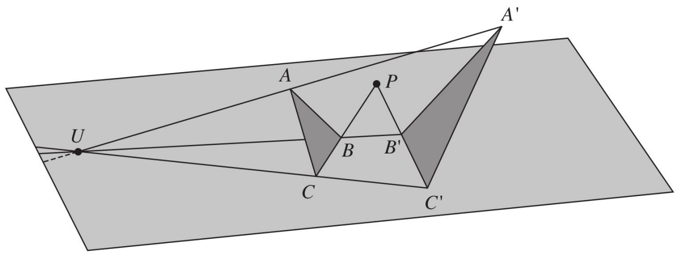
3.1.3 Desargues' Theorem

The idea that information in three dimensions can be related to information in two dimensions, and vice versa, plays an important role in mathematics just as it does in Art. For example, consider the following three-dimensional figure that consists of two triangles $\triangle ABC$ and $\triangle A'B'C'$ which are in perspective from a point U. For the moment we shall assume that no pair of corresponding sides BC and $B'C', CA$ and $C'A'$, and AB and $A'B'$, are parallel.

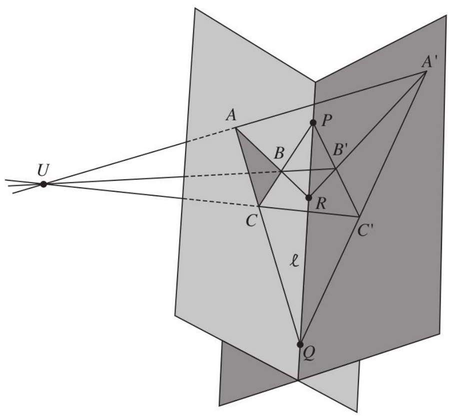


We shall show that this three-dimensional figure has the property that BC and $B'C', CA$ and $C'A', AB$ and $A'B'$ meet at P, Q, R, respectively, where P, Q and R are collinear. This will enable us to formulate an equivalent two- dimensional result, known as Desargues' Theorem.

To prove the three-dimensional result, observe that both BC and $B'C'$ lie in the plane that passes through the points U, B and C. Since BC and $B'C'$ are coplanar but not parallel, they must meet at some point P.



Similarly, the sides CA and $C'A'$ meet at some point Q, and the sides AB and $A'B'$ meet at some point R. Since the points P, Q and R lie both on the plane which contains the triangle $\triangle ABC$ and on the plane which contains the triangle $\triangle A'B'C'$, they must lie on the line ℓ where the two planes meet. It follows that P, Q and R are collinear.



To obtain the equivalent two-dimensional result, imagine that you are viewing the three-dimensional configuration through a transparent screen. Since this viewing process will not alter the collinearity of points or the coincidence of lines, we may reinterpret the three-dimensional result in terms of the image on the screen to obtain the following theorem.

Theorem 1. Desargues' Theorem Let $\triangle ABC$ and $\triangle A'B'C'$ be triangles in $\mathbb{R}^2$ such that the lines AA', BB' and CC' meet at a point U. Let BC and $B'C'$ meet at P, CA and $C'A'$ meet at Q, and AB and $A'B'$ meet at R. Then P, Q and R are collinear.

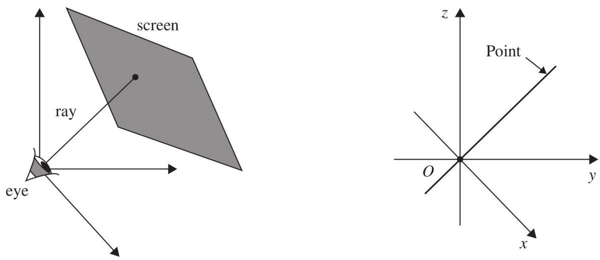
Strictly speaking, we have not proved this theorem since it is not immediately obvious that $\triangle ABC$ and $\triangle A'B'C'$ can be obtained as images of triangles in $\mathbb{R}^3$ which have corresponding sides that are not parallel. Nevertheless, the above argument does provide reasonably convincing evidence that the theorem is true.

One remarkable feature of the above argument is the way in which the geometry of the figure on the transparent screen is characterized by the rays of light that enter an eye. Thus a point on the screen corresponds to a single ray of light that enters the eye, a line on the screen corresponds to a plane of rays of light that enter the eye, and so on. The geometry of the figure can be investigated entirely in terms of these rays of light. The screen is needed only to interpret the result in terms of a two-dimensional figure.

3.2 The Projective Plane $\mathbb{RP}^2$

3.2.1 Projective Points

Imagine an eye situated at the origin of $\mathbb{R}^3$ looking at a fixed screen. As we mentioned in Subsection 3.1.1, each point of the screen corresponds to the ray of light that enters the eye from the point. This correspondence between points of the screen and rays of light through the origin is the clue that we need to define a space of points for our new geometry.



Rather than use the points of the screen directly, we use the rays of light that enable an eye to 'see' the points from the origin. We can express this idea mathematically by defining a projective point to be a Euclidean line in $\mathbb{R}^3$ that passes through the origin. In order to avoid confusion with Euclidean points of $\mathbb{R}^3$, we write Point with a capital P whenever we mean a projective point.

Definition. A **Point** (or **projective point**) is a line in $\mathbb{R}^3$ that passes through the origin of $\mathbb{R}^3$. The **real projective plane** $\mathbb{RP}^2$ is the set of all such Points.

In order to prove results in projective geometry algebraically, we need to have an algebraic notation that can be used to specify the Points of $\mathbb{RP}^2$. To do this, we use the fact that a line ℓ through the origin O in $\mathbb{R}^3$ is uniquely determined once we have specified a Euclidean point (other than O) that lies on ℓ. For example, there is a unique line ℓ in $\mathbb{R}^3$ through O and the point with Euclidean coordinates (4, 2, 6), so we can use these coordinates to specify a projective point. When doing this we write the coordinates in the form $[4, 2, 6]$, with square brackets to indicate that the coordinates refer to a projective point.

Remark 3.2.1. Often we abuse our notation slightly, by talking about 'the Point $[a, b, c]'$ when strictly speaking we should say 'the Point with homogeneous coordinates $[a, b, c]'$.

Notice that the homogeneous coordinates of a Point are not unique. For example, the Point with homogeneous coordinates $[4, 2, 6]$ consists of a line that passes through $(0, 0, 0)$ and $(4, 2, 6)$. But this line also passes through $(-2, -1, -3)$, so $[4, 2, 6]$ and $[-2, -1, -3]$ both represent the same Point.

In general, if (a, b, c) is any point on a line through the origin, and λ is any real number, then $(\lambda a, \lambda b, \lambda c)$ also lies on the line. Moreover, if (a, b, c) is not at the origin and $\lambda \neq 0$, then $(\ \lambda a, \lambda b, \lambda c\ )$ is not at the origin either. It follows that $[a, b, c]$ and $[\lambda a, \lambda b, \lambda c]$ both represent the same Point, for any $\lambda \neq 0$. We express this by writing

$$[a, b, c] = [\lambda a, \lambda b, \lambda c], \qquad \text{for any } \lambda \neq 0 \tag{1}$$

Conversely, if there is no non-zero real number λ such that

$$(a', b', c') = (\lambda a, \lambda b, \lambda c)$$

then (a, b, c) and (a', b', c') cannot lie on the same line through the origin, and so the homogeneous coordinates $[a, b, c]$ and $[a', b', c']$ must represent different Points in $\mathbb{RP}^2$.

Example 1. Which of the following homogeneous coordinates represent the same Point in $\mathbb{RP}^2$ as $[6, 3, 2]$?
(a) $[18, 9, 6]$ (b) $[12, -6, 4]$ (c) $\left[1, \frac{1}{2}, \frac{1}{3}\right]$ (d) $[1, 2, 3]$

Solution: (a) This represents the same Point as $[6, 3, 2]$, for if $\lambda = 3$, then

$$[18, 9, 6] = [6\lambda, 3\lambda, 2\lambda] = [6, 3, 2]$$

(b) This represents a Point different from $[6, 3, 2]$, for there is no λ that satisfies the simultaneous equations

$$12 = 6\lambda, -6 = 3\lambda, 4 = 2\lambda$$

(c) This represents the same Point as $[6, 3, 2]$, for if $\lambda = \frac{1}{6}$, then

$$\left[1, \frac{1}{2}, \frac{1}{3}\right] = [6\lambda, 3\lambda, 2\lambda] = [6, 3, 2]$$

(d) This represents a Point different from $[6, 3, 2]$, for there is no λ that satisfies the simultaneous equations

$$1 = 6\lambda, \quad 2 = 3\lambda, \quad 3 = 2\lambda$$

$\square$

Problem 1. Which of the following homogcncous coordinates represent the same Point in $\mathbb{RP}^2$ as $[1, 2, 3]$?
(a) $[2, 4, 6]$ (b) $[1, 2, -3]$ (c) $[-1, -2, -3]$ (d) $[11, 12, 13]$

Solution:

(a) This represents the same Point as $[1, 2, 3]$, for if $\lambda = 2$, then

$$[2, 4, 6] = [\lambda, 2\lambda, 3\lambda] = [1, 2, 3]$$

(b) This does not represent the same Point as $[1, 2, 3]$, for there is no λ that satisfies

$$1 = \lambda, \quad 2 = 2\lambda, \quad -3 = 3\lambda$$

(c) This represents the same Point as $[1, 2, 3]$, for if $\lambda = -1$, then

$$[-1,-2,-3] = [\lambda, 2\lambda, 3\lambda] = [1,2,3]$$

(d) This does not represent the same Point as $[1,2,3]$, for there is no λ that satisfies

$$11 = \lambda, \quad 12 = 2\lambda, \quad 13 = 3\lambda$$

$\square$

At first sight it may seem rather unsatisfactory that the coordinates of a Point are not unique. However, this ambiguity can often be turned to our advantage. For example, if a calculation yields a Point of $\mathbb{RP}^2$ with fractional homogeneous coordinates such as $\left[1, \frac{1}{2}, \frac{1}{3}\right]$, then the rest of the calculation may be simpler if we 'clear' the fractions and represent the Point by the integer homogeneous coordinates $[6,3,2]$ instead.

Problem 2. For each of the following homogeneous coordinates, find integer homogeneous coordinates which represent the same Point.

(a) $\left[\frac{3}{4}, \frac{1}{2}, -\frac{1}{8}\right]$ (b) $\left[0, 4, \frac{2}{3}\right]$ (c) $\left[\frac{1}{6}, -\frac{1}{3}, -\frac{1}{2}\right]$

Solution: In each case we multiply by the least common multiple of the denominators (or any integer multiple of the least common multiple) to obtain:

(a) $\left[\frac{3}{4}, \frac{1}{2}, -\frac{1}{8}\right] = [6,4,-1]$ (multiply by 8);

(b) $\left[0, 4, \frac{2}{3}\right] = [0, 12, 2]$ (multiply by 3) $= [0, 6, 1]$;

(c) $\left[\frac{1}{6}, -\frac{1}{3}, -\frac{1}{2}\right] = [1, -2, -3]$ (multiply by 6).

$\square$

Given a collection of homogeneous coordinates, it is not always easy to spot those that represent the same Point. In such cases it is sometimes possible to rewrite the coordinates in a form that makes the comparison easier.

Example 2. Determine homogeneous coordinates of the form $[a, b, 1]$ for the Points

$$\begin{aligned}&[2, -1, 4], \quad [4, 2, 8], \quad [2\pi, -\pi, 4\pi],\\ &[200, 100, 400], \quad \left[-\frac{1}{2}, -\frac{1}{4}, -1\right], \quad [6, -9, -12]\end{aligned}$$

Hence decide which homogeneous coordinates represent the same Points.

Solution: According to equation (1), a Point of $\mathbb{RP}^2$ is unchanged if its homogeneous coordinates are multiplied (or divided) by any non-zero real number. Since the third coordinate of each Point is non-zero, we may divide by this third coordinate to obtain homogeneous coordinates of the form $[a, b, 1]$ as follows:

$$[2,-1,4]=\left[\frac{1}{2},-\frac{1}{4},1\right]; \quad [4,2,8]=\left[\frac{1}{2},\frac{1}{4},1\right];$$
$$[2\pi,-\pi,4\pi]=\left[\frac{1}{2},-\frac{1}{4},1\right]; \quad [200,100,400]=\left[\frac{1}{2},\frac{1}{4},1\right];$$
$$\left[-\frac{1}{2},-\frac{1}{4},-1\right]=\left[\frac{1}{2},\frac{1}{4},1\right]; \quad [6,-9,-12]=\left[-\frac{1}{2},\frac{3}{4},1\right].$$

Since $[a,b,1]=[a',b',1]$ if and only if $a=a'$ and $b=b'$, it follows that:

$[2,-1,4]$ and $[2\pi,-\pi,4\pi]$ represent the same Point;

$[4,2,8],[200,100,400]$ and $\left[-\frac{1}{2},-\frac{1}{4},-1\right]$ represent the same Point;

$[6,-9,-12]$ represents none of the other Points. □

Notice that the method used in Example 2 works only if the third coordinates of all the Points are non-zero. If this is not the case, then you may still be able to apply the technique using the first or second coordinates.

Problem 3. Determine homogeneous coordinates of the form $[1,b,c]$ for the Points

$$[2,3,-5], \quad [-8,-12,20], \quad [\sqrt{2},\sqrt{3},-\sqrt{5}]$$
$$[4,-6,10], \quad [-20,-30,50], \quad [74,148,0]$$

Hence decide which homogeneous coordinates represent the same Points.

Solution: Dividing by the first coordinate in each, we obtain:

$$\begin{aligned}
[2,3,-5] &= \left[1,\frac{3}{2},-\frac{5}{2}\right];\\
[-8,-12,20] &= \left[1,\frac{3}{2},-\frac{5}{2}\right];\\
[\sqrt{2},\sqrt{3},-\sqrt{5}] &= \left[1,\sqrt{\frac{3}{2}},-\sqrt{\frac{5}{2}}\right];\\
[4,-6,10] &= \left[1,-\frac{3}{2},\frac{5}{2}\right];\\
[-20,-30,50] &= \left[1,\frac{3}{2},-\frac{5}{2}\right];\\
[74,148,0] &= [1,2,0].
\end{aligned}$$

Hence the homogeneous coordinates

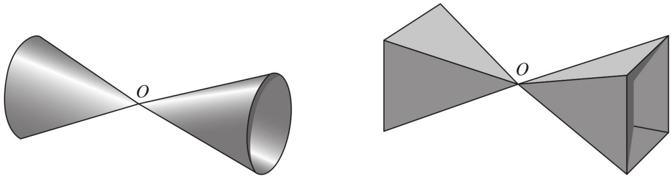
$$[2,3,-5],\quad [-8,-12,20],\quad [-20,-30,50]$$

all represent the same Point. The other homogeneous coordinates represent different Points. □

Having defined projective points, we are now in a position to define a projective figure. Just as a figure in Euclidean geometry is defined to be a subset of $\mathbb{R}^2$, so figures in projective geometry are defined to be subsets of $\mathbb{RP}^2$.

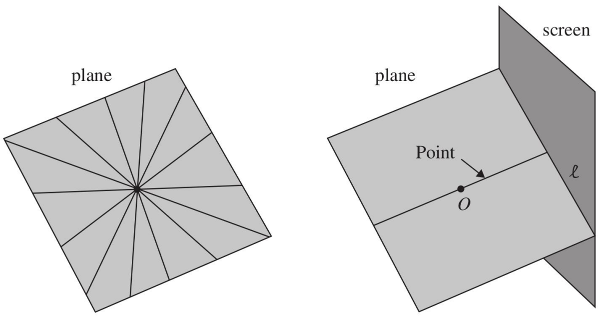
Definition. A projective figure is a subset of $\mathbb{RP}^2$.

Projective figures are just sets of lines in $\mathbb{R}^3$ that pass through the origin. Thus a double cone with a vertex at O, and a double square pyramid with a vertex at O, are both examples of projective figures, for they can both be formed from sets of lines that pass through the origin of $\mathbb{R}^3$.



3.2.2 Projective Lines

A particularly simple type of projective figure is a plane through the origin. Such a plane is a projective figure because it can be formed from the set of all Points (lines through the origin of $\mathbb{R}^3$) that lie on the plane. Since all but one of these Points can be thought of as rays of light that come from a line on a screen, it seems reasonable to define any plane through the origin to be a projective line.



Just as we use 'Point' to refer to a 'projective point', so we use 'Line' to refer to a 'projective line'. The use of a capital L avoids any confusion with lines in $\mathbb{R}^3$.

Definition. A Line (or projective line) in $\mathbb{RP}^2$ is a plane in $\mathbb{R}^3$ that passes through the origin. Points in $\mathbb{RP}^2$ are collinear if they lie on a Line.

Since a Line in $\mathbb{RP}^2$ is simply a plane in $\mathbb{R}^3$ that passes through the origin, it must consist of the set of Euclidean points (x, y, z) that satisfy an equation of the form

$$ax + by + cz = 0$$

where a, b and c are real and not all zero. We can interpret this fact in terms of $\mathbb{RP}^2$ as follows.

Theorem 1. The general equation of a Line in $\mathbb{RP}^2$ is

$$ax + by + cz = 0 \tag{2}$$

where a, b, c are real and not all zero.

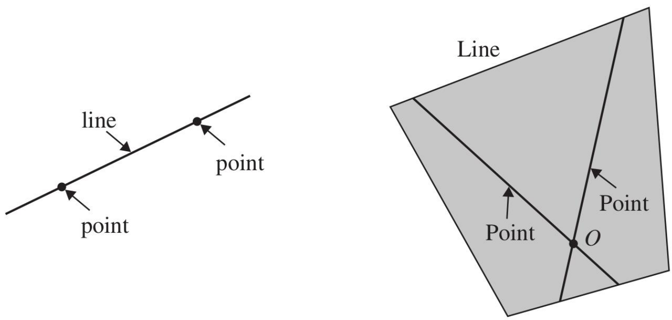
Remark

1. The equation of a Line is not unique, for, if $\lambda \neq 0$, then $\lambda ax + \lambda by + \lambda cz = 0$ is also an equation for the Line. We

can use this fact to 'clear fractions' from the coefficients just as we did for the homogeneous coordinates of a Point.

2. From the figure in the margin it is clear that a Point lies on a Line, or a Line passes through a Point, if and only if the Point has homogeneous coordinates $[x, y, z]$ which satisfy the equation of the Line. For example, $[1, -1, 1]$ lies on the Line $3x + y - 2z = 0$, but $[0, 1, 3]$ does not.

In Euclidean geometry there is a unique line that passes through any two distinct points, as illustrated on the left of the figure below. Similarly, in projective geometry two distinct Points (lines through the origin) lie on a unique Line (plane through the origin).



We express this observation in the form of a theorem, as follows.

Theorem 2. Collinearity Property of $\mathbb{RP}^2$ Any two distinct Points of $\mathbb{RP}^2$ lie on a unique Line.

It is sometimes possible to find an equation for the Line that passes through two distinct Points of $\mathbb{RP}^2$ simply by spotting an equation of the form (2) that is satisfied by the homogeneous coordinates of both Points.

Example 3. For each of the following pairs of Points, write down an equation for the Line that passes through them.

(a) $[3, 2, 0]$ and $[3, 4, 0]$

(b) $[1, 2, 1]$ and $[3, 0, 3]$

(c) $[1, 0, 0]$ and $[0, 0, 1]$

Solution: (a) Both the Points have a z -coordinate equal to 0 , so the homogeneous coordinates must satisfy the equation $z = 0$. This equation is of the form (2) with $a = 0, b = 0$ and $c = 1$, so it must be the required equation for the Line.

(b) The homogeneous coordinates of both Points satisfy $x = z$. This equation is of the form (2) with $a = 1, b = 0$ and $c = -1$. It must therefore be the required equation for the Line.

(c) The homogeneous coordinates of both Points satisfy $y = 0$. This equation is of the form (2) with $a = 0, b = 1$ and $c = 0$, so it must be the required equation for the Line. $\square$

Problem 4. For each of the following pairs of Points, write down an equation for the Line that passes through them.

(a) $[0, 1, 0]$ and $[0, 0, 1]$ (b) $[2, 2, 3]$ and $[3, 3, 7]$

Solution: In each case we seek an equation of the form $ax + by + cz = 0$ which is satisfied by the homogeneous coordinates of the given pair of Points.

(a) An equation for the Line through $[0, 1, 0]$ and $[0, 0, 1]$ is $x = 0$.

(b) An equation for the Line through $[2, 2, 3]$ and $[3, 3, 7]$ is $x = y$. $\square$

But how do we find an equation for a Line through two given Points in cases where it cannot be found by inspection? As an example, consider the Points $[2, -1, 4]$ and $[1, -1, 1]$. We could certainly substitute the values $x = 2$, $y = -1, z = 4$ and $x = 1, y = -1, z = 1$ into equation (2), to obtain the pair of simultaneous equations

$$2a - b + 4c = 0,$$
$$a - b + c = 0.$$

Then subtracting twice the second equation from the first, we obtain $b = -2c$. So from the second equation it follows that $a = -3c$. If we set $c = -1$, say, then $a = 3$ and $b = 2$, so an equation for the Line is

$$3x + 2y - z = 0$$

In this case the calculations are fairly straightforward, but there is an alternative method that is often simpler. Notice that the Line in $\mathbb{RP}^2$ through the Points $[2, -1, 4]$ and $[1, -1, 1]$ is the Euclidean plane in $\mathbb{R}^3$ that contains the position vectors of the points $(2, -1, 4)$ and $(1, -1, 1)$ in $\mathbb{R}^3$. A point (x, y, z) lies in this plane if and only if the vector (x, y, z) is a linear combination of the vectors $(2, -1, 4)$ and $(1, -1, 1)$; in other words, if and only if the vectors $(x, y, z), (2, -1, 4)$ and $(1, -1, 1)$ are linearly dependent.

But three vectors in $\mathbb{R}^3$ are linearly dependent if and only if the 3×3 determinant that has these vectors as its rows is zero. It follows that (x, y, z) lies in the plane containing the position vectors $(2, -1, 4)$ and $(1, -1, 1)$ if and only if

$$\begin{vmatrix} x & y & z \\ 2 & -1 & 4 \\ 1 & -1 & 1 \end{vmatrix} = 0.$$

Translating this statement back into a statement concerning $\mathbb{RP}^2$, we deduce that the Point $[x, y, z]$ lies on the Line through the Points $[2, -1, 4]$ and $[1, -1, 1]$ if and only if

$$\begin{vmatrix} x & y & z \\ 2 & -1 & 4 \\ 1 & -1 & 1 \end{vmatrix} = 0.$$

Expanding this determinant in terms of the entries in its first row. we obtain

$$\begin{aligned} \begin{vmatrix} x & y & z \\ 2 & -1 & 4 \\ 1 & -1 & 1 \end{vmatrix} &= x \begin{vmatrix} -1 & 4 \\ -1 & 1 \end{vmatrix} - y \begin{vmatrix} 2 & 4 \\ 1 & 1 \end{vmatrix} + z \begin{vmatrix} 2 & -1 \\ 1 & -1 \end{vmatrix} \\ &= 3x + 2y - z \end{aligned}$$

Hence an equation for the required Line in $\mathbb{RP}^{\nvDash}$ is

$$3x + 2y - z = 0 \qquad (3)$$

Remark

It is always sensible to check your arithmetic by checking that the two given Points actually lie on the Line that you have found. For instance, the answer above is correct, since equation (3) is a homogeneous linear equation in x, y and z, and the equation is satisfied by $x = 2, y = -1, z = 4$ and by $x = 1$, $y = -1, z = 1$.

We may summarize the above method in the form of a strategy, as follows.

> **Strategy.** To determine an equation for the Line in $\mathbb{RP}^2$ through the Points $[d, c, f]$ and $[g, h, k]$:
> 1. write down the equation
>
> $$\begin{vmatrix} x & y & z \\ d & e & f \\ g & h & k \end{vmatrix} = 0$$
>
> 2. expand the determinant in terms of the entries in its first row to obtain the required equation in the form $ax + by + cz = 0$.

Example 4. Find an equation for the Line that passes through the Points $[1, 2, 3]$ and $[2, -1, 4]$.

Solution: An equation for the Line is

$$\begin{vmatrix} x & y & z \\ 1 & 2 & 3 \\ 2 & -1 & 4 \end{vmatrix} = 0$$

Now

$$\begin{aligned} \begin{vmatrix} x & y & z \\ 1 & 2 & 3 \\ 2 & -1 & 4 \end{vmatrix} &= x \begin{vmatrix} 2 & 3 \\ -1 & 4 \end{vmatrix} - y \begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix} + z \begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix} \\ &= 11x + 2y - 5z. \end{aligned}$$

An equation for the Line is therefore

$$11x + 2y - 5z = 0$$

$\square$

Problem 5. Determine an equation for each of the following Lines in $\mathbb{RP}^2$:

(a) the Line through the Points $[2, 5, 4]$ and $[3, 1, 7]$;

(b) the Line through the Points $[-2, -4, 5]$ and $[3, -2, -4]$.

Solution: We use the strategy for determining an equation for the Line through two given Points given in Subsection 3.2.2.

(a) An equation for the Line through the Points $[2, 5, 4]$ and $[3, 1, 7]$ is

$$\begin{vmatrix} x & y & z \\ 2 & 5 & 4 \\ 3 & 1 & 7 \end{vmatrix} = 0$$

Now

$$\begin{vmatrix} x & y & z \\ 2 & 5 & 4 \\ 3 & 1 & 7 \end{vmatrix} = x \begin{vmatrix} 5 & 4 \\ 1 & 7 \end{vmatrix} - y \begin{vmatrix} 2 & 4 \\ 3 & 7 \end{vmatrix} + z \begin{vmatrix} 2 & 5 \\ 3 & 1 \end{vmatrix}$$
$$= 31x - 2y - 13z,$$

so an equation for the Line is

$$31x - 2y - 13z = 0$$

(b) An equation for the Line through the Points $[-2, -4, 5]$ and $[3, -2, -4]$ is

$$\begin{vmatrix} x & y & z \\ -2 & -4 & 5 \\ 3 & -2 & -4 \end{vmatrix} = 0$$

Now

$$\begin{aligned} \begin{vmatrix} x & y & z \\ -2 & -4 & 5 \\ 3 & -2 & -4 \end{vmatrix} &= x \begin{vmatrix} -4 & 5 \\ -2 & -4 \end{vmatrix} - y \begin{vmatrix} -2 & 5 \\ 3 & -4 \end{vmatrix} + z \begin{vmatrix} -2 & -4 \\ 3 & -2 \end{vmatrix} \\ &= 26x + 7y + 16z \end{aligned}$$

so an equation for the Line is

$$26x + 7y + 16z = 0$$

$\square$

A similar technique can be used to check whether three given Points are collinear. Indeed, three Points $[a, b, c], [d, e, f], [g, h, k]$ are collinear if and only if the position vectors of the points $(a, b, c), (d, e, f), (g, h, k)$ are linearly dependent; that is, if and only if

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & k \end{vmatrix} = 0$$

Example 5. Determine whether the Points $[2, 1, 3]$, $[1, 2, 1]$ and $[-1, 4, -3]$ are collinear.

Solution: We have

$$\begin{aligned} \begin{vmatrix} 2 & 1 & 3 \\ 1 & 2 & 1 \\ -1 & 4 & -3 \end{vmatrix} &= 2\begin{vmatrix} 2 & 1 \\ 4 & -3 \end{vmatrix} - 1\begin{vmatrix} 1 & 1 \\ -1 & -3 \end{vmatrix} + 3\begin{vmatrix} 1 & 2 \\ -1 & 4 \end{vmatrix} \\ &= 2(-6-4) - (-3+1) + 3(4+2) \\ &= -20 + 2 + 18 \\ &= 0. \end{aligned}$$

Since this is zero it follows that $[2, 1, 3]$, $[1, 2, 1]$ and $[-1, 4, -3]$ are collinear. $\square$

We summarize the method of Example 5 in the following strategy.

Strategy. To determine whether three Points $[a, b, c]$, $[d, e, f]$, $[g, h, k]$ are collinear:

1. evaluate the determinant $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & k \end{vmatrix}$;
2. the Points $[a, b, c]$, $[d, e, f]$, $[g, h, k]$ are collinear if and

only if this determinant is zero.

Problem 6. Determine whether the following sets of Points are collinear.

(a) $[1,2,3],[1,1,-2],[2,1,-9]$ (b) $[1,2,-1],[2,1,0],[0,-1,3]$

Solution: We use the strategy for deciding whether three Points are collinear given in Subsection 3.2.2.

(a) The Points $[1,2,3]$, $[1,1,-2]$ and $[2,1,-9]$ are collinear if and only if

$$\begin{vmatrix} 1 & 2 & 3 \\ 1 & 1 & -2 \\ 2 & 1 & -9 \end{vmatrix} = 0$$

Now

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 3 \\ 1 & 1 & -2 \\ 2 & 1 & -9 \end{vmatrix} &= 1\begin{vmatrix} 1 & -2 \\ 1 & -9 \end{vmatrix} - 2\begin{vmatrix} 1 & -2 \\ 2 & -9 \end{vmatrix} + 3\begin{vmatrix} 1 & 1 \\ 2 & 1 \end{vmatrix} \\ &= 1(-9+2) - 2(-9+4) + 3(1-2) \\ &= -7 + 10 - 3 \\ &= 0. \end{aligned}$$

It follows that the three given Points are collinear.

(b) The Points $[1,2,-1]$, $[2,1,0]$ and $[0,-1,3]$ are collinear if and only if

$$\begin{vmatrix} 1 & 2 & -1 \\ 2 & 1 & 0 \\ 0 & -1 & 3 \end{vmatrix} = 0$$

Now

$$\begin{aligned}
\begin{vmatrix} 1 & 2 & -1 \\ 2 & 1 & 0 \\ 0 & -1 & 3 \end{vmatrix} &= 1\begin{vmatrix} 1 & 0 \\ -1 & 3 \end{vmatrix} - 2\begin{vmatrix} 2 & 0 \\ 0 & 3 \end{vmatrix} - 1\begin{vmatrix} 2 & 1 \\ 0 & -1 \end{vmatrix} \\
&= 1(3-0) - 2(6-0) - 1(-2-0) \\
&= 3 - 12 + 2 \\
&= -7 \neq 0
\end{aligned}$$

It follows that the three given Points are not collinear.

□

Before rushing to solve a problem using determinants, you should always stop to see if you can solve the problem more easily by inspection. For example, suppose that you are asked to check whether the Points $[1,0,0], [0,1,0]$, $[1,1,1]$ are collinear. Clearly, $[1,0,0]$ and $[0,1,0]$ lie on the Line $z = 0$, whereas $[1,1,1]$ does not, so the Points are not collinear.

Problem 7. Verify that no three of the Points $[1,0,0], [0,1,0]$ $[0,0,1]$ and $[1,1,1]$ are collinear.

Solution: We have already shown that $[1,0,0], [0,1,0]$, $[1,1,1]$ are not collinear, so this leaves three other cases to consider.

First we check that $[1,0,0],[0,0,1],[1,1,1]$ are not collinear. This follows because $[1,0,0]$ and $[0,0,1]$ lie on the Line $y=0$, whereas $[1,1,1]$ does not.

Next we check that $[0,1,0],[0,0,1],[1,1,1]$ are not collinear. This follows because $[0,1,0]$ and $[0,0,1]$ lie on the Line $x=0$, whereas $[1,1,1]$ does not.

Finally we check that $[1,0,0],[0,1,0]$, $[0,0,1]$ are not collinear. This follows because $[1,0,0],[0,1,0]$ lie on the Line $z=0$, whereas $[0,0,1]$ does not. $\square$

The Points that you considered in Problem 7 play an important part in our development of the theory of projective geometry, so we give them special names.

Definition. The Points $[1,0,0],[0,1,0],[0,0,1]$ are known as the triangle of reference. The Point $[1,1,1]$ is called the unit Point.

Next, observe that any two distinct Lines necessarily meet at a unique Point. Indeed, a Line in $\mathbb{RP}^2$ is simply a plane in $\mathbb{R}^3$ that passes through the origin, and two distinct planes through the origin of $\mathbb{R}^3$ must intersect in a unique Euclidean line through the origin; that is, in a Point. This is very different to the situation in Euclidean geometry where parallel lines do not meet.

Theorem 3. Incidence Property of $\mathbb{RP}^2$ Any two distinct Lines in $\mathbb{RP}^2$ intersect in a unique Point of $\mathbb{RP}^2$

We can determine the Point of intersection of two Lines simply by solving the equations of the two Lines as a pair of simultaneous equations.

Example 6. Determine the Point of intersection of the Lines in $\mathbb{RP}^2$ with equations $x + 6y - 5z = 0$ and $x - 2y + z = 0$.

Solution: At the Point of intersection $[x, y, z]$ of the two Lines, we have

$$x + 6y - 5z = 0$$
$$x - 2y + z = 0$$

Subtracting the second equation from the first, we obtain

$$8y - 6z = 0$$

so that $y = \frac{3}{4}z$. Substituting this into the second equation, we obtain $x = \frac{1}{2}z$.

It follows that the Point of intersection has homogeneous coordinates $\left[\frac{1}{2}z, \frac{3}{4}z, z\right]$ which we can rewrite in the form $\left[\frac{1}{2}, \frac{3}{4}, 1\right]$ or $[2, 3, 4]$. $\square$

Problem 8. Determine the Point of intersection of each of the following pairs of Lines in $\mathbb{RP}^2$:

(a) the Lines with equations $x-y-z = 0$ and $x+5y+2z = 0$;

(b) the Lines with equations $x+2y-z=0$ and $2x+y-4z=0$.

Solution:

(a) At the Point of intersection $[x, y, z]$ of the two Lines, we have

$$x - y - z = 0, \text{ and} \qquad (*)$$
$$x + 5y + 2z = 0 \qquad (**)$$

Subtracting equation $(*)$ from equation $(**)$, we obtain $6y + 3z = 0$, so $z = -2y$

Next, substituting $-2y$ in place of z in equation $(*)$, we obtain $x - y + 2y = 0$, so $x = -y$.

It follows that the homogeneous coordinates of the Point of intersection are $[-y, y, -2y]$ (where $y \neq 0$), which we may rewrite equivalently as $[-1, 1, -2]$

(b) At the Point of intersection $[x, y, z]$ of the two Lines, we have

$$x + 2y - z = 0, \text{ and} \qquad (a)$$
$$2x + y - 4z = 0 \qquad (b)$$

Subtracting twice equation (a) from equation (b), we obtain $-3y - 2z = 0$, so $y = -\frac{2}{3}z$.

Next, substituting $-\frac{2}{3}z$ in place of y in equation (a), we obtain $x - \frac{4}{3}z - z = 0$, so $x = \frac{7}{3}z$.

It follows that the homogeneous coordinates of the Point of intersection are $\left[\frac{7}{3}z, -\frac{2}{3}z, z\right]$ (where $z \neq 0$), which we may rewrite equivalently as $[7, -2, 3]$.

$\square$

Problem 9. Determine the Point of $\mathbb{RP}^2$ at which the Line through the Points $[1, 2, -3]$ and $[2, -1, 0]$ meets the Line through the Points $[1, 0, -1]$ and $[1, 1, 1]$.

Solution: First, we find equations for the two Lines, using the determinant formula.

An equation for the Line through the Points $[1, 2, -3]$ and $[2, -1, 0]$ is

$$\begin{vmatrix} x & y & z \\ 1 & 2 & -3 \\ 2 & -1 & 0 \end{vmatrix} = 0$$

Now

$$\begin{aligned} \begin{vmatrix} x & y & z \\ 1 & 2 & -3 \\ 2 & -1 & 0 \end{vmatrix} &= x \begin{vmatrix} 2 & -3 \\ -1 & 0 \end{vmatrix} - y \begin{vmatrix} 1 & -3 \\ 2 & 0 \end{vmatrix} + z \begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix} \\ &= x(0-3) - y(0+6) + z(-1-4) \\ &= -3x - 6y - 5z. \end{aligned}$$

Hence an equation for the Line may be written as

$$3x + 6y + 5z = 0 \qquad (i)$$

Next, an equation for the Line through the Points $[1, 0, -1]$ and $[1, 1, 1]$ is

$$\begin{vmatrix} x & y & z \\ 1 & 0 & -1 \\ 1 & 1 & 1 \end{vmatrix} = 0$$

Now

$$\begin{aligned} \begin{vmatrix} x & y & z \\ 1 & 0 & -1 \\ 1 & 1 & 1 \end{vmatrix} &= x \begin{vmatrix} 0 & -1 \\ 1 & 1 \end{vmatrix} - y \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} + z \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} \\ &= x(0+1) - y(1+1) + z(1-0) \\ &= x - 2y + z \end{aligned}$$

Hence an equation for the Line may be written as

$$x - 2y + z = 0 \qquad (ii)$$

At the Point of intersection $[x, y, z]$ of the two Lines, both equations (i) and (ii) hold.

Adding three times equation (ii) to equation (i), we obtain $6x + 8z = 0$, so $z = -\frac{3}{4}x$. Next, substituting $z = -\frac{3}{4}x$ into equation (ii), we obtain $x - 2y - \frac{3}{4}x = 0$, so $y = \frac{1}{8}x$

Hence the homogeneous coordinates for the Point of intersection of the two Lines are $\left[x, \frac{1}{8}x, -\frac{3}{4}x\right]$ (where $x \neq 0$), or, equivalently, $[8, 1, -6]$. $\square$

In some cases we can write down the Point at which two Lines intersect without having to solve any equations at all. For example, the Lines with equations $x = 0$ and $y = 0$ clearly meet at the Point $[0, 0, 1]$.

Problem 10. Determine the Point of $\mathbb{RP}^2$ at which the Line through the Points $[1, 0, 0]$ and $[0, 1, 0]$ meets the Line through the Points $[0, 0, 1]$ and $[1, 1, 1]$.

Solution: In this particular case, the homogeneous coordinates of the Points are particularly simple, so we can write down equations for the two Lines without using determinants.

An equation for the Line through the Points $[1, 0, 0]$ and $[0, 1, 0]$ is $z = 0$ (since this is of the right form, and passes through the two Points).

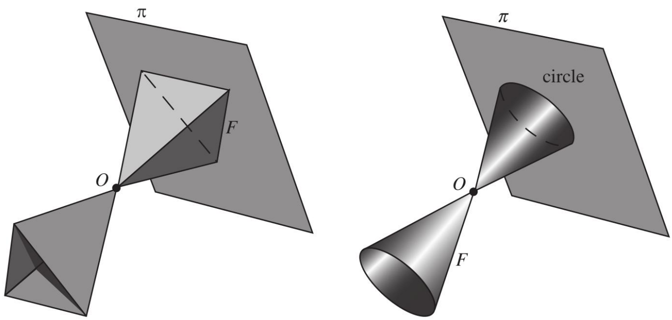
An equation for the Line through the Points $[0, 0, 1]$ and $[1, 1, 1]$ is $x = y$ (since this is of the right form, and passes through the two Points).

The two Lines meet where $z = 0$ and $x = y$, so the homogeneous coordinates for their Point of intersection are $[x, x, 0]$ (where $x \neq 0$), or, equivalently, $[1, 1, 0]$. $\square$

3.2.3 Embedding Planes

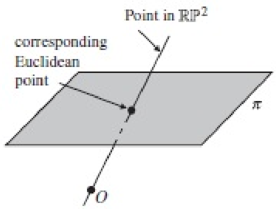
So far we have used three-dimensional space to develop the theory of projective geometry. In practice, however, we want to use projective geometry to study two-dimensional figures in a plane. In order to do this, we now investigate a way of associating figures in a plane with figures in $\mathbb{RP}^2$, and vice versa.

Suppose that a plane π contains a figure F. We can place π into $\mathbb{R}^3$, making sure that it does not pass through the origin, and then construct a corresponding projective figure by drawing in all the Points of $\mathbb{RP}^2$ that pass through the points of F. For example, if F is the triangle shown on the left below, then the corresponding projective figure is a double triangular pyramid. Note that if we change the position of π in $\mathbb{R}^3$, we obtain a different projective figure corresponding to F.

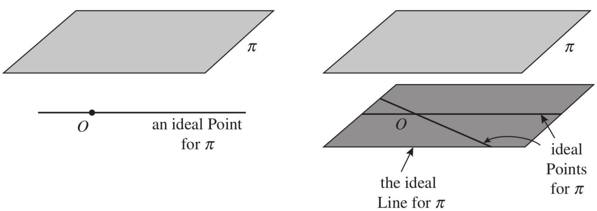


Conversely, suppose that we start with a projective figure F. The corresponding Euclidean figure in π consists of the Euclidean points where the Points of F pierce π. For example, if F is a double cone whose axis is at right angles to the embedding plane, as shown on the right above, then the corresponding Euclidean figure is a circle. Note that if we change the position of π in $\mathbb{R}^3$, we obtain a different plane figure corresponding to F.

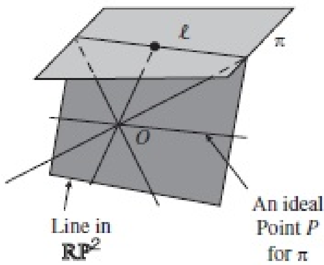
This correspondence between projective figures and Euclidean figures works well provided that each Point of the projective figure pierces the plane π, as shown in the margin. Unfortunately, any Point of $\mathbb{RP}^2$ that consists of a line through the origin parallel to π does not pierce π, and so cannot be associated with a point of π. Such a Point is called an ideal Point for π



All the ideal Points for π lie on a plane through O parallel to π. This plane is a projective line known as the ideal Line for π.



How can we represent a projective figure on π if the figure includes some of the ideal Points for π ? As a simple example, consider the Line illustrated in the margin. This is a projective figure which intersects π in a line ℓ. Every Point of the Line pierces the embedding plane at a point of ℓ except for the ideal Point P which cannot be represented on π. In order to represent the Line completely, we need not only the line ℓ but also the ideal Point P. In other words, the Line is represented by $\ell \cup \{P\}$.



In general, a projective figure can be represented by a figure in π provided that we are prepared to include a subset of Points taken from the ideal Line for π. In order to allow for these additional ideal Points, we introduce the concept of an embedding plane.

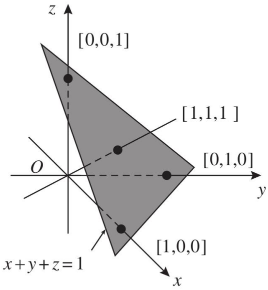
Definition. An embedding plane is a plane, π, which does not pass through the origin, together with the set of all ideal Points for π. The plane in $\mathbb{R}^3$ with equation $z = 1$ is called the standard embedding plane. The mapping of $\mathbb{RP}^2$

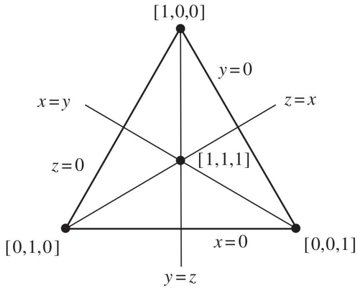
into the standard embedding plane is called the standard embedding of $\mathbb{RP}^2$

We may summarize the above discussion by saying that for a given embedding plane, every projective figure in $\mathbb{RP}^2$ corresponds to a figure in the embedding plane, and vice versa. The figure in the embedding plane may include some ideal Points but is otherwise a Euclidean figure.

If two embedding planes are parallel to each other, the same Points of $\mathbb{RP}^2$ correspond to ideal Points of the embeddings; whereas, if the embed- ding planes are not parallel, different Points of $\mathbb{RP}^2$ correspond to ideal Points of the two embedding planes.

Once we have represented a projective figure in an embedding plane, we can investigate the relationship between its Points and Lines without having to refer to three-dimensional space at all. For example, consider the representation of the triangle of reference and unit Point on the embedding plane $x + y + z = 1$, shown on the left below. If we extract the embedding plane from $\mathbb{R}^3$, as shown on the right, we can use the algebraic theory developed earlier to write down an equation for the Line through any two given Points, without reference to $\mathbb{R}^3$.





Similarly, we can use the algebraic techniques to calculate the homogeneous coordinates of the Point of intersection of any two given Lines.

Problem 11. On the right-hand diagram above, insert the homogeneous coordinates of the Points where the Lines through $[1,1,1]$ meet the sides of the triangle of reference.

Solution: In Problem 10 we found that the Point of intersection of $z = 0$ and $x = y$ is $[1,1,0]$. Similarly, the Point of intersection of $y = 0$

and $z = x$ is $[x, 0, x]$ or $[1, 0, 1]$, and the Point of intersection of $x = 0$ and $y = z$ is $[0, z, z]$ or $[0, 1, 1]$. $\square$

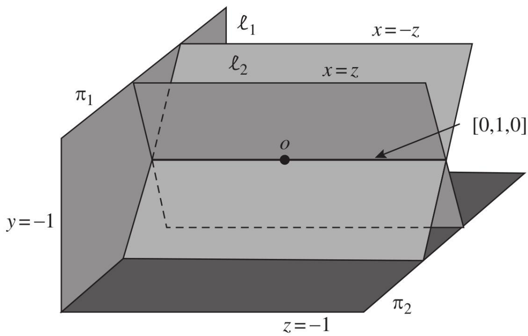
Any plane may be used as an embedding plane provided that it does not pass through the origin. For example, if we take π to be the plane $z = -1$, then the ideal Line for π has equation $z = 0$, and the ideal Points are Points of the form $[a, b, 0]$, where a and b are not both zero. Any other Point $[a, b, c]$ has $c \neq 0$ and can therefore be represented in π by the Euclidean point $(-a/c, -b/c, -1)$.

Problem 12. Let π be the embedding plane $y = -1$. Describe the ideal Points for π, and specify the Euclidean point of π which represents the Point $[2, 4, 6]$.

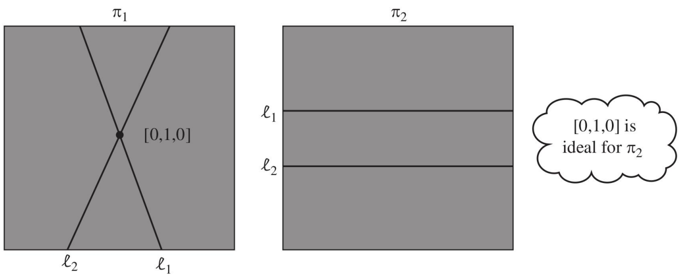
Solution: The ideal Points for π consist of lines through the origin of $\mathbb{R}^3$ that are parallel to π. These are the Points that lie on the ideal Line $y = 0$.

The Euclidean point of π which corresponds to the Point $[2, 4, 6]$ is that multiple of $(2, 4, 6)$ which lies on the plane $y = -1$. That is, $-\frac{1}{4}(2, 4, 6) = \left(-\frac{1}{2}, -1, -\frac{3}{2}\right)$ $\square$

Although we can choose any embedding plane to represent figures of $\mathbb{RP}^2$, the representation does depend on the choice. For example, suppose that π_1 is the embedding plane $y = -1$, and that π_2 is the embedding plane $z = -1$. Now consider the projective figure which consists of two Lines ℓ_1 and ℓ_2 with equations $x = -z$ and $x = z$, respectively. These Lines intersect at the Point $[0, 1, 0]$



On the embedding plane π_1 the Lines ℓ_1 and ℓ_2 are represented by two lines that can be seen to meet at the point corresponding to $[0, 1, 0]$. However, on the embedding plane π_2 the Point of intersection $[0, 1, 0]$ is an ideal Point and so the Lines ℓ_1 and ℓ_2 are represented by parallel lines that do not appear to meet. The contrast between the two representations of ℓ_1 and ℓ_2 is particularly striking if we extract the two embedding planes from $\mathbb{R}^3$ and lay them side by side, as follows.



This example illustrates that Lines which appear to be parallel in one embedding plane may not appear to be parallel in another embedding plane. In the next section you will see that the transformations of projective geometry are chosen so as to ensure that the projective properties of a figure are unaffected by the choice of embedding plane. Since parallelism does depend on the choice of cmbedding planc, it cannot be a projcctive propcrty, so the concept of parallel Lines is meaningless in projective geometry.

3.2.4 An equivalent definition of Projective Geometry

In our work on projective geometry, we have used Euclidean points in a plane in $\mathbb{R}^3$ to construct the projective points (Points) of the geometry $\mathbb{RP}^2$, homogeneous coordinates for those Points, and projective lines (Lines).

Equivalently, we could have defined $\mathbb{RP}^2$ as the set of ordered triples $[a, b, c]$, where a, b, c are real and not all zero, with the convention that we regard $[\ \lambda a, \lambda b, \lambda c]$ and $[a, b, c]$ (where $\lambda \neq 0$) as the same Point in the geometry. We would then have defined projective lines (Lines) as the set of points $[x, y, z]$ in $\mathbb{RP}^2$ that satisfy an equation of the form $ax + by + cz = 0$, where a, b, c are real and not all zero, Then we would continue to develop the theory of projective geometry in the same way as we have done here.

However, we chose to start our work by looking at a model of $\mathbb{RP}^2$ obtained by using an embedding plane π in $\mathbb{R}^3$ that does not pass through the origin. We modeled the projective points $[a, b, c]$ by the Euclidean lines through the origin and the corresponding Euclidean points (a, b, c), plus 'points at infinity' (the ideal Points); and we modeled the projective lines by Euclidean planes through the origin, For convenience, we chose often to use Euclidean points (a, b, c) on a given embedding plane to describe the Euclidean model.

The formal method of defining projective geometry, though, is less intuitive than the description motivated by the $\mathbb{R}^3$ model!

3.3 Projective Transformations

3.3.1 The Group of Projective Transformations

By now you should be familiar with the idea that a geometry consists of a space of points together with a group of transformations which act on that space.

Having introduced the space of projective points $\mathbb{RP}^2$ in Section 3.2, we are now in a position to describe the transformations of $\mathbb{RP}^2$. First we shall define the transformations algebraically, then we give a geometrical interpretation of the transformations using the ideas of perspectivity introduced in Section 3.1, and finally meet the Fundamental Theorem of Projective Geometry.

Recall that a point of $\mathbb{R}^3$ (other than the origin) on an embedding plane π (that does not pass through the origin) has coordinates $\mathbf{x} = (x, y, z)$ with respect to the standard basis of $\mathbb{R}^3$, and homogeneous coordinates of the corresponding Point $[\mathbf{x}]$ in $\mathbb{RP}^2$ (which represents the points $[\lambda\mathbf{x} : \lambda \subset \mathbb{R}]$) are $[\lambda x, \lambda y, \lambda z]$ for some real $\lambda \neq 0$. Since the Points of $\mathbb{RP}^2$ are just lines through the origin of $\mathbb{R}^3$, we need a group of transformations that map the lines through the origin of $\mathbb{R}^3$ onto the lines through the origin of $\mathbb{R}^3$. Suitable transformations of $\mathbb{R}^3$ that do this are the invertible linear transformations.

If $\mathbf{A}$ is the matrix of an invertible linear transformation of $\mathbb{R}^3$ to itself, the transformation maps points $\mathbf{x} = (x, y, z)$ of $\mathbb{R}^3$ to points $\mathbf{Ax}$ of $\mathbb{R}^3$; then the projective transformation with matrix A maps Points $[\ \mathbf{x}]$ of $\mathbb{RP}^2$ to Points $[\mathbf{Ax}]$ of $\mathbb{RP}^2$. This suggests that we define the transformations of projective geometry as follows.

Definition. A projective transformation of $\mathbb{RP}^2$ is a function $t : \mathbb{RP}^2 \to \mathbb{RP}^2$ of the form

$$t : [\mathbf{x}] \mapsto [\mathbf{Ax}]$$

where $\mathbf{A}$ is an invertible 3×3 matrix. We say that $\mathbf{A}$ is a matrix associated with t. The set of all projective transformations of $\mathbb{RP}^2$ is denoted by $P(2)$.

Example 1. Show that the function $t : \mathbb{RP}^2 \to \mathbb{RP}^2$ defined by

$$t : [x, y, z] \mapsto [2x + z, -x + 2y - 3z, x - y + 5z]$$

is a projective transformation, and find the image of $[1, 2, 3]$ under t.

Solution: The transformation t has the form $t : [\mathbf{x}] \mapsto [\mathbf{Ax}]$, where $\mathbf{x} = (x, y, z)$ and

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 2 & -3 \\ 1 & -1 & 5 \end{pmatrix}$$

Now

$$\begin{aligned} det\ \mathbf{A} &= \begin{vmatrix} 2 & 0 & 1 \\ -1 & 2 & -3 \\ 1 & -1 & 5 \end{vmatrix} \\ &= 2(10-3) - 0 + (1-2) \\ &= 13 \neq 0 \end{aligned}$$

So A is invertible. It follows that t is a projective transformation. We have

$$t([1,2,3]) = [2+3, -1+4-9, 1-2+15] = [5, -6, 14]\ .$$

□

Problem 1. Decide which of the following functions t from $\mathbb{RP}^2$ to itself are projective transformations. For those that are projective transformations, write down a matrix associated with t.

(a) $t : [x, y, z] \mapsto [-2y + 3z, -x + 5y - z, -3x]$

(b) $t : [x, y, z] \mapsto [x - 7y + 4z, -x + 5y - z, x - 9y + 7z]$

(c) $t : [x, y, z] \mapsto [x - 1 + z, 2y - 4z + 5, 2x]$

Solution:

(a) The mapping $t : [x, y, z] \mapsto [-2y + 3z, -x + 5y - z, -3x]$ can be expressed in the form $[\mathbf{x}] \mapsto [\mathbf{A}\mathbf{x}]$, where

$$\mathbf{A} = \begin{pmatrix} 0 & -2 & 3 \\ -1 & 5 & -1 \\ -3 & 0 & 0 \end{pmatrix}$$

Now

$$\begin{aligned} \det \mathbf{A} &= \begin{vmatrix} 0 & -2 & 3 \\ -1 & 5 & -1 \\ -3 & 0 & 0 \end{vmatrix} \\ &= 0 \begin{vmatrix} 5 & -1 \\ 0 & 0 \end{vmatrix} - (-2) \begin{vmatrix} -1 & -1 \\ -3 & 0 \end{vmatrix} + 3 \begin{vmatrix} -1 & 5 \\ -3 & 0 \end{vmatrix} \\ &= 0 + 2 \times (-3) + 3 \times 15 \\ &= 39 \neq 0, \end{aligned}$$

so A is invertible. It follows that t is projective transformation, and that $\mathbf{A}$ is a matrix associated with t.

(b) The mapping

$$t : [x, y, z] \mapsto [x - 7y + 4z, -x + 5y - z, x - 9y + 7z]$$

can be expressed in the form $[\mathbf{x}] \mapsto [\mathbf{Ax}]$ where

$$\mathbf{A} = \begin{pmatrix} 1 & -7 & 4 \\ -1 & 5 & -1 \\ 1 & -9 & 7 \end{pmatrix}$$

Now

$$\begin{aligned} \det \mathbf{A} &= \begin{vmatrix} 1 & -7 & 4 \\ -1 & 5 & -1 \\ 1 & -9 & 7 \end{vmatrix} \\ &= 1 \begin{vmatrix} 5 & -1 \\ -9 & 7 \end{vmatrix} + 7 \begin{vmatrix} -1 & -1 \\ 1 & 7 \end{vmatrix} + 4 \begin{vmatrix} -1 & 5 \\ 1 & -9 \end{vmatrix} \\ &= 1 \times 26 + 7 \times (-6) + 4 \times 4 \\ &= 0, \end{aligned}$$

so A is not invertible. It follows that t is not a projective transformation.

(c) The mapping

$$t : [x, y, z] \mapsto [x - 1 + z, 2y - 4z + 5, 2x]$$

cannot be expressed in the form $[\mathbf{x}] \mapsto [\mathbf{Ax}]$, where $\mathbf{A}$ is a 3×3 matrix whose entries are real numbers. Hence t cannot be a projective transformation.

$\square$

Problem 2. Let t be the projcctive transformation associated with the matrix.

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & -1 \\ -1 & -2 & 1 \\ 4 & -3 & 4 \end{pmatrix}$$

Determine the image under t of each of the following Points.

(a) $[1, 2, -1]$ (b) $[1, 0, 0]$ (c) $[0, 1, 0]$

(d) $[0, 0, 1]$ (e) $[1, 1, 1]$

Solution:

(a) The image of the Point $[1, 2, -1]$ under t is given by

$$\left[\begin{pmatrix} 1 & 1 & -1 \\ -1 & -2 & 1 \\ 4 & -3 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}\right] = \left[\begin{pmatrix} 4 \\ -6 \\ -6 \end{pmatrix}\right]$$

that is, the Point $[4, -6, -6] = [-2, 3, 3]$

(b) The image of the Point $[1, 0, 0]$ under t is given by

$$\left[\begin{pmatrix} 1 & 1 & -1 \\ -1 & -2 & 1 \\ 4 & -3 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right] = \left[\begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix}\right]$$

that is, the Point $[1, -1, 4]$.

(c) The image of the Point $[0, 1, 0]$ under t is given by

$$\left[\begin{pmatrix} 1 & 1 & -1 \\ -1 & -2 & 1 \\ 4 & -3 & 4 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right] = \left[\begin{pmatrix} 1 \\ -2 \\ -3 \end{pmatrix}\right]$$

that is, the Point $[1, -2, -3]$.

(d) The image of the Point $[0, 0, 1]$ under t is given by

$$\left[\begin{pmatrix} 1 & 1 & -1 \\ -1 & -2 & 1 \\ 4 & -3 & 4 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right] = \left[\begin{pmatrix} -1 \\ 1 \\ 4 \end{pmatrix}\right]$$

that is, the Point $[-1, 1, 4]$.

(e) The image of the Point $[1, 1, 1]$ under t is given by

$$\left[\begin{pmatrix} 1 & 1 & -1 \\ -1 & -2 & 1 \\ 4 & -3 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right] = \left[\begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix}\right]$$

that is the Point $[1, -2, 5]$.

$\square$

Since we can multiply the homogeneous coordinates of Points in $\mathbb{RP}^2$ by any non-zero real number λ without altering the Point itself, it follows that if $\mathbf{A}$ is a matrix associated with a particular projective transformation then so is the matrix $\lambda\mathbf{A}$, provided that $\lambda \neq 0$. For example, another matrix associated with the transformation in Example 1 is

$$\mathbf{B} = \begin{pmatrix} -4 & 0 & -2 \\ 2 & -4 & 6 \\ -2 & 2 & -10 \end{pmatrix}$$

for we have $\mathbf{B} = -2\mathbf{A}$.

Problem 3. Write down a matrix with top left-hand entry $\frac{1}{2}$ which is associated with the transformation in Example 1 .

Solution: Since the matrix $\mathbf{A}$ which represents the transformation in Example 1 in Subsection 3.3.1 has 2 as its top left-hand entry, we obtain the required matrix by dividing each entry of $\mathbf{A}$ by 4. This gives the matrix

$$\begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{4} \\ -\frac{1}{4} & \frac{1}{2} & -\frac{3}{4} \\ \frac{1}{4} & -\frac{1}{4} & \frac{5}{4} \end{pmatrix}$$

□

Before we can use the projective transformations to define projective geometry, we must first check that they form a group.

Theorem 1. The set of projective transformations $P(2)$

forms a group under the operation of composition of functions.

Proof: We check that the four group axioms hold.

G1 CLOSURE: Let t_1 and t_2 be projective transformations defined by

$$t_1 : [\mathbf{x}] \mapsto [\mathbf{A}_1\mathbf{x}] \quad \text{and} \quad t_2 : [\mathbf{x}] \mapsto [\mathbf{A}_2\mathbf{x}]$$

where $\mathbf{A}_1$ and $\mathbf{A}_2$ are invertible 3×3 matrices. Then

$$\begin{aligned} t_1 \circ t_2([\mathbf{x}]) &= t_1\left(t_2([\mathbf{x}])\right) \\ &= t_1\left([\mathbf{A}_2\mathbf{x}]\right) \\ &= \left[(\mathbf{A_1}\mathbf{A_2})\,\mathbf{x}\right]. \end{aligned}$$

Since $\mathbf{A}_1$ and $\mathbf{A}_2$ are invertible, it follows that $\mathbf{A}_1\mathbf{A_2}$ is invertible. So by definition $t_1 \circ t_2$ is a projective transformation.

G2 IDENTITY: Let $i : \mathbb{RP}^2 \to \mathbb{RP}^2$ be the transformation defined by $i : [\mathbf{x}] \mapsto [\mathbf{Ix}]$ where $\mathbf{I}$ is the 3×3 identity matrix; this is a projective transformation, since $\mathbf{I}$ is invertible.

Let $t : \mathbb{RP}^2 \to \mathbb{RP}^2$ be an arbitrary projective transformation, defined by $t : [\mathbf{x}] \mapsto [\mathbf{Ax}]$, for some invertible 3 $\times 3$ matrix $\mathbf{A}$. Then for any $[\mathbf{x}] \in \mathbb{RP}^2$

$$t \circ i([\mathbf{x}]) = [\mathbf{A}(\mathbf{Ix})] = [\mathbf{Ax}]$$

and

$$i \circ t([\mathbf{x}]) = [\mathbf{I}(\mathbf{A}\mathbf{x})] = [\mathbf{A}\mathbf{x}]$$

Thus $t \circ i = i \circ t = t$. Hence i is the identity transformation.

G3 INVERSES: Let $t : \mathbb{RP}^2 \to \mathbb{RP}^2$ be an arbitrary projective transformation defined by

$$t : [\mathbf{x}] \mapsto [\mathbf{A}\mathbf{x}]$$

for some invertible 3×3 matrix $\mathbf{A}$. Then we can define another projective transformation $t' : \mathbb{RP}^2 \to \mathbb{RP}^2$ by

$$t' : [\mathbf{x}] \mapsto \left[\mathbf{A}^{-1}\mathbf{x}\right]$$

Now, for each $[\mathbf{x}] \in \mathbb{RP}^2$, we have

$$t \circ t'([\mathbf{x}]) = \mathbf{t}\left(\left[\mathbf{A}^{-1}\mathbf{x}\right]\right) = \left[\mathbf{A}\left(\mathbf{A}^{-1}\mathbf{x}\right)\right] = [\mathbf{x}]$$

and

$$t' \circ t([\mathbf{x}]) = t'([\mathbf{A}\mathbf{x}]) = \left[\mathbf{A}^{-1}(\mathbf{A}\mathbf{x})\right] = [\mathbf{x}]$$

Thus t' is an inverse for t.

G4 ASSOCIATIVITY: Composition of functions is always associative.

It follows that the set of projective transformations $P(2)$ forms a group. □

The above proof shows that if t_1 and t_2 are projective transformations with associated matrices $\mathbf{A}_1$ and $\mathbf{A}_2$, respectively, then $t_1 \circ t_2$ is a projective transformation with an associated matrix $\mathbf{A}_1\mathbf{A}_2$. We therefore have the following strategy for composing projective transformations.

Strategy. To compose two projective transformations t_1 and t_2 :

1. write down matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ associated with t_1 and t_2.

2. calculate $\mathbf{A}_1\mathbf{A}_2$.

3. write down the composite $t_1 \circ t_2$ with which $\mathbf{A}_1\mathbf{A}_2$ is associated.

The proof also shows that if t is a projective transformation with an associated matrix $\mathbf{A}$, then t^{-1} is a projective transformation with associated matrix $\mathbf{A}^{-1}$. We therefore have the following strategy for calculating the inverse of a projective transformation.

Strategy. To find the inverse of a projective transformation t :

1. write down a matrix $\mathbf{A}$ associated with t.

2. calculate $\mathbf{A}^{-1}$.

3. write down the inverse t^{-1} with which $\mathbf{A}^{-1}$ is associated.

Example 2. Let t_1 and t_2 be projective transformations defined by

$$t_1 : [x, y, z] \mapsto [x + z, x + y + 3z, -2x + z]$$
$$t_2 : [x, y, z] \mapsto [2x, x + y + z, 4x + 2y]$$

Determine the projective transformations $t_2 \circ t_1$ and t_1^{-1}.

Solution: The transformations t_1 and t_2 have associated matrices

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 3 \\ -2 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{A}_2 = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 1 \\ 4 & 2 & 0 \end{pmatrix}$$

respectively. It follows that $t_2 \circ t_1$ has an associated matrix

$$\mathbf{A}_2\mathbf{A}_1 = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 1 \\ 4 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 3 \\ -2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 2 \\ 0 & 1 & 5 \\ 6 & 2 & 10 \end{pmatrix}$$

so

$$t_2 \circ t_1 : [x, y, z] \mapsto [2x + 2z, y + 5z, 6x + 2y + 10z]$$

Next, t_1^{-1} has an associated matrix $\mathbf{A}_1^{-1}$ given by

$$\mathbf{A}_1^{-1} = \begin{pmatrix} \frac{1}{3} & 0 & -\frac{1}{3} \\ -\frac{7}{3} & 1 & -\frac{2}{3} \\ \frac{2}{3} & 0 & \frac{1}{3} \end{pmatrix}$$

a simpler matrix associated with t_1^{-1} is then

$$\begin{pmatrix} 1 & 0 & -1 \\ -7 & 3 & -2 \\ 2 & 0 & 1 \end{pmatrix}$$

so

$$t_1^{-1} : [x, y, z] \mapsto [x - z, -7x + 3y - 2z, 2x + z]$$

□

Problem 4. Let t_1 and t_2 be projective transformations defined by

$$t_1 : [x, y, z] \mapsto [2x + y, -x + z, y + z]$$
$$t_2 : [x, y, z] \mapsto [5x + 8y, 3x + 5y, 2z]$$

Determine the projective transformations $t_1 \circ t_2$ and t_1^{-1}.

Solution: Matrices associated with t_1 and t_2 are

$$\mathbf{A}_1 = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \text{ and } \mathbf{A}_2 = \begin{pmatrix} 5 & 8 & 0 \\ 3 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

respectively. It follows that a matrix associated with the projective transformation $t_1 \circ t_2$ is $\mathbf{A}_1\mathbf{A}_2$. Now

$$\mathbf{A}_1\mathbf{A}_2 = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 5 & 8 & 0 \\ 3 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$
$$= \begin{pmatrix} 13 & 21 & 0 \\ -5 & -8 & 2 \\ 3 & 5 & 2 \end{pmatrix}$$

so we conclude that $t_1 \circ t_2$ is the transformation

$$[x, y, z] \mapsto [13x + 21y, -5x - 8y + 2z, 3x + 5y + 2z]$$

Next, t_1^{-1} has an associated matrix $\mathbf{A}_1^{-1}$ given by

$$\mathbf{A}_1^{-1} = \begin{pmatrix} 1 & 1 & -1 \\ -1 & -2 & 2 \\ 1 & 2 & -1 \end{pmatrix}$$

The projective transformation t_1^{-1} is therefore

$$[x, y, z] \longmapsto [x + y - z, -x - 2y + 2z, x + 2y - z]$$

$\square$

Having shown that the set of projective transformations forms a group under composition of functions, we can now define projective geometry to be the study of those properties of figures in $\mathbb{RP}^2$ that are preserved by projective transformations.

Those properties that are preserved by projective transformations are known as projective properties.

3.3.2 Some Properties of Projective Transformations

We now check two important properties of projective transformations, namely, that they preserve collinearity and incidence.

A Line in $\mathbb{RP}^2$ is a plane in $\mathbb{R}^3$ that passes through the origin. It therefore consists of the set of points (x, y, z) of $\mathbb{R}^3$ that satisfy an equation of the form

$$ax + by + cz = 0$$

wherc a, b and c arc not all zcro. Wc can writc this condition cquivalently in the matrix form $\mathbf{Lx} = 0$, where $\mathbf{L}$ is the non-zero row matrix (abc) and $\mathbf{x} = (xyz)^T$.

Now let t be a projective transformation defined by $t : [\mathbf{x}] \mapsto [\mathbf{Ax}]$, where $\mathbf{A}$ is an invertible 3×3 matrix, and let $[\mathbf{x}]$ be an arbitrary Point on the Line $\mathbf{Lx} = 0$. Then the image of $[\mathbf{x}]$ under t is a Point $[\mathbf{x}']$ where $\mathbf{x}' = \mathbf{Ax}$. Since $\mathbf{x}$ satisfies the equation $\mathbf{Ix} = 0$, it follows that $\mathbf{x}'$ satisfies $\mathbf{I}\left(\mathbf{A}^{-1}\mathbf{x}'\right) = 0$, or $\left(\mathrm{LA}^{-1}\right)\mathbf{x}' = 0$. Dropping the dash, we conclude that the image of the Line $\mathbf{Lx} = 0$ under t is the Line with equation

$$\left(\mathbf{LA}^{-1}\right)\mathbf{x} = 0$$

Since the image of a Line in $\mathbb{RP}^2$ is a Line, it follows that collinearity is preserved under a projective transformation.

Notice that if $\mathbf{B}$ is any matrix associated with t^{-1}, then $\mathbf{B} = \lambda\mathbf{A}^{-1}$ for some non-zero real number λ, and so $\left(\mathbf{LA}^{-1}\right)\mathbf{x} = 0$ if and only if $(\mathbf{LB})\mathbf{x} = 0$. It follows that the image of the Line can equally well be written as $(\mathbf{LB})\mathbf{x} = 0$. (For instance, since $\mathbf{A}^{-1} = \operatorname{adj}(\mathbf{A})/\det(\mathbf{A})$ so that t^{-1} also has $\operatorname{adj}(\mathbf{A})$ as an associated matrix, we can express the image of the Line as $(\mathbf{L}\operatorname{adj}(\mathbf{A}))\mathbf{x} = 0$.)

We therefore summarize the above discussion in the form of a strategy, as follows.

Strategy. To find the image of a Line

$$ax + by + cz = 0$$

under a projective transformation $t : [\mathbf{x}] \mapsto [\mathbf{AX}]$:

1. write the equation of the Line in the form $\mathbf{Lx} = 0$, where $\mathbf{L}$ is the matrix (abc);
2. find a matrix $\mathbf{B}$ associated with t^{-1};
3. write down the equation of the image as $(\mathbf{LB})\mathbf{x} = 0$.

Example 3. Find the image of the Line $2x+y-3z=0$ under the projective transformation t_1 defined by

$$t_1 : [x, y, z] \mapsto [x+z, x+y+3z, -2x+z]$$

Solution: The equation of the Line can be written in the form $\mathbf{Lx} = 0$, where

$$\mathbf{L} = \begin{pmatrix} 2 & 1 & -3 \end{pmatrix}$$

In Example 2 we showed that t_1^{-1} has an associated matrix

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & -1 \\ -7 & 3 & -2 \\ 2 & 0 & 1 \end{pmatrix}$$

So

$$\mathbf{LB} = \left(\quad 2 \quad 1 - 3) \right. \begin{pmatrix} 1 & 0 & -1 \\ -7 & 3 & -2 \\ 2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -11 & 3-7 \end{pmatrix}$$

It follows that the required image has equation

$$-11x + 3y - 7z = 0$$

$\square$

Problem 5. Find the image of the Line $x + 2y - z = 0$ under the projective transformation t_1 defined by

$$t_1 : [x, y, z] \mapsto [2x + y, -x + z, y + z]$$

Solution: The equation of the Line can be written in the form $\mathbf{Lx} = 0$, where

$$\mathbf{L} = \begin{pmatrix} 1 & 2 & -1 \end{pmatrix}$$

From Problem 4 we know that t_1^{-1} has an associated matrix

$$\mathbf{A}_1^{-1} = \begin{pmatrix} 1 & 1 & -1 \\ -1 & -2 & 2 \\ 1 & 2 & -1 \end{pmatrix}$$

so

$$\begin{aligned} \mathbf{LA}_1^{-1} &= \begin{pmatrix} 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ -1 & -2 & 2 \\ 1 & 2 & -1 \end{pmatrix} \\ &= \begin{pmatrix} -2 & -5 & 4 \end{pmatrix}. \end{aligned}$$

The required image is therefore the Line

$$-2x - 5y + 4z = 0$$

$\square$

Next, we consider the incidence property. If two Lines intersect at the Point P, then P lies on both Lines. So if t is a projective transformation, then $t(P)$ lies on the images of both

Lines. It follows that the image under t of the Point of intersection of the two Lines is the Point of intersection of the images of the two Lines. In other words, incidence is also preserved under a projective transformation.

Theorem 2. Collinearity and incidence are both projective properties.

3.3.3 Fundamental Theorem of Projective Geometry

In Chapter 2 we discussed the Fundamental Theorem of Affine Geometry which states that given any two sets of three non-collinear points of $\mathbb{R}^2$ there is a unique affine transformation which maps the points in one set to the corresponding points in the other set. So an affine transformation is uniquely determined by its effect on any given triangle.

In this subsection we explore an analogous result for projective geometry known as the Fundamental Theorem of Projective Geometry. We begin by asking you to tackle the following problem.

Problem 6. Let t_1 and t_2 be the projective transformations with associated matrices

$$\mathbf{A}_1 = \begin{pmatrix} -4 & -1 & 1 \\ -3 & -2 & 1 \\ 4 & 2 & -1 \end{pmatrix} \text{ and } \mathbf{A}_2 = \begin{pmatrix} -8 & -6 & -2 \\ -3 & 4 & 7 \\ 6 & 0 & -4 \end{pmatrix}$$

respectively. Find the images of the Points $[1,-1,1],[1,-2,2]$ and $[-1,2,-1]$ under t_1 and t_2.

Solution: First we consider the images under t_1. The image of the Point $[1,-1,1]$ under t_1 is

$$\left[\begin{pmatrix} -4 & -1 & 1 \\ -3 & -2 & 1 \\ 4 & 2 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}\right] = \left[\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}\right]$$

that is, the Point $[-2,0,1]$. Similarly, the image of the Point $[1,-2,2]$ under t_1 is

$$\left[\begin{pmatrix} -4 & -1 & 1 \\ -3 & -2 & 1 \\ 4 & 2 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}\right] = \left[\begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix}\right]$$

that is, the Point $[0,3,-2]$. Finally, the image of the Point $[-1,2,-1]$ under t_1 is

$$\left[\begin{pmatrix} -4 & -1 & 1 \\ -3 & -2 & 1 \\ 4 & 2 & -1 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}\right] = \left[\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}\right]$$

that is, the Point $[1,-2,1]$. Next, we consider the images under t_2. The image of the Point $[1,-1,1]$ under t_2 is

$$\left[\begin{pmatrix} -8 & -6 & -2 \\ -3 & 4 & 7 \\ 6 & 0 & -4 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}\right] = \left[\begin{pmatrix} -4 \\ 0 \\ 2 \end{pmatrix}\right]$$

that is, the Point with homogeneous coordinates $[-4,0,2]$ or (equivalently) $[-2,0,1]$.

Similarly, the image of the Point $[1,-2,2]$ under t_2 is

$$\left[\begin{pmatrix} -8 & -6 & -2 \\ -3 & 4 & 7 \\ 6 & 0 & -4 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}\right] = \left[\begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix}\right]$$

that is, the Point $[0,3,-2]$. Finally, the image of the Point $[-1,2,-1]$ under t_2 is

$$\left[\begin{pmatrix} -8 & -6 & -2 \\ -3 & 4 & 7 \\ 6 & 0 & -4 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}\right] = \left[\begin{pmatrix} -2 \\ 4 \\ -2 \end{pmatrix}\right]$$

that is, the Point with homogeneous coordinates $[-2,4,-2]$ or (equivalently) $[1,-2,1]$ □

You should have found that both of the projective transformations t_1 and t_2 map the Points $[1,-1,1], [1,-2,2]$ and $[-1,2,-1]$ to the Points $[-2,0,1]$, $[0,3,-2]$ and $[1,-2,1]$, respectively. Notice, however, that t_1 and t_2 are not the same projective transformation, since their matrices are not multiples of each other. It follows that, unlike affine transformations, projective transformations are not uniquely determined by their effect on three (non-collinear) Points.

This raises the question as to whether it is possible to specify how many Points are required to determine a projective transformation. According to the Fundamental Theorem of Projective Geometry, the answer is four. In fact the theorem states that given any two sets of four Points, no three of which are collinear, there is a unique projective transformation that maps the Points in one set to the corresponding Points in the second set. Thus, in projective geometry a transformation is uniquely determined by its effect on a quadrilateral.

To understand why a triangle is insufficient to determine a projective transformation uniquely, consider what happens when we look for a projective transformation that maps the triangle of reference to three given non-collinear Points.

Example 4. Find a projective transformation t that maps the Points $[1, 0, 0]$, $[0, 1, 0]$ and $[0, 0, 1]$ to the non collinear Points $[1, -1, 1], [1, -2, 2]$ and $[-1, 2, -1]$, respectively.

Solution: Let A be a matrix associated with t, and let the first column of $\mathbf{A}$ be $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$. Then since

$$\left[\begin{pmatrix} a & * & * \\ b & * & * \\ c & * & * \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right] = \left[\begin{pmatrix} a \\ b \\ c \end{pmatrix}\right] = \left[\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}\right]$$

it follows that we may take $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ as the first column of $\mathbf{A}$.

Similarly, since

$$\left[\begin{pmatrix} * & d & * \\ * & e & * \\ * & f & * \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right] = \left[\begin{pmatrix} d \\ e \\ f \end{pmatrix}\right] = \left[\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}\right]$$

Projective Transformations and

$$\left[\begin{pmatrix} * & * & g \\ * & * & h \\ * & * & k \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right] = \left[\begin{pmatrix} g \\ h \\ k \end{pmatrix}\right] = \left[\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}\right]$$

it follows that a suitahle transformation is given hy $t : [\mathbf{x}] \mapsto$

$[\mathbf{A}x]$ where

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & -1 \\ -1 & -2 & 2 \\ 1 & 2 & -1 \end{pmatrix}.$$

$\square$

This example illustrates the fact that we can always find a projective transformation $t : [\mathbf{x}] \mapsto [\mathbf{Ax}]$ which maps the triangle of reference to three non-collinear Points simply by writing the homogeneous coordinates of the Points as the columns of $\mathbf{A}$. Notice, however, that the transformation we obtain is not unique. Indeed, if the Points $[1, -1, 1], [1, -2, 2]$ and $[-1, 2, -1]$ in Example 4 are rewritten in the form $[u, -u, u], [v, -2v, 2v]$ and $[-w, 2w, -w]$, for some non-zero real numbers u, v, w, then the matrix becumes

$$\mathbf{A} = \begin{pmatrix} u & v & -w \\ -u & -2v & 2w \\ u & 2v & -w \end{pmatrix}$$

The corresponding transformation $t : [\mathbf{x}] \mapsto [\mathbf{Ax}]$ still maps the triangle of reference to the Points $[1, -1, 1], [1, -2, 2]$ and $[-1, 2, -1]$, as required, but the effect that t has on the other Points of $\mathbb{RP}^2$ depends on the numbers u, v and w.

So if we wish to specify t uniquely we need to assign particular values to u, v and w. We can do this by specifying the effect that t has on a fourth Point $[1, 1, 1]$.

Example 5. Find the projective transformation t which maps the Points $[1,0,0]$, $[0,1,0]$, $[0,0,1]$ and $[1,1,1]$ to the Points $[1,-1,1],[1,-2,2]$, $[-1,2,-1]$ and $[0,1,2]$, respectively.

Solution: If $\mathbf{A}$ is the matrix associated with t, then its columns must be multiples of the homogeneous coordinates $[1,-1,1],[1,-2,2],[-1,2,-1]$; that is,

$$\mathbf{A} = \begin{pmatrix} u & v & -w \\ -u & -2v & 2w \\ u & 2v & -w \end{pmatrix}$$

Also, to ensure that t maps $[1,1,1]$ to $[0,1,2]$ we must choose u, v and w so that

$$\left[\begin{pmatrix} u & v & -w \\ -u & -2v & 2w \\ u & 2v & -w \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right] = \left[\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}\right]$$

We can do this by solving the equations

$$\begin{aligned} u + v - w &= 0, \\ -u - 2v + 2w &= 1, \\ u + 2v - w &= 2. \end{aligned}$$

Adding the second and third equations we obtain $w = 3$. If we then subtract the first equation from the third we obtain $v = 2$. Finally, if we substitute v and w into the first equation we obtain $u = 1$. The required projective transformation is therefore given by $t : [\mathbf{x}] \mapsto [\mathbf{Ax}]$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & -3 \\ -1 & -4 & 6 \\ 1 & 4 & -3 \end{pmatrix}$$

$\square$

It is natural to ask whether the method used in this example can be adapted to find a projcetive transformation which maps the triangle of refcrence and unit Point to any four given Points. The answer is usually yes, but since collinearity is a projective property, and since no three of the Points $[1, 0, 0], [0, 1, 0], [0,,$ $0, 1], [1, 1, 1]$ are collinear, the method must fail if three of the four given Points lie on a Line. Provided we exclude this possibility. the answer is yes!

Strategy. To find the projective transformation which maps

$$[1, 0, 0] \text{ to } [a_1, a_2, a_3]$$
$$[0, 1, 0] \text{ to } [b_1, b_2, b_3]$$
$$[0, 0, 1] \text{ to } [c_1, c_2, c_3]$$
$$[1, 1, 1] \text{ to } [d_1, d_2, d_3]$$

where

no three of $[a_1, a_2, a_3], [b_1, b_2, b_3], [c_1, c_2, c_3], [d_1, d_2, d_3]$ are collinear:

1. find u, v, w such that

$$\begin{pmatrix} a_1 u & b_1 v & c_1 w \\ a_2 u & b_2 v & c_2 w \\ a_3 u & b_3 v & c_3 w \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$$

2. write down the required projective transformation in the form $t : [\mathbf{x}] \mapsto [\mathbf{Ax}]$, where $\mathbf{A}$ is any non-zero real multiple of the matrix

$$\begin{pmatrix} a_1 u & b_1 v & c_1 w \\ a_2 u & b_2 v & c_2 w \\ a_3 u & b_3 v & c_3 w \end{pmatrix}$$

Remark

To see why this strategy always works, notice that we can rewrite the equation from Step 1 in the form

$$u \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + v \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} + w \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$$

From this we can make the following observations.

(a) The equation in Step 1 must have a unique solution for u, v, w because the required values of u, v and w are simply the coordinates of (d_1, d_2, d_3) with respect to the basis of $\mathbb{R}^3$ formed from the three linearly independent vectors $(a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3)$

(b) The values of u, v and w must all be non-zero, because otherwise three of the vectors $(a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3), (d_1, d_2, d_3)$ would be linearly dependent.

(c) Since the columns of $\mathbf{A}$ are non-zero, multiples of the linearly independent vectors $(a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3)$ it follows that $\mathbf{A}$ is invertible, and hence that t is a projective transformation.

There is no need to check whether any three of the four given Points are collincar becausc any failure of this condition will cmerge in the process of applying the strategy. Indeed, if the equation in Step 1 fails to yield unique non-zero values for u, v and w, then it must be because three of the Points $(a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3), (d_1, d_2, d_3)$ lie on a Line.

Problem 7. Use the above strategy to find the projective transformation which maps the Points $[1, 0, 0], [0, 1, 0], [0, 0, 1]$ and $[1, 1, 1]$ to the Points:

(a) $[-1, 0, 0], [-3, 2, 0], [2, 0, 4]$ and $[1, 2, -5]$, respectively

(b) $[1, 0, 0], [0, 0, 1], [0, 1, 0]$ and $[3, 4, 5]$, respectively;

(c) $[2, 1, 0], [1, 0, -1], [0, 3, -1]$ and $[3, -1, 2]$, respectively.

Solution: We use the strategy preceding the problem.

(a) Let $\mathbf{A}$ be the matrix

$$\begin{pmatrix} -u & -3v & 2w \\ 0 & 2v & 0 \\ 0 & 0 & 4w \end{pmatrix}.$$

We wish to choose u, v, w such that

$$\begin{pmatrix} -u & -3v & 2w \\ 0 & 2v & 0 \\ 0 & 0 & 4w \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -5 \end{pmatrix},$$

that is,

$$\begin{pmatrix} -u - 3v + 2w \\ 2v \\ 4w \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -5 \end{pmatrix}$$

It follows that $w = -\frac{5}{4}$ and $v = 1$. Also,

$$-u - 3v + 2w = 1, \text{ so } u = -\frac{13}{2}$$

Thus

$$\mathbf{A} = \begin{pmatrix} \frac{13}{2} & -3 & -\frac{5}{2} \\ 0 & 2 & 0 \\ 0 & 0 & -5 \end{pmatrix}.$$

A simpler matrix for the projective transformation is the matrix

$$2\mathbf{A} = \begin{pmatrix} 13 & -6 & -5 \\ 0 & 4 & 0 \\ 0 & 0 & -10 \end{pmatrix}.$$

(b) Let $\mathbf{A}$ be the matrix

$$\begin{pmatrix} u & 0 & 0 \\ 0 & 0 & w \\ 0 & v & 0 \end{pmatrix}.$$

We wish to choose u, v, w such that

$$\begin{pmatrix} u & 0 & 0 \\ 0 & 0 & w \\ 0 & v & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix},$$

that is,

$$\begin{pmatrix} u \\ w \\ v \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}.$$

It follows that $u = 3, v = 5$ and $w = 4$. Thus

$$\mathbf{A} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 5 & 0 \end{pmatrix}.$$

(c) Let $\mathbf{A}$ be the matrix

$$\begin{pmatrix} 2u & v & 0 \\ u & 0 & 3w \\ 0 & -v & -w \end{pmatrix}$$

We wish to choose u, v, w such that

$$\begin{pmatrix} 2u & v & 0 \\ u & 0 & 3w \\ 0 & -v & -w \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}$$

that is,

$$\begin{pmatrix} 2u + v \\ u + 3w \\ -v - w \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}$$

It follows that

$$2u + v = 3 \qquad (a)$$
$$u + 3w = -1 \qquad (b)$$
$$-v - w = 2 \qquad (c)$$

Adding equations (a) and (c) in order to eliminate v, we obtain

$$2u - w = 5 \qquad (d)$$

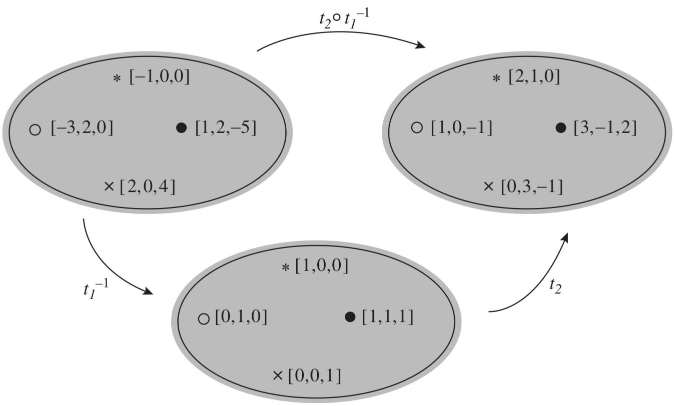
Subtracting equation (d) from twice equation (b) in order to eliminate u, we obtain $7w = -7$ or $w = -1$

It follows from equation (d) that $u = 2$, and from equation (a) that $v = -1$. Thus

$$\mathbf{A} = \begin{pmatrix} 4 & -1 & 0 \\ 2 & 0 & -3 \\ 0 & 1 & 1 \end{pmatrix}$$

$\square$

Now consider the transformation t_1 in Problem 7 (a). The inverse of this, t_1^{-1}, is a projective transformation which maps the Points $[-1, 0, 0], [-3, 2, 0]$ $[2, 0, 4]$ and $[1, 2, -5]$ back to the triangle of reference and unit Point. So if, after applying this inverse, we apply the projective transformation t_2 in Problem 7 (c), then the overall effect of the composite $t_2 \circ t_1^{-1}$ is that of a projective transformation which sends the Points $[-1, 0, 0], [-3, 2, 0], [2, 0, 4]$ and $[1, 2, -5]$ directly to the Points $[2, 1, 0], [1, 0, -1], [0, 3, -1]$ and $[3, -1, 2]$, respectively



In a similar way we can find a projective transformation which maps any set of four Points to any other set of four Points. The only constraint is that no three of the Points in either set can be collinear. In the following statement of the Fundamental Theorem we express this constraint by requiring that each of the four sets of Points lie at the vertices of some quadrilateral, where a quadrilateral is defined as follows. A quadrilateral is a set of four Points A, B, C and D (no three of which are collinear), together with the Lines AB, BC, CD and DA.

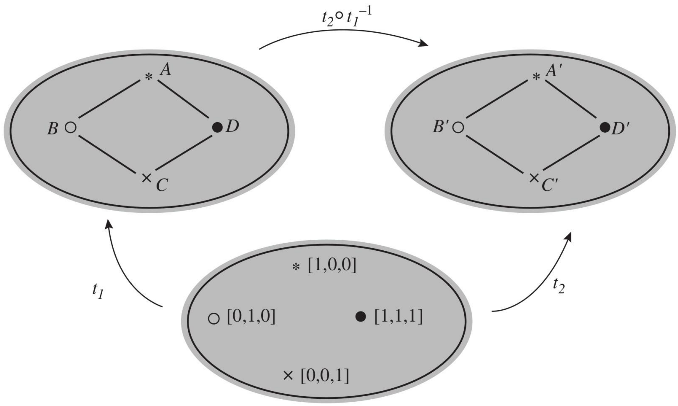
Theorem 3. (The Fundamental Theorem of Projective Geometry) Let $ABCD$ and $A'B'C'D'$ be two quadrilaterals in $\mathbb{RP}^2$. Then:

(a) there is a projective transformation t which maps

$$A \text{ to } A', B \text{ to } B', C \text{ to } C', D \text{ to } D'$$

(b) the projective transformation t is unique.

Proof: According to the strategy above, there is a projective transformation t_1 which maps the Points $[1,0,0], [0,1,0], [0,0,1], [1,1,1]$ to the Points A, B, C, D, respectively. Similarly, there is a projective transformation t_2 which maps the Points $[1,0,0], [0,1,0], [0,0,1], [1,1,1]$ to the Points A', B', C', D', respectively.



(a) The composite $t = t_2 \circ t_1^{-1}$ is then a projective transformation which maps A to A', B to B', C to C', D to D'.

(b) To check uniqueness of t, we first check that the identity transformation is the only projective transformation which maps each of the Points $[1,0,0], [0,1,0], [0,0,1], [1,1,1]$ to themselves. In fact any projective transformation with this property must have an associated matrix which is some non-zero multiple of the matrix

$$\begin{pmatrix} u & 0 & 0 \\ 0 & v & 0 \\ 0 & 0 & w \end{pmatrix}, \quad \text{where} \begin{pmatrix} u & 0 & 0 \\ 0 & v & 0 \\ 0 & 0 & w \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Such a matrix must be (a non-zero multiple of) the identity matrix, and so the transformation must indeed be the identity.

Next suppose that t and t' are two projective transformations which satisfy the conditions of the theorem. Then the composites $t_2^{-1} \circ t \circ t_1$ and $t_2^{-1} \circ$ t$'$ - t_1 must both be projective transformations which map each of the Points $[1,0,0],[0,1,0],[0,0,1],[1,1,1]$ to themselves. Since this implies that both composites are equal to the identity, we deduce that

$$t_2^{-1} \circ t \circ t_1 = t_2^{-1} \circ t' \circ t_1$$

If we now compose both sides of this equation with t_2 on the left and with t_1^{-1} on the right, then we obtain $t = t'$, as required. □

The Fundamental Theorem tells us that there is a projective transformation which maps any given quadrilateral onto any other given quadrilateral. So we have the following corollary.

Corollary. All quadrilaterals are projective-congruent.

If we actually need to find the projective transformation which maps one given quadrilateral onto another given quadrilateral, we simply follow the strategy used to prove part (a) of the Fundamental Theorem.

Strategy. To determine the projective transformation t which maps the vertices of the quadrilateral $ABCD$ to the corresponding vertices of the quadrilateral $A'B'C'D'$:

1. find the projective transformation t_1 which maps the triangle of reference and unit Point to the Points A, B, C, D, respectively;

2. find the projective transformation t_2 which maps the triangle of reference and unit Point to the Points A', B', C', D', respectively;

3. calculate $t = t_2 \circ t_1^{-1}$.

Example 6. Find the projective transformation t which maps the Points $[1, -1, 2]$, $[1, -2, 1]$, $[5, -1, 2]$, $[1, 0, 1]$ to the Points $[-1, 3, -2], [-3, 7, -5], [2, -5, 4], [-3, 8, -5]$, respectively.

Solution: We follow the steps in the above strategy.

(a) Any matrix associated with the projective transformation t_1 which maps the Points $[1, 0, 0], [0, 1, 0], [0, 0, 1], [1, 1, 1]$ to the Points $[1, -1, 2]$, $[1, -2, 1], [5, -1, 2], [1, 0, 1]$, respectively, must be a multiple of the matrix

$$\begin{pmatrix} u & v & 5w \\ -u & -2v & -w \\ 2u & v & 2w \end{pmatrix}, \quad \text{where} \begin{pmatrix} u & v & 5w \\ -u & -2v & -w \\ 2u & v & 2w \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Solving the equations

$$\begin{aligned} u + v + 5w &= 1 \\ -u - 2v - w &= 0 \\ 2u + v + 2w &= 1 \end{aligned}$$

we obtain $u = \frac{1}{2}, v = -\frac{1}{3}, w = \frac{1}{6}$. So a suitable choice of matrix for t_1 is $\begin{pmatrix} \frac{1}{2} & -\frac{1}{3} & \frac{5}{6} \\ -\frac{1}{2} & \frac{2}{3} & -\frac{1}{6} \\ 1 & -\frac{1}{3} & \frac{1}{3} \end{pmatrix}$, or more simply

$$\mathbf{A}_1 = \begin{pmatrix} 3 & -2 & 5 \\ -3 & 4 & -1 \\ 6 & -2 & 2 \end{pmatrix}$$

(b) Any matrix associated with the projective transformation t_2 which maps the Points $[1,0,0], [0,1,0], [0,0,1], [1,1,1]$ to the Points $[-1,3,-2], [-3,7,-5], [2,-5,4], [-3,8,-5]$, respectively, must be a multiple of the matrix

$$\begin{pmatrix} -u & -3v & 2w \\ 3u & 7v & -5w \\ -2u & -5v & 4w \end{pmatrix}, \text{ where } \begin{pmatrix} -u & -3v & 2w \\ 3u & 7v & -5w \\ -2u & -5v & 4w \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 8 \\ -5 \end{pmatrix}$$

Solving the equations

$$-u - 3v + 2w = -3$$
$$3u + 7v - 5w = 8$$
$$-2u - 5v + 4w = -5,$$

we obtain $u = 2, v = 1, w = 1$. So a suitable choice of matrix for t_2 is

$$\mathbf{A}_2 = \begin{pmatrix} -2 & -3 & 2 \\ 6 & 7 & -5 \\ -4 & -5 & 4 \end{pmatrix}.$$

(c) A matrix associated with the inverse, t_1^{-1}, of t_1 is $\mathbf{A}_1^{-1}$, which we can calculate to be

$$\mathbf{A}_1^{-1} = \begin{pmatrix} -\frac{1}{12} & \frac{1}{12} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{12} & -\frac{1}{12} \end{pmatrix}$$

then a simpler matrix associated with t_1^{-1} is

$$\mathbf{B} = \begin{pmatrix} 1 & -1 & -3 \\ 0 & -4 & -2 \\ -3 & -1 & 1 \end{pmatrix}$$

The required projective transformation is therefore $t : [\mathbf{x}] \mapsto [\mathbf{A}x]$, where

$$\begin{aligned}\mathbf{A} = \mathbf{A}_2\mathbf{B} &= \begin{pmatrix} -2 & -3 & 2 \\ 6 & 7 & -5 \\ 4 & 5 & 4 \end{pmatrix} \begin{pmatrix} 1 & -1 & -3 \\ 0 & -4 & -2 \\ 3 & -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -8 & 12 & 14 \\ 21 & -29 & -37 \\ -16 & 20 & 26 \end{pmatrix}.\end{aligned}$$

□

3.4 Use of Fundamental Theorem of Projective Geometry

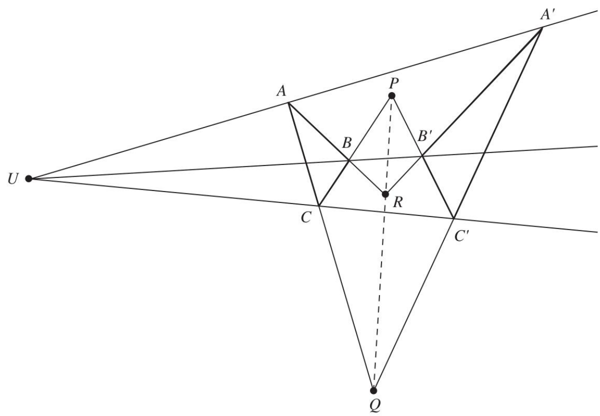
In Section 3.2 we described how an embedding plane π can be used to represent projective space $\mathbb{RP}^2$. The Points of $\mathbb{RP}^2$ are represented by Euclidean points in π and the Lines of $\mathbb{RP}^2$ are represented by Euclidean lines in π.

In general, any Euclidean figure in an embedding plane corresponds to a projective figure in $\mathbb{RP}^2$, and visa versa. This correspondence enables us to compare Euclidean theorems about a figure in an embedding plane with projective theorems about the corresponding projective figure. Provided that the theorems are concerned exclusively with projective properties, such as collinearity and incidence, then a Euclidean theorem will hold if and only if the corresponding projective theorem holds.

3.4.1 Desargues' Theorem and Pappus' Theorem

The advantage of interpreting a Euclidean theorem as a projective theorem in this way is that we can often obtain a much simpler proof of the theorem than would be possible using Euclidean geometry directly. We illustrate this by using projective geometry to prove the theorem of Desargues.

Theorem 1. (Desargues' Theorem) Let $\triangle ABC$ and $\Delta A'B'C'$ be triangles in $\mathbb{R}^2$ such that the lines AA', BB' and CC' meet at a point U. Let BC and $B'C'$ meet at P, CA and $C'A'$ meet at Q, and AB and $A'B'$ meet at R. Then P, Q and R are collinear.



Proof: Because this theorem is concerned exclusively with the projective properties of collinearity and incidence we can interpret it as a projective theorem in $\mathbb{RP}^2$. Moreover, by the Fundamental Theorem of Projective Geometry we know that any configuration of the theorem is projective-congruent to a configuration of the theorem in which $A = [1,0,0], B = [0,1,0], C = [0,0,1]$ and $U = [1,1,1]$. If we can prove the theorem in this special case. then we can use the fact that projective-congruence preserves projective properties to deduce that the theorem holds in general.

To prove the special case we use the algebraic techniques described in Section 3.2. First observe that the Line AU passes through the Points $[1, 0, 0]$ and $[1, 1, 1]$, and therefore has equation $y = z$. Since A' is a Point on AU, it must have homogeneous coordinates of the form $[a, b, b]$, for some real numbers a and b. Now, $b \neq 0$, since $A \neq A'$; so we may write the homogeneous coordinates of A' in the form $[p, 1, 1]$ (where $p = a/b$).

Similarly, the homogeneous coordinates of the Points B' and C' may be written in the form $[1, q, 1]$ and $[1, 1, r]$, respectively, for some real numbers q and r.

We now find the Point P where BC and $B'C'$ intersect. The Line BC has equation $x = 0$. Since the Line $B'C'$ passes through the Points $B' = [1, q, 1]$ and $C' = [1, 1, r]$, it must have equation

$$\begin{vmatrix} x & y & z \\ 1 & q & 1 \\ 1 & 1 & r \end{vmatrix} = 0$$

which we may rewrite in the form

$$(qr - 1)x - (r - 1)y + (1 - q)z = 0$$

It follows that at the Point P of intersection of the Lines BC and $B'C'$ we must have $x = 0$ and $(r - 1)y = (1 - q)z$, so that

P has homogeneous coordinates $[0, 1-q, r-1]$ Similarly, the Points Q and R have homogeneous coordinates $[1-p, 0, r-1]$ and $[1-p, q-1, 0]$, respectively. Now, the Points P, Q and R are collinear if

$$\begin{vmatrix} 0 & 1-q & r-1 \\ 1-p & 0 & r-1 \\ 1-p & q-1 & 0 \end{vmatrix} = 0$$

But

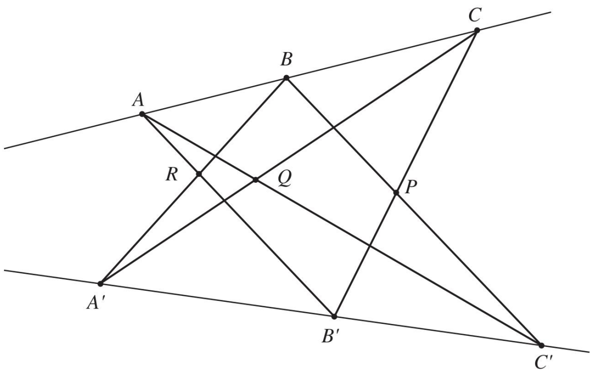
$$\begin{aligned} &\begin{vmatrix} 0 & 1-q & r-1 \\ 1-p & 0 & r-1 \\ 1-p & q-1 & 0 \end{vmatrix} \\ &= -(1-q)\begin{vmatrix} 1-p & r-1 \\ 1-p & 0 \end{vmatrix} + (r-1)\begin{vmatrix} 1-p & 0 \\ 1-p & q-1 \end{vmatrix} \\ &= -(1-q)(1-p)(1-r) + (r-1)(1-p)(q-1) \\ &= 0. \end{aligned}$$

It follows that P, Q and R are collinear, as asserted. The general result now holds, by projective-congruence. $\square$

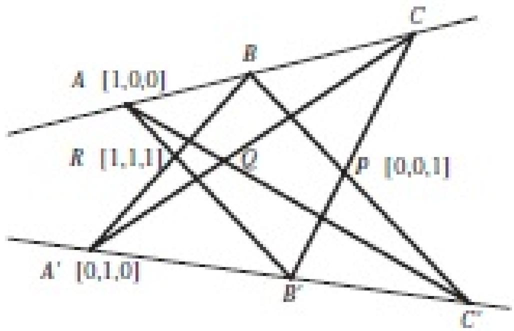
When using the Fundamental Theorem to simplify proofs of results in projective geometry, we do not usually refer to projective-congruence. Instead, so long as the properties involved are projective properties, we content ourselves with an initial remark of the type: 'By the Fundamental Theorem of Projective Geometry, we may choose the four Points..., no three of which are collinear, to be the triangle of reference and the unit Point; that is, to have homogeneous coordinates $[1,0,0], [0,1,0], [0,0,1]$ and $[1,1,1]$, respectively.

Next we use the Fundamental Theorem of Projective Geometry to prove Pappus' Theorem.

Theorem 2. (Pappus' Theorem) Let A, B and C be three points on a line in $\mathbb{R}^2$, and let A', B' and C' be three points on another line. Let BC' and $B'C$ meet at P, CA' and $C'A$ meet at Q, and AB' and $A'B$ meet at R. Then P, Q, R are collinear.



Proof: We interpret the theorem as a projective theorem, so: by the Fundamental Theorem of Projective Geometry we may choose the four Points A, A', P, R, no three of which are collinear, to be the triangle of reference and the unit Point; that is, to have homogeneous coordinates $[1,0,0], [0,1,0]$, $[0,0,1]$ and $[1,1,1]$, respectively.



First observe that the Line AR passes through the Points $[1, 0, 0]$ and $[1, 1, 1]$, and must therefore have equation $y = z$. Since B' is a Point on AR, it must have homogeneous coordinates of the form $[a, b, b]$ for some real numbers a and b. Now, $b \neq 0$ since $A \neq B'$, so we may write the homogeneous coordinates of B' in the form $[r, 1, 1]$ (where $r = a/b$).

Similarly, the Point B lies on the Line $x = z$ through the Points $A' = [0, 1, 0]$ and $R = [1, 1, 1]$, so it must have homogeneous coordinates of the form $[1, s, 1]$

Next we find the Point C where the Line AB intersects the Line $B'P$. Since the Line AB passes through the Points $A = [1, 0, 0]$ and $B = [1, s, 1]$, it must have equation $y = sz$. Also since the Line $B'P$ passes through the Points $B' = [r, 1, 1]$ and $P = [0, 0, 1]$ it must have equation $x = ry$. At the Point C where AB meets $B'P$ we have $y = sz$ and $x = ry$, so $C = [rs, s, 1]$

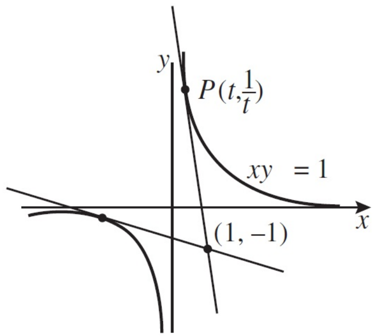
Similarly, C' is the point where the Line BP intersects the Line $A'B'$. Since $B = [1, s, 1]$ and $P = [0, 0, 1], BP$ has equation $y = sx$; and, since $A' = [0, 1, 0]$ and $B' = [r, 1, 1], A'B'$ has equation $x = rz$. It follows that $C' = [r, rs, 1]$

Finally we find the point Q where AC' intersects $A'C$. Since the Line AC' passes through the Points $A = [1, 0, 0]$ and $C' = [r, rs, 1]$ it must have equation $y = rsz$. Also the Line $A'C$ passes through the Points $A' = [0, 1, 0]$ and $C = [rs, s, 1]$ so it must have equation $x = rsz$. At the Point Q where AC' intersects A'C we have $y = rsz$ and $x = rsz$, so $Q = [rs, rs, 1]$

To complete the proof we simply observe that the Points $R = [1, 1, 1]$, $Q = [rs, rs, 1]$ and $P = [0, 0, 1]$ all lie on the Line $x = y$. It follows that P, Q and R are collinear. $\square$

Although we can sometimes simplify the proof of a Euclidean theorem by using projective geometry, there is another more subtle reason for interpreting a Euclidean theorem as a projective theorem. By doing so we can often avoid having to make special provision for exceptional cases, such as when two lines are parallel. In projective geometry, Lines which correspond to a pair of parallel lines in an embedding plane actually meet and are therefore no different to any other Lines.

As an example, consider the diagram . This illustrates the situation that occurs in Pappus' Theorem when the Point of intersection R of $A'B$ and AB' is an ideal Point for the embedding plane. The above proof of Pappus' Theorem is able to cope with this situation because it uses arguments from $\mathbb{RP}^2$! Our interpretation of the theorem on an embedding plane in this situation is that the Points P and Q must be collinear with the ideal Point R at which $A'B$ and AB' meet. That is, PQ must be parallel to both $A'B$ and AB'.

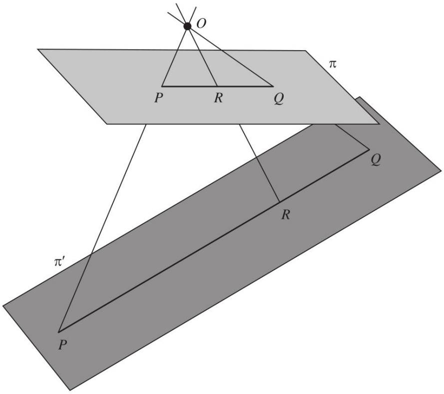


3.5 Cross-Ratio

3.5.1 Another Projective Property

Earlier. in Subsection 2.2.1. we noted that ratio of lengths along a line is an affine property. Thus, in affine geometry, if we are given two points P and Q on a line ℓ, then we can locate the position of a third point R along ℓ by specifying the ratio $PR : RQ$. In particular, it is possible to talk about the point midway between P and Q.

In projective geometry it is meaningless to talk about the Point midway between two other Points. In one embedding plane π a Point R may appear to be midway between the Points P and Q, whereas in another embedding plane π' the ratio $PR : RQ$ may be very different.



This ambiguity arises from the fact that perspectivities do not preserve the ratio of lengths along a line, so: ratio of lengths along a line is not a projective property.

In some embedding planes, such as the plane π' illustrated in the margin, the Point R does not even appear to lie between P and Q, so betweenness is not a projective property either!

Fortunately, there is a quantity, known as cross-ratio, that is preserved under all projective transformations. To see how this is defined, consider four collinear Points $A = [\mathbf{a}], B = [\mathbf{b}], C = [\mathbf{c}], D = [\mathbf{d}]$ in $\mathbb{RP}^2$. We can express the fact that A, B, C, D are collinear by writing $\mathbf{c}$ and $\mathbf{d}$ as linear combinations of $\mathbf{a}$ and $\mathbf{b}$. Thus we can write $\mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b}$ and $\mathbf{d} = \gamma\mathbf{a} + \delta\mathbf{b}$ for suitable real numbers $\alpha, \beta, \gamma, \delta$.

The cross-ratio is then defined to be the ratio of the ratios $\frac{\beta}{\alpha}$ and $\frac{\delta}{\gamma}$.

Definition. Let A, B, C, D be four collinear Points in $\mathbb{RP}^2$ represented by position vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$, and let

$$\mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b} \quad \text{and} \quad \mathbf{d} = \gamma\mathbf{a} + \delta\mathbf{b}$$

Then the cross-ratio of A, B, C, D is

$$(ABCD) = \frac{\beta}{\alpha} \bigg/ \frac{\delta}{\gamma}.$$

Of course, before we can be sure that this definition makes sense, we must ensure that it does not depend on the particular choice of position vectors $\mathbf{a}$, $\mathbf{b}, \mathbf{c}, \mathbf{d}$ that are used to represent the Points A, B, C, D. We shall check this shortly, but first we illustrate how cross-ratios are calculated.

Example 1. Let $A = [1, 2, 3], B = [1, 1, 2], C = [3, 5, 8], D = [1, -1, 0]$ be Points of $\mathbb{RP}^2$. Calculate the cross-ratio $(ABCD)$.

Solution: First, we have to find real numbers α and β such that the following vector equation holds:

$$(3, 5, 8) = \alpha(1, 2, 3) + \beta(1, 1, 2)$$

Comparing corresponding coordinates on both sides of this vector equation, we deduce that

$$3 = \alpha + \beta, \quad 5 = 2\alpha + \beta \quad \text{and} \quad 8 = 3\alpha + 2\beta$$

Solving these equations gives $\alpha = 2, \beta = 1$. Next, we find real numbers γ and δ such that the vector equation

$$(1, -1, 0) = \gamma(1, 2, 3) + \delta(1, 1, 2)$$

holds. Comparing corresponding coordinates on both sides of this vector equation, we deduce that

$$1 = \gamma + \delta, -1 = 2\gamma + \delta \quad \text{and} \quad 0 = 3\gamma + 2\delta$$

Solving these equations gives $\gamma = -2, \delta = 3$. It follows from the definition of cross-ratio that

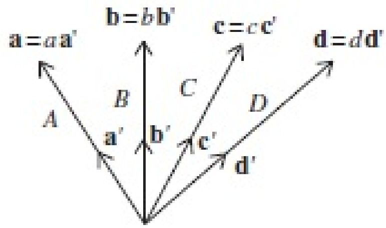
$$(ABCD) = \frac{\beta}{\alpha} \Big/ \frac{\delta}{\gamma} = \frac{1}{2} \Big/ \frac{3}{-2} = -\frac{1}{3}.$$

$\square$

Theorem 1. The cross-ratio $(ABCD)$ is independent of the homogeneous coordinates that are used to represent the collinear Points A, B, C, D.

Proof: Suppose that $A = [\mathbf{a}], B = [\mathbf{b}], C = [\mathbf{c}], D = [\mathbf{d}]$, and let

$$\mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b} \quad \text{and} \quad \mathbf{d} = \gamma\mathbf{a} + \delta\mathbf{b} \tag{1}$$



Now suppose that $A = [\mathbf{a}'], B = [\mathbf{b}'] . C = [\mathbf{c}'], D = [\mathbf{d}']$. Then

$$\mathbf{a} = a\mathbf{a}', \quad \mathbf{b} = b\mathbf{b}', \quad \mathbf{c} = c\mathbf{c}', \quad \mathbf{d} = d\mathbf{d}'$$

where a, b, c, d are some non-zero real numbers. By substituting these expressions into the equations (1), we obtain

$$c\mathbf{c}' = \alpha a\mathbf{a}' + \beta b\mathbf{b}' \quad \text{and} \quad d\mathbf{d}' = \gamma a\mathbf{a}' + \delta b\mathbf{b}'$$

which we can rewrite in the form

$$\mathbf{c}' = \alpha'\mathbf{a}' + \beta'\mathbf{b}' \quad \text{and} \quad \mathbf{d}' = \gamma'\mathbf{a}' + \delta'\mathbf{b}' \qquad (2)$$

where $\alpha' = \alpha a/c, \beta' = \beta b/c, \gamma' = \gamma a/d, \delta' = \delta b/d$

We can now check that equations (1) and (2) yield the same value for the cross-ratio:

$$\left.\frac{\beta'}{\alpha'}\right/\frac{\delta'}{\gamma'} = \left.\frac{\beta b/c}{\alpha a/c}\right/\frac{\delta b/d}{\gamma a/d}$$
$$= \left.\frac{\beta b}{\alpha a}\right/\frac{\delta b}{\gamma a}$$
$$= \left.\frac{\beta}{\alpha}\right/\frac{\delta}{\gamma}.$$

So, as expected, the cross-ratio is independent of the choice of homogeneous coordinates. □

The next problem illustrates that although the value of the cross-ratio $(ABCD)$ is independent of the choice of homogeneous coordinates that are used to represent $A, B,$ C$, D$, the value of the cross-ratio does depend on the order in which the Points A, B, C, D appear.

Theorem 2. Let A, B, C, D be four distinct collinear Points in $\mathbb{RP}^2$, and let $(ABCD) = k$. Then

$$(BACD) = (ABDC) = 1/k$$
$$(ACBD) = (DBCA) = 1 - k$$

Proof: Let $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ be any position vectors in $\mathbb{R}^3$ in the directions of the Points A, B, C, D, respectively, of $\mathbb{RP}^2$, and let $\alpha, \beta, \gamma, \delta$ be real numbers such that

$$\mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b} \quad \text{and} \quad \mathbf{d} = \gamma\mathbf{a} + \delta\mathbf{b}.$$

Then, by definition of cross-ratio, the cross-ratio $(ABCD)$ of the four Points A, B, C, D is the quantity

$$(ABCD) = \frac{\beta}{\alpha}\Big/\frac{\delta}{\gamma} = \frac{\beta\gamma}{\alpha\delta} = k \text{ , say.}$$

To determine $(BACD)$, we interchange the roles of A and B in the evaluation of $ABCD$ above; it follows that, since

$$\mathbf{c} = \beta\mathbf{b} + \alpha\mathbf{a} \quad \text{and} \quad \mathbf{d} = \delta\mathbf{d} + \gamma\mathbf{a}$$

the cross-ratio $(BACD)$ is the quantity

$$(BACD) = \frac{\alpha}{\beta}\Big/\frac{\gamma}{\delta} = \frac{\alpha\delta}{\beta\gamma} = \frac{1}{k}.$$

To determine $(ABDC)$, we interchange the roles of C and D in the evaluation of $(ABCD)$ above; it follows that, since

$$\mathbf{d} = \gamma\mathbf{a} + \delta\mathbf{b} \quad \text{and} \quad \mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b},$$

the cross-ratio $(ABDC)$ is the quantity

$$(ABDC) = \frac{\delta}{\gamma}\Big/\frac{\beta}{\alpha} = \frac{\alpha\delta}{\beta\gamma} = \frac{1}{k}.$$

To evaluate $(ACBD)$, we use the equations

$$\mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b} \quad \text{and} \quad \mathbf{d} = \gamma\mathbf{a} + \delta\mathbf{b} \tag{3}$$

to express $\mathbf{b}$ and $\mathbf{d}$ in terms of $\mathbf{a}$ and $\mathbf{c}$, as follows.

From the first equation in (3) we have

$$\begin{aligned} \mathbf{b} &= (\mathbf{c} - \alpha\mathbf{a})/\beta \\ &= (-\alpha/\beta)\mathbf{a} + (1/\beta)\mathbf{c} \qquad (4) \end{aligned}$$

If we then substitute this expression for **b** into the second equation in (3), we obtain

$$\begin{aligned} \mathbf{d} &= \gamma\mathbf{a} + \delta((-\alpha/\beta)\mathbf{a} + (1/\beta)\mathbf{c}) \\ &= ((\beta\gamma - \alpha\delta)/\beta)\mathbf{a} + (\delta/\beta)\mathbf{c} \qquad (5) \end{aligned}$$

It follows from the coefficients of **a** and **c** in equations (4) and (5) that

$$\begin{aligned} (ACBD) &= \frac{1/\beta}{-\alpha/\beta}\Big/\frac{\delta/\beta}{(\beta\gamma - \alpha\delta)/\beta} \\ &= -\left(\frac{\beta\gamma - \alpha\delta}{\alpha\delta}\right) \\ &= 1 - \frac{\beta\gamma}{\alpha\delta} \\ &= 1 - k \end{aligned}$$

Finally, we can use the previous parts of the proof to evaluate $(DBCA)$, as follows:

$$
\begin{array}{rcll}
\text{(DBCA)} & = & 1\ /\text{(BDCA)} & \text{(swap first two Points)} \\
& = & \text{(BDAC)} & \text{(swap last two Points)} \\
& = & \text{1-(BADC)} & \text{(swap middle two Points)} \\
& = & \text{1- 1/(ABDC)} & \text{(swap first two Points)} \\
& = & \text{1-(ABCD)} & \text{(swap last two Points)} \\
& = & \text{1-k} &
\end{array}
$$

□

Earlier, we showed that the cross-ratio $(ABCD)$ of the four collinear Points $A = [1, 2, 3], B = [1, 1, 2], C = [3, 5, 8], D = [1, -1, 0]$ in $\mathbb{RP}^2$ is $-\frac{1}{3}$. Theorem 2 enables us to deduce that

$$
\begin{array}{ll}
(BACD) = -3, & (ABDC) = -3 \\
(ACBD) = \frac{4}{3}, & (DBCA) = \frac{4}{3}.
\end{array}
$$

Problem 1. Let the P uints $A = [1, -1, -1], B = [1, 3, -21], C = [3.5, -5], D = [1, -5.0]$ be collinear Points of $\mathbb{RP}^2$. By applying Theorem 2 to the solution of Problem 1(a), determine the values of the cross-ratios $(ABDC), (DBCA)$ and $(ACBD)$.

The next theorem confirms that cross-ratio is preserved by projective transformations.

Theorem 3. Let t be a projective transformation, and let A, B, C, D be any four collinear Points in $\mathbb{RP}^2$. If $A' =$

$t(A), B' = t(B), C' = t(C), D' = t(D)$, then

$$(ABCD) = (A'B'C'D')$$

Proof: Let t be the projective transformation $t : [\mathbf{x}] \mapsto [\mathbf{Ax}]$, where $\mathbf{A}$ is an invertible 3×3 matrix. If $A = [\mathbf{a}], B = [\mathbf{b}], C = [\mathbf{c}], D = [\mathbf{d}]$, and

$$\mathbf{a}' = \mathbf{Aa}, \mathbf{b}' = \mathbf{Ab}, \mathbf{c}' = \mathbf{Ac}, \mathbf{d}' = \mathbf{Ad}$$
$$\text{then } A' = [\mathbf{a}'] , B' = [\mathbf{b}'] , C' = [\mathbf{c}'] , D' = [\mathbf{d}']$$

Since A, B, C, D are collinear, we can write

$$\mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b} \quad \text{and} \quad \mathbf{d} = \gamma\mathbf{a} + \delta\mathbf{b} \ , \tag{6}$$

so

$$(ABCD) = \frac{\beta}{\alpha}/\frac{\delta}{\gamma}$$

Multiplying each equation in (6) through by A, we obtain

$$\mathbf{c}' = \alpha\mathbf{a}' + \beta\mathbf{b}' \quad \text{and} \quad \mathbf{d}' = \gamma\mathbf{a}' + \delta\mathbf{b}',$$

so that

$$(A'B'C'D') = \frac{\beta}{\alpha}/\frac{\delta}{\gamma}$$

It follows that

$$(A'B'C'D') = (ABCD)$$

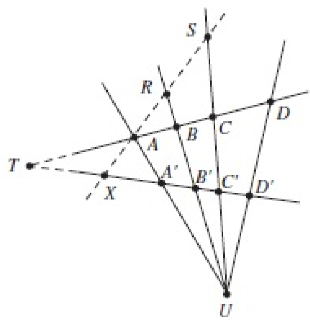
$\square$

We now use Theorem 3 to prove that if four distinct Points on a Line are in perspective with four distinct Points on another Line, then the cross-ratios of the four Points on each Line are equal.

Theorem 4. Let A, B, C, D be four distinct Points on a Line, and let A', B', C', D' be four distinct Points on another Line such that AA', BB', CC', DD' all meet at a Point U. Then

$$(ABCD) = (A'B'C'D')$$

Proof: By the Fundamental Theorem of Projective Geometry, there is a unique projective transformation t which maps B to B', C to C', B' to B, and C' to C. We shall show that $t(A) = A'$ and $t(D) = D'$, and hence by Theorem 3 it follows that $(ABCD) = (A'B'C'D')$



First observe that the composite $t \circ t$ fixes the Points B, C, B' and C'. By the Fundamental Theorem of Projective Geometry, the only projective transformation which does this is the identity transformation, so $t \circ t = i$ and $t = t^{-1}$.

Next observe that t maps the Line BC onto the Line $B'C'$, and vice versa; so the Point T at which BC and $B'C'$ intersect must be fixed by t. Also, t maps the Lines BB' and CC' onto themselves, so their Point of intersection U must be fixed by t.

Now let X be the image of A under t. Then X lies on $B'C'$. We want to show that $X = A'$.

Suppose that $X \neq A'$; then AX cannot pass through U so it must intersect BB' at R and CC' at S, where R, S and U are distinct Points.

Since t is self-inverse, it maps X back to A and therefore maps AX onto itself. But this implies that t fixes the four Points R, S, T, U; so by the Fundamental Theorem of Projective Geometry t must be the identity transformation. This is a contradiction with the hypothesis that the Lines $ABCD$ and $A'B'C'D'$ are different. It follows that we must conclude that $X = A'$, that is, $t(A) = A'$. A similar argument shows that $t(D) = D'$.

Finally, it follows by Theorem 3 that $(ABCD) = (A'B'C'D')$, as required. □

In affine geometry, if we are given two points A and B, then the ratio AC/CB uniquely determines a third point C on the line AB. We now explore the analogous result for projective geometry, namely that if we are given any three collinear Points A, B, C in $\mathbb{RP}^2$, then the value of the cross-ratio $(ABCD)$ uniquely determines a fourth Point D.

Theorem 5. (Unique Fourth Point Theorem) Let A, B, C, X, Y be collinear Points in $\mathbb{RP}^2$ such that

$$(ABCX) = (ABCY)$$

Then $X = Y$.

Proof: Let $A = [\mathbf{a}], B = [\mathbf{b}], C = [\mathbf{c}], X = [\mathbf{x}], Y = [\mathbf{y}]$. Since A, B, C, X, Y are collinear, it follows that there are real numbers $\alpha, \beta, \gamma, \delta, \lambda, \mu$ such that

$$\mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b}, \quad \mathbf{x} = \gamma\mathbf{a} + \delta\mathbf{b} \quad \text{and} \quad \mathbf{y} = \lambda\mathbf{a} + \mu\mathbf{b} \tag{7}$$

Then

$$(ABCX) = \frac{\beta\gamma}{\alpha\delta} \quad \text{and} \quad (ABCY) = \frac{\beta\lambda}{\alpha\mu}.$$

Since $(ABCX) = (ABCY)$, it follows that

$$\frac{\gamma}{\delta} = \frac{\lambda}{\mu}$$

so $\lambda = \gamma\mu/\delta$. If we substitute this value of λ into the expression for $\mathbf{y}$ in equation (7), we obtain

$$\mathbf{y} = (\gamma\mu/\delta)\mathbf{a} + \mu\mathbf{b} = (\mu/\delta)(\gamma\mathbf{a} + \delta\mathbf{b}) = (\mu/\delta)\mathbf{x}.$$

Since $\mathbf{y}$ is a scalar multiple of $\mathbf{x}$, it follows that $X = Y$, as required. $\square$

In Theorem 4 we showed that the cross-ratios $(ABCD)$ and $(A'B'C'D')$ are equal if the Points A', B', C', D' are in perspective with the Points A, B, C, D. Our next result is a partial converse of this result.

Theorem 6. Let A, B, C, D and A, E, F, G be two sets of collinear Points (on different Lines in $\mathbb{RP}^2$) such that the cross-ratios $(ABCD)$ and $(AEFG)$ are equal. Then the Lines BE, CF and DG are concurrent.

Proof: Let P be the Point at which the Lines BE and CF meet, and let X be the Point at which the Line PG meets the Line $ABCD$. Then the Points A, B, C and X are in perspective from P with the Points A, E, F and G, so that

$$(ABCX) = (AEFG)$$

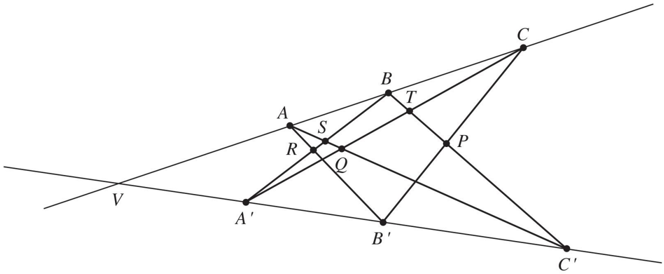
Since we know that $(AEFG) = (ABCD)$, it follows that

$$(ABCX) = (ABCD)$$

By Theorem 5, we must therefore have $X = D$. Hence the Points A, B, C, D and the Points A, E, F, G are in perspective from P. $\square$

We can now use Theorem 6 together with the other properties of cross-ratio to give a second proof of Pappus' Theorem.

Theorem 7. (Pappus' Theorem) Let A, B and C be three Points on a Line in $\mathbb{RP}^2$, and let A', B' and C' be three Points on another Line. Let BC' and $B'C$ meet at P, CA' and $C'A$ meet at Q, and AB' and $A'B$ meet at R. Then P, Q and R are collinear.



Proof: Let V be the Point of intersection of the two given Lines. Also let the Lines BA' and AC' meet at the Point S, and the Lines BC' and CA' meet at the Point T.

Now, the Points V, A', B', C' are in perspective from A with the Points B, A', R, S, so that

$$(VA'B'C') = (BA'RS) \tag{8}$$

Similarly, the Points V, A', B', C' are in perspective from C with the Points B, T, P, C', so that

$$(VA'B'C') = (BTPC') \tag{9}$$

It follows from equations (8) and (9) that

$$(BA'RS) = (BTPC')$$

so that by Theorem 6 the Lines $A'T, RP, SC'$ are concurrent.

We may rephrase this statement as follows: the Line RP passes through the Point where $A'T$ meets SC'; that is, the Line RP passes through Q. In other words, P, Q and R are collinear. $\square$

3.5.2 Cross-Ratio on Embedding Planes

So far, we have calculated a given cross-ratio $(ABCD)$ by applying the definition of cross-ratio directly to the Points A, B, C, D. However, it is sometimes convenient to evaluate the cross-ratio by examining the representation of the Points on some embedding plane.

Suppose that four collinear Points of $\mathbb{RP}^2$ pierce an embedding plane π at the points A, B, C, D with position vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$, respectively. According to the Section Formula, we can write $\mathbf{c}$ and $\mathbf{d}$ in the form

$$\mathbf{c} = \lambda\mathbf{a} + (1-\lambda)\mathbf{b} \quad \text{and} \quad \mathbf{d} = \mu\mathbf{a} + (1-\mu)\mathbf{b}.$$

where $(1-\lambda) : \lambda$ is the ratio $AC : CB$, and $(1-\mu) : \mu$ is the ratio $AD : DB$. Then from the definition of cross-ratio

$$(ABCD) = \frac{1-\lambda}{\lambda} \bigg/ \frac{1-\mu}{\mu},$$

so

$$(ABCD) = \frac{AC}{CB} \bigg/ \frac{AD}{DB}. \tag{10}$$

Example 2. In an embedding plane, the points A, B, C, D lie in order along a line with the distances AB, BC, CD being 1 unit, 3 units and 2 units, respectively. Determine the cross-ratios $(ABCD), (BACD)$ and $(ACBD)$.

Solution: Using equation (10) and the sign convention for ratios, we have

$$\begin{aligned}(ABCD) &= \frac{AC}{CB}/\frac{AD}{DB} = \left(-\frac{4}{3}\right)/\left(-\frac{6}{5}\right) = \frac{10}{9}\\ (BACD) &= \frac{BC}{CA}/\frac{BD}{DA} = \left(-\frac{3}{4}\right)/\left(-\frac{5}{6}\right) = \frac{9}{10}\end{aligned}$$

and

$$(ACBD) = \frac{AB}{BC}/\frac{AD}{DC} = \left(\frac{1}{3}\right)/\left(-\frac{6}{2}\right) = -\frac{1}{9}.$$

$\square$

Problem 2. The points A, B, C, D lie in order along a line with the distances AB, BC, CD being 2 units, 1 unit and 3 units, respectively. Determine the cross-ratios $(ABCD)$ and $(DBCA)$.

Sometimes one of the Points whose cross-ratio we are trying to find turns out to be an ideal Point for the embedding plane. In such cases, formula (10) cannot he used since some of the distances in the formula will not he defined.

To be specific, suppose that the Points A, B, C, D are collinear, but that A is an ideal Point for the embedding plane π, as shown in the margin. As before, we can let $\mathbf{b}, \mathbf{c}, \mathbf{d}$ be the position vectors of the points B, C, D on π, but we take a to be a unit vector along A. Then

$$\mathbf{c} = -(CB)\mathbf{a} + \mathbf{b} \quad \text{and} \quad \mathbf{d} = -(DB)\mathbf{a} + \mathbf{b}$$

From the definition of cross-ratio, it follows that

$$(ABCD) = \frac{1}{-CB} \Big/ \frac{1}{-DB} = \frac{DB}{CB} \tag{11}$$

We can now obtain the corresponding formulas for the cases where B, C or D is an ideal Point, by applying Theorem 2. For example, if B is an ideal Point, then

$$\begin{aligned} (ABCD) &= \tfrac{1}{(BACD)} && \text{(swap first two terms)} \\ &= (BADC) && \text{(swap last two terms)} \\ &= \tfrac{CA}{DA} && \text{by equation (11).} \end{aligned}$$

We now summarize the various formulas for cross-ratio in the form of a strategy, as follows.

Strategy. To use an embedding plane to calculate the cross-ratio of four collinear Points:

1. if the four Points pierce the embedding plane at A, B, C, D, then

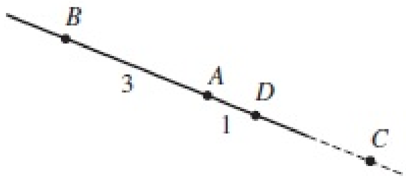
$$(ABCD) = \frac{AC}{CB} / \frac{AD}{DB}$$

2. if one of the Points is an ideal Point for the embedding plane, then

$$\begin{aligned}
(ABCD) &= \frac{DB}{CB} \text{ if } A \text{ is ideal,}\\
(ABCD) &= \frac{CA}{DA} \text{ if } B \text{ is ideal,}\\
(ABCD) &= \frac{BD}{AD} \text{ if } C \text{ is ideal,}\\
(ABCD) &= \frac{AC}{BC} \text{ if } D \text{ is ideal.}
\end{aligned}$$

Example 3. Determine $(ABCD)$ for the collinear points A, B, C, D illustrated in the margin, where C is an ideal Point.

Solution:



Since C is an ideal Point, we have

$$(ABCD) = \frac{BD}{AD} = \frac{4}{1} = 4$$

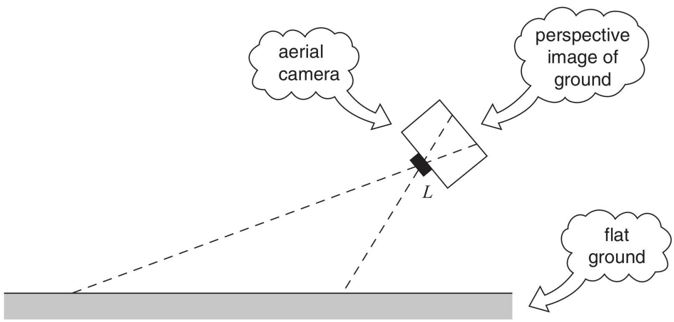
$\square$

3.5.3 An Application of Cross-Ratio

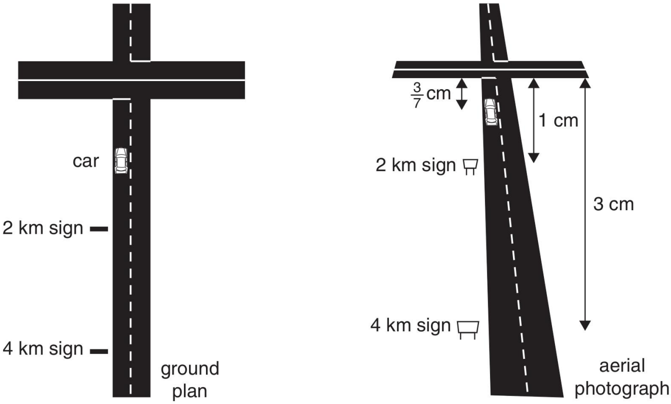
Earlier, we described how projective geometry can be used to obtain two dimensional representations of three-dimensional scenes. We now describe how cross-ratios can be used to obtain information about a three-dimensional scene from a two-dimensional representation of the scene. We do this in the context of aerial photography.

For simplicity, consider an aerial camera that takes pictures on a flat film behind its lens, L, of features on a flat piece of land in front of L. Since a point on the ground lies on the same line through L as its image on the film, we can regard the process of taking a photograph as a perspectivity centred at L.

Since collinearity is invariant under a perspectivity, the image of any line ℓ on the ground is a line on the film. Moreover, the cross-ratio of any four points on ℓ must be equal to the cross-ratio of their images on the film.

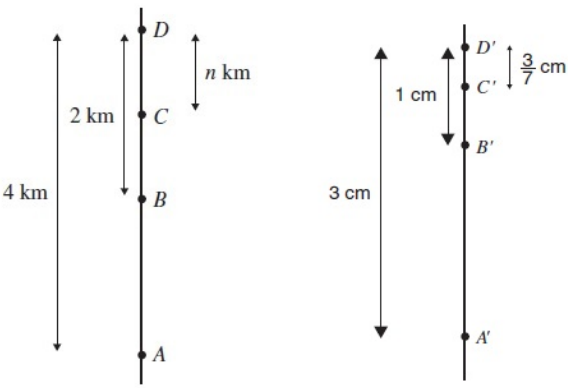


Example 4. An aerial camera photographs a car travelling along a straight road on flat ground towards a junction. Before the junction there are two warning signs at distances of 4 km and 2 km from the juncrion. On the film the signs are 1 cm and 3 cm from the junction, and the car is $\frac{3}{7}$ cm from the junction. How far is the car from the junction on the ground?



Solution: Let A and B denote the signs, C denote the car, and D denote the junction, and let A', B', C', D' be their images on the film. Then

$$\begin{aligned}(A'B'C'D') &= \frac{A'C'}{C'B'}\bigg/\frac{A'D'}{D'B'}\\ &= \left(-\frac{18/7}{4/7}\right)\bigg/\left(-\frac{3}{1}\right)\\ &= \frac{3}{2}.\end{aligned}$$



Now let the car be n km from the junction. Then

$$\begin{aligned}(ABCD) &= \frac{AC}{CB}\bigg/\frac{AD}{DB} \\ &= \left(-\frac{4-n}{2-n}\right)\bigg/\left(-\frac{4}{2}\right) \\ &= \frac{4-n}{2(2-n)}\end{aligned}$$

Since $(ABCD)$ and $(A'B'C'D')$ must be equal, it follows that

$$\frac{4-n}{2(2-n)} = \frac{3}{2}$$

Hence

$$4 - n = 3(2 - n)$$

and so $n = 1$. That is, the car is 1 km from the junction. $\square$

3.6 Exercises

Section 3.2

1. (a) Write down numbers a, b, c and d such that

$$[1, a, b] = \left[-\frac{1}{2}, 3, 4\right] \text{ and } [c, d, 2] = [3, 0, 1]$$

(b) Which of the following homogeneous coordinates represent the same Point of $\mathbb{RP}^2$ as $[4,-8,2]$?

(i) $[1,4,-2]$ (ii) $\left[\frac{1}{4},-\frac{1}{2},\frac{1}{8}\right]$ (iii) $\left[-\frac{1}{2},-2,1\right]$
(iv) $[-2,4,-1]$ (v) $\left[-\frac{1}{8},-\frac{1}{2},\frac{1}{4}\right]$

2. Determine an equation for each of the following Lines in $\mathbb{RP}^2$:

(a) the Line through the Points $[1,2,3]$ and $[3,0,-2]$;

(b) the Line through the Points $[1,-1,-1]$ and $[2,1,-3]$.

3. Determine whether each of the following sets of Points are collinear:

(a) $[1,-1,0],[1,0,-1]$ and $[2,-1,-1]$;

(b) $[1,0,1],[0,1,2]$ and $[1,2,3]$.

4. Determine the Point of intersection of each of the following pairs of Lines in $\mathbb{RP}^2$:

(a) the Lines with equations $x-2y+z=0$ and $x-y-z=0$;

(b) the Lines with equations $x+2y+5z=0$ and $3x-y+z=0$.

5. Determine the Point of $\mathbb{RP}^2$ at which the Line through the Points $[8,-1,2]$ and $[1,-2,-1]$ meets the Line through the Points $[0,1,-1]$ and $[2,3,1]$.

6. Determine the Point of $\mathbb{RP}^2$ at which the Line through the Points $[1,2,2]$ and $[2,3,3]$ meets the Line through the Points $[0,1,2]$ and $[0,1,3]$.

Section 3.3

In these exercises, you may find the following list of matrices and their inverses useful.

$$\mathbf{A}: \begin{pmatrix} 2 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad \begin{pmatrix} -2 & 0 & 1 \\ 0 & 3 & -2 \\ 1 & -3 & 1 \end{pmatrix} \begin{pmatrix} 0 & 3 & -1 \\ 2 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 3 & 4 \\ -1 & 3 & 2 \\ 3 & -3 & 3 \end{pmatrix}$$

$$\mathbf{A}^{-1}: \begin{pmatrix} 1 & 1 & -1 \\ -1 & -2 & 2 \\ 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} -1 & -1 & -1 \\ -\frac{2}{3} & -1 & -\frac{4}{3} \\ -1 & -2 & -2 \end{pmatrix}$$

$$\begin{pmatrix} 5 & -7 & -2 \\ 3 & -4 & -\frac{4}{3} \\ -2 & 3 & 1 \end{pmatrix}$$

1. Determine which of the following transformations t of

$\mathbb{RP}^2$ are projective transformations. For those that are projective transformations, write down a matrix associated with t.

(a) $t : [x, y, z] \mapsto [2x, y + 3z, 1]$

(b) $t : [x, y, z] \mapsto [x, x - y + 3z, x + y]$

(c) $t : [x, y, z] \mapsto [2y, y - 4z, x]$

(d) $t : [x, y, z] \mapsto [x + y - z, y + 3z, x + 2y + 2z]$

2. Determine the images of the Points $[1, 2, 3], [0, 1, 0]$ and $[1, -1, 1]$ under the projective transformation t associated with the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

3. Let

$$t_1 : \lfloor x, y, z] \mapsto [2x + y, -x + z, y + z]$$
$$t_2 : [x, y, z] \mapsto [x + y, 3x - z, 4y - 2z]$$

be projective transformations from $\mathbb{RP}^2$ to $\mathbb{RP}^2$.

(a) Write down matrices associated with each of t_1 and t_2.

(b) Determine formulas for $t_2 \circ t_1$ and $t_2 \circ t_1^{-1}$.

4. Find the image of the Line $x + 2y + 3z = 0$ under the projective transformation t_1 defined in Exercise 3.

5. Find the projective transformation t that maps the Points $[-1,0,0]$, $[-3,2,0]$, $[2,0,4]$, $[1,2,-5]$ to the Points $[2,1,0], [1,0,-1]$, $[0,3,-1], [3,-1,2]$, respectively.

6. Find the projective transformation t that maps the Points $[1,0,-3]$, $[1,1,-2]$, $[3,3,-5]$, $[6,4,-13]$ to the Points $[3,-5,3]$, $[\frac{1}{2},-1,0], [3,-5,6], [8,-13,12]$, respectively.

7. Determine matrices for the projective transformations which map the Points $[1,0,0], [0,1,0], [0,0,1]$ and $[1,1,1]$ onto the following Points:

   (a) $[-2,0,1], [0,1,-1], [-1,2,-1]$ and $[-1,1,-1]$;

   (b) $[0,1,0], [1,0,0], [-1,-1,1]$ and $[2,1,1]$;

   (c) $[0,1,-3], [1,1,-1], [4,2,3]$ and $[7,4,3]$.

8. Use the results of Exercise 7 to determine the projective transformations that map:

   (a) the Points

   $$[-2,0,1], [0,1,-1], [-1,2,-1], [-1,1,-1]$$

   to the Points

   $$[0,1,0], [1,0,0], [-1,-1,1], [2,1,1]$$

   respectively;

(b) the Points

$$[0, 1, 0], [1, 0, 0], [-1, -1, 1], [2, 1, 1]$$

to the Points

$$[0, 1, -3], [1, 1, -1], [4, 2, 3], [7, 4, 3]$$

respectively;

(c) the Points

$$[0, 1, -3], [1, 1, -1], [4, 2, 3], [7, 4, 3]$$

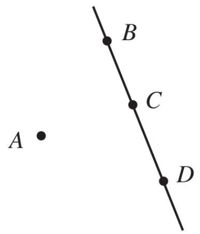
to the Points

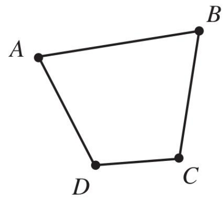
$$[-2.0.1].[0, 1, -1].[-1.2, -1].[-1, 1, -1]$$

respectively.

Section 3.4

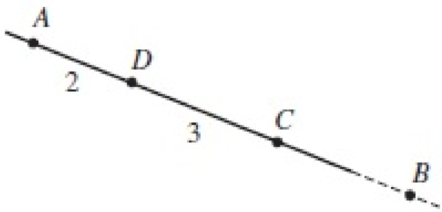
1. For which of the following configurations of Points A, B, C and D in $\mathbb{RP}^2$ is there a projective transformation sending A, B, C to the triangle of reference and D to the unit Point?





2. Let ΔABC be a triangle in $\mathbb{R}^2$, and let U be any point of $\mathbb{R}^2$ that is not collinear with any two of the points A, B, C. Let the Lines BC and AU meet at P, CA and BU meet at Q, and AB and CU meet at R. Prove that P, Q, R cannot be collinear.

3. Determine $(ABCD)$ for the collinear points A, B, C, D illustrated in the margin, where B is an ideal Point



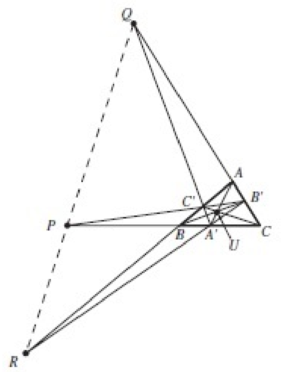
4. Prove that:

   (a) $(ABCD) = \frac{AC}{BC}$ if D is an ideal Point;

   (b) $(ABCD) = \frac{BD}{AD}$ if C is an ideal Point.

5. Give a Euclidean interpretation of Desargues' Theorem on an embedding plane π in the case where Q is an ideal Point for π.

6. Let ΔABC be a triangle in $\mathbb{R}^2$, and let U be any point of $\mathbb{R}^2$ that is not collinear with any two of the points A, B and C. Let the lines AU, BU and CU meet the lines BC, CA and AB at the points A', B' and C', respectively. Next, let the lines BC and $B'C'$ meet at P, AC and $A'C'$ meet at Q, and AB and $A'B'$ meet at R. Prove that P, Q and R are collinear.



Hint: Let A, B, C be the vertices of the triangle of reference, and let U be the unit Point. Then determine the homogeneous coordinates of the Points A', B' and C'.

Section 3.5

1. For each of the following sets of Points A, B, C, D, calculate the cross-ratio $(ABCD)$.

   (a) $A = [2, 1, 3], B = [1, 2, 3], C = [8, 1, 9], D = [4, -1, 3]$

(b) $A = [2,1,1], B = [-1,1,-1], C = [1,2,0], D = [-1,4,-2]$

(c) $A = [-1,1,1], B = [0,0,2], C = [5,-5,3], D = [-3,3,7]$

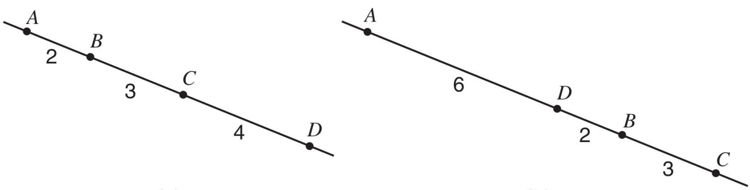
2. For the Points A, B, C, D in Exercise l(a), determine the cross-ratios $(BACD), (BDCA)$ and $(ADBC)$

3. Calculate the cross-ratio $(ABCD)$ for each of theh following sets of collinear Points in $R\mathbb{P}^2$.

(a) $A = [1,-1,-1], B = [1,3,-2], C = [3,5,-5], D = [1,-5,0]$

(b) $A = [1,2,3], B = [2,2,4], C = [-3,-5,-8], D = [3,-3,0]$

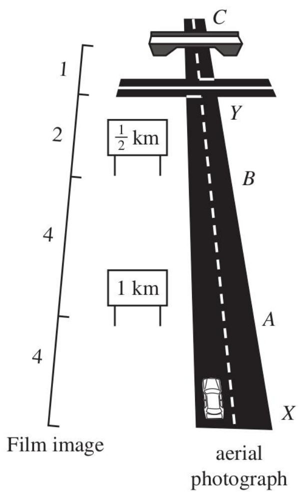
4. For each set of collinear points A, B, C, D illustrated below, calculate the cross-ratio $(ABCD)$.



(a) (b)

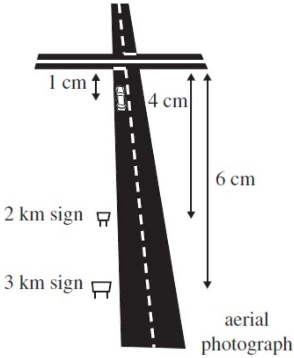
5. Calculate the cross-ratio $(ABCD)$ for the collinear points A, B, C, D illustrated below, where D is an ideal Point.

6. The following diagram represents an aerial photograph of a straight road on flat ground. At A there is a sign 'Junction 1 km ', at B a sign 'Junction $\frac{1}{2}$ km$'$, and C is the road junction. Also, a police patrol car is at X, and a bridge is at Y. The distances marked on the left of the diagram are measured in cm from the photograph.



Calculate the actual distances (in km) of the patrol car and the bridge from the junction.

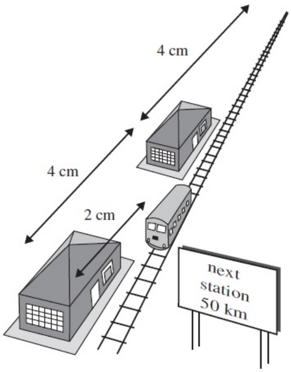
7. An aerial camera photographs a car travelling along a straight road on flat ground towards a junction. Before the junction there are two warning signs, at distances of 2 km and 3 km from the junction



On the film the signs are 4 cm and 6 cm from the junction, and the car is 1 cm from the junction. How far is the car from the junction on the ground?

If two lines that are known to be parallel on the ground appear to meet on the film, then the point of intersection on the film corresponds to the ideal Point where the 'parallel lines meet'. We can therefore use the above technique even when one of the Points is ideal, for we can use the second part of the strategy in Subsection 3.5.2 to calculate the cross-ratio whenever one of the Points is ideal.

8. An aerial camera photographs a train travelling between two stations along a straight traek 8n flat ground. The stations are 50 km apart. When the film is inspected, the stations are 4 cm apart, the train is midway between the stations, and the rails appear to meet (or vanish) 4 cm beyond the station towards which the train is travelling. How far has the train to travel to the next station?


4 cm
4 cm
2 cm
next
station
50 km