5
Curves

5.1 Introduction

Curves arise in many applications such as art, industrial design, mathematics, architecture, and engineering, and numerous computer drawing packages and computer-aided design packages have been developed to facilitate the creation of curves. A particularly illustrative application is that of computer fonts which are defined by curves that specify the outline of each character in the font. Different font sizes are obtained by applying scaling transformations. Special font effects can be obtained by applying other transformations such as shears and rotations. Likewise, in other applications there is a need to perform various tasks such as modifying, analyzing, and visualizing the curves. In order to execute such operations a mathematical representation for curves is required. In this chapter, curve representations are introduced and the simplest types of curve, namely lines and conics, are described. Chapters 6–8 explore Bézier and B-spline curves, two important representations that are widely used in CAD and computer graphics. The representations of curves lead naturally to representations of surfaces in Chapter 9. Conics also emerge as the silhouettes of quadric surfaces in Section 11.5.

Definition 5.1

Three representations of curves are considered.

Parametric: The coordinates of points of a *parametric* curve are expressed as functions of a variable or *parameter* such as t. A curve in the plane has the form $\mathbf{C}(t) = (x(t), y(t))$, and a curve in space has the form $\mathbf{C}(t) = (x(t), y(t), z(t))$. The functions $x(t)$, $y(t)$, and $z(t)$ are called the *coordinate functions*. The image of $\mathbf{C}(t)$ is called the *trace* of $\mathbf{C}$, and $\mathbf{C}(t)$ is called a *parametrization* of $\mathbf{C}$. A subset of a curve ~~**C** which is also a~~ curve is called a *curve segment*. A parametric curve defined by polynomial coordinate functions is called a *polynomial curve*. The *degree* of a polynomial curve is the highest power of the variable occurring in any coordinate function. A function $p(t)/q(t)$ is said to be *rational* if $p(t)$ and $q(t)$ are polynomials. A parametric curve defined by rational coordinate functions is called a *rational curve*. The *degree* of a rational curve is the highest power of the variable occurring in the numerator or denominator of any coordinate function. Most of the curves considered in this book are parametric.

Non-parametric explicit: The coordinates (x, y) of points of a *non-parametric explicit* planar curve satisfy $y = f(x)$ or $x = g(y)$. Such curves have the parametric form $\mathbf{C}(t) = (t, f(t))$ or $\mathbf{C}(t) = (g(t), t)$. For non-parametric explicit spatial curves, two of the coordinates are expressed in terms of the third: for instance, $x = f(z)$, $y = g(z)$.

Implicit: The coordinates (x, y) of points of an *implicit* curve satisfy $F(x, y) = 0$, for some function F. When F is a polynomial in variables x and y the curve is called an *algebraic curve*. An implicitly defined spatial curve must satisfy (at least) two conditions $F(x, y, z) = 0$ and $G(x, y, z) = 0$ simultaneously. Implicit curves defined by polynomials of degree two are considered in Section 5.6.

Example 5.2 (Parametric Curves)

1. Parabola: (t, t^2), for $t \in \mathbb{R}$, is a polynomial curve of degree 2. See Figure 5.1(a).

2. Quarter circle: $\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$, for $t \in [0, 1]$, is a rational curve of degree 2.

3. Unit radius circle: $(\cos(t), \sin(t))$, for $t \in [0, 2\pi]$, see Figure 5.1(b).

4. Twisted space cubic: (t, t^2, t^3), for $t \in \mathbb{R}$ is a polynomial curve of degree 3. See Figure 5.1(c).

5. Helix: $(r\cos(t), r\sin(t), at)$, for $t \in \mathbb{R}$, $r > 0$, $a \neq 0$. See Figure 5.1(d).

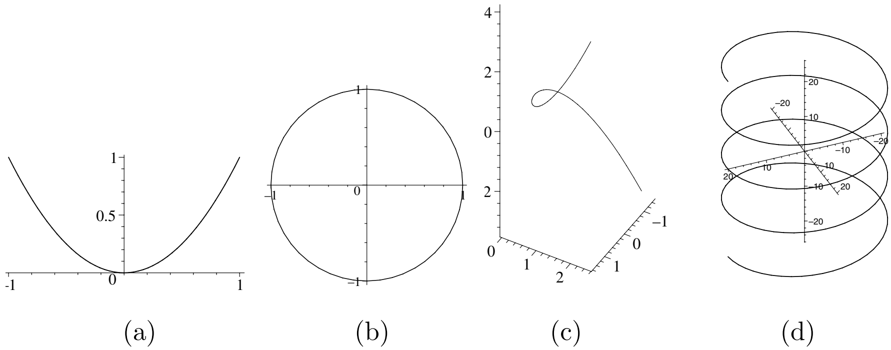


Figure 5.1 Parametric curves: (a) parabola, (b) unit circle, (c) twisted cubic, and (d) helix

Example 5.3 (Non-parametric Implicit Curves)

1. Parabola: $y = x^2$, $x \in \mathbb{R}$.
2. Circular arc: $y = \sqrt{1 - x^2}$, $x \in [-1, 1]$.
3. Twisted space cubic: $y = x^2$, $z = x^3$, $x \in \mathbb{R}$.

Example 5.4 (Implicit Curves)

1. Unit radius circle: $x^2 + y^2 - 1 = 0$.
2. Cuspidal cubic: $y^2 - x^3 = 0$, see Figure 5.2.

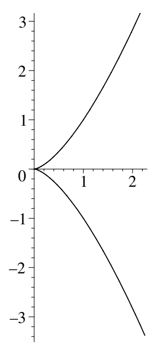


Figure 5.2 Cuspidal cubic $y^2 - x^3 = 0$

5.2 Curve Rendering

The process of drawing a curve is called *rendering*. Parametric curves are the most widely used in computer graphics and geometric modelling since points on the curve are easily computed. In contrast, the evaluation of points on an implicitly defined curve is substantially more difficult.

A curve of the form $\mathbf{C}(t) = (x(t), y(t))$ defined on the interval $[a, b]$ is rendered by evaluating $n+1$ points $(x(t_i), y(t_i))$, where $t_0 < t_1 < \cdots < t_n$ and $t_0 = a$, $t_n = b$. The points are joined in sequence by line segments to give a linear approximation to the curve as shown in Figure 5.3. If the resulting approximation is too jagged then a smoother curve can be obtained by increasing the number of evaluated points.

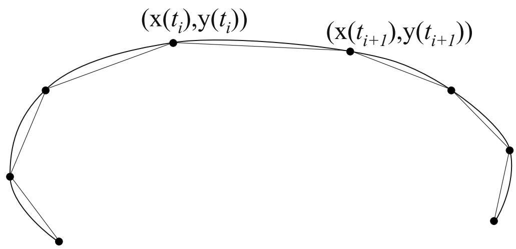


Figure 5.3 Linear approximation to a parametric curve

Points on polynomial and rational curves can be evaluated using a reasonable number of arithmetical operations. Points on curves defined by functions such as square roots, trigonometric functions, exponential and logarithmic functions are more computationally expensive to calculate.

The most economical way to evaluate a polynomial is to use *Horner's method*. Consider the polynomial $1 + 2t + 3t^2 + 4t^3$. If the polynomial is computed as $1 + 2 \cdot t + 3 \cdot t \cdot t + 4 \cdot t \cdot t \cdot t$ then 3 additions and 6 multiplications are required. However, if the polynomial is computed as $((4 \cdot t + 3) \cdot t + 2) \cdot t + 1$, then only 3 additions and 3 multiplications are required yielding a saving of 3 multiplications. For polynomials of higher degree the saving is even greater.

In general, a polynomial of the form $a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$ can be expressed in the form

$$(((a_n t + a_{n-1})\, t + a_{n-2})\, t + \cdots)\, t + a_0 \, .$$

A computer algorithm to evaluate a polynomial, based on Horner's method, is easily implemented.

EXERCISES

5.1. Express $3-5t+4t^2-2t^3+6t^4$ in Horner's form. Determine the difference in the number of $\pm$ and $\times$ required to evaluate the polynomial in its original and new form.

5.2. Write a computer program which takes as input the coefficients of a polynomial and a parameter value t, and which outputs the value of the polynomial at the given parameter value using Horner's method.

5.3. Determine the number of operations $\pm$ and $\times$ saved by evaluating a polynomial of degree n using Horner's method.

5.4. Write a computer program which renders a parametric curve. Alternatively, learn how to plot curves using a computer package. Plot some of the curves given in the examples.

5.3 Parametric Curves

Let $\mathbf{C}(t) = (x(t), y(t))$ be a curve defined on an open interval (a, b). Then $\mathbf{C}(t)$ is said to be C^k-continuous (or just C^k) if the first k derivatives of $x(t)$ and $y(t)$ exist and are continuous. If infinitely many derivatives exist then $\mathbf{C}(t)$ is said to be C^∞. A curve $\mathbf{C}(t) = (x(t), y(t))$ defined on a closed interval $[a, b]$ is said to be C^k-continuous if there exists an open interval (c, d) containing the interval $[a, b]$, and a C^k-continuous curve $\mathbf{D}(t)$ defined on (c, d), such that $\mathbf{C}(t) = \mathbf{D}(t)$ for all $t \in [a, b]$. Curves defined on a closed interval need to be "extendable" to a curve on an open interval in order to differentiate $x(t)$ and $y(t)$ at the ends of the interval. (Another type of continuity called G^k-continuity, which is important for CAD applications, is introduced in Definition 7.14.)

Suppose $\mathbf{C}(t)$ is a C^1 curve defined on an interval I, then the function $\nu(t) = \sqrt{(x'(t))^2 + (y'(t))^2}$ is called the *speed* of the curve $\mathbf{C}(t)$. If $\nu(t) \neq 0$, for all $t \in I$, then $\mathbf{C}(t)$ is said to be a *regular* curve. If $\nu(t) = 1$ for all $t \in I$, then $\mathbf{C}(t)$ is said to be a *unit speed curve.*

Example 5.5

1. Let $(x(t), y(t)) = (t, t^2)$. Then $(x'(t), y'(t)) = (1, 2t)$, and

$$\nu(t) = \sqrt{1^2 + (2t)^2} = \sqrt{1 + 4t^2}\ .$$

2. Let $(x(t), y(t), z(t)) = (\cos t, \sin t, t^2)$. Then

$$(x'(t), y'(t), z'(t)) = (-\sin t, \cos t, 2t) \ ,$$

and

$$\nu = \sqrt{(-\sin t)^2 + (\cos t)^2 + (2t)^2} = \sqrt{1+4t^2} \ .$$

Let $\mathbf{C}(t) = (x(t), y(t))$ be a regular parametric curve, and suppose $\mathbf{P}$ and $\mathbf{Q}$ are the points with coordinates $(x(t), y(t))$ and $(x(t+\delta t), y(t+\delta t))$ respectively. Let

$$\mathbf{t}_{\delta t} = \frac{\overrightarrow{\mathbf{PQ}}}{\delta t} = \frac{(x(t+\delta t), y(t+\delta t)) - (x(t), y(t))}{\delta t}$$

as shown in Figure 5.4. Then as $\delta t \to 0$, $\mathbf{Q} \to \mathbf{P}$ and $\mathbf{t}_{\delta t}$ tends to the limiting vector

$$\lim_{\delta t \to 0} \mathbf{t}_{\delta t} = \left(\lim_{\delta t \to 0} \frac{x(t+\delta t) - x(t)}{\delta t}, \lim_{\delta t \to 0} \frac{y(t+\delta t) - y(t)}{\delta t} \right) = (x'(t), y'(t)) \ .$$

$\mathbf{C}'(t) = (x'(t), y'(t))$ is called the *tangent vector*. The *unit tangent vector* is defined to be

$$\mathbf{t}(t) = \frac{1}{|(x'(t), y'(t))|}(x'(t), y'(t)) = \left(\frac{x'(t)}{\sqrt{x'(t)^2 + y'(t)^2}}, \frac{y'(t)}{\sqrt{x'(t)^2 + y'(t)^2}} \right) \ .$$

The line through the point $(x(t), y(t))$ in the direction of the tangent vector is called the *tangent line* and has the equation

$$y'(t)(x - x(t)) - x'(t)(y - y(t)) = 0 \ . \tag{5.1}$$

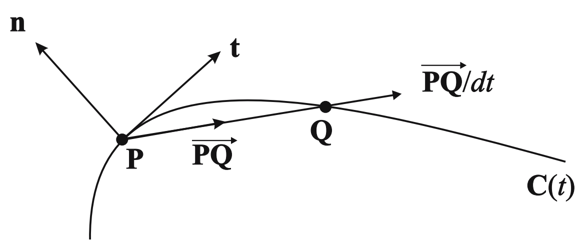


Figure 5.4 Tangent and normal to a curve

If the tangent vector $\mathbf{C}'(t) = (x'(t), y'(t))$ is rotated through an angle $\pi/2$ radians (in an anticlockwise direction), then the *normal vector* $(-y'(t), x'(t))$ is obtained. The *unit normal vector* of $\mathbf{C}(t)$ is defined to be

$$\mathbf{n}(t) = \frac{(-y'(t), x'(t))}{|(-y'(t), x'(t))|} = \left(\frac{-y'(t)}{\sqrt{x'(t)^2 + y'(t)^2}}, \frac{x'(t)}{\sqrt{x'(t)^2 + y'(t)^2}} \right) \ .$$

Example 5.6

Let $\mathbf{C}(t) = (\cos(t), \sin(t))$. Then the tangent vector is $\mathbf{C}'(t) = (-\sin(t), \cos(t))$ and the normal vector is $(-\cos(t), -\sin(t))$. Since $|\mathbf{C}'(t)| = 1$ these vectors are also the unit tangent and normal vectors. At the point $(\cos(\pi/4), \sin(\pi/4)) = \left(1/\sqrt{2}\,, 1/\sqrt{2}\right)$ the unit tangent vector is

$$(-\sin(\pi/4), \cos(\pi/4)) = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) ,$$

and the unit normal vector is

$$(-\cos(\pi/4), -\sin(\pi/4)) = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) .$$

The tangent line to $\mathbf{C}(t)$ at $\left(1/\sqrt{2}\,, 1/\sqrt{2}\right)$ is

$$\cos(\pi/4)(x - \cos(\pi/4)) + \sin(\pi/4)(y - \sin(\pi/4)) = 0 ,$$

which simplifies to $x + y - \sqrt{2} = 0$.

The derivation of the tangent vector can be extended to space curves: for a curve $\mathbf{C}(t) = (x(t), y(t), z(t))$, the tangent vector is $\mathbf{C}'(t) = (x'(t), y'(t), z'(t))$, and the unit tangent vector is $\mathbf{t}(t) = \mathbf{C}'(t)/\sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}$.

EXERCISES

5.5. Let $\mathbf{C}(t) = (x(t), y(t))$ be a regular curve.

(a) Determine the parametric equation of the tangent line to $\mathbf{C}$ at the point $(x(t), y(t))$.

(b) Prove that the tangent line to $\mathbf{C}(t)$ at a point $(x(t), y(t))$ is given by Equation (5.1).

(c) The normal line to $\mathbf{C}$ at a point $\mathbf{p} = (x(t), y(t))$ is the line through $\mathbf{p}$ perpendicular to the tangent. Determine the implicit equation of the normal line.

5.6. For each of the curves below, determine (i) the unit tangent vector, (ii) the unit normal vector, and (iii) the implicit equation of the tangent line.

(a) (t, t^2) at the point $(1, 1)$.

(b) (t^2, t^3) at the point $(4, 8)$.

(c) *logarithmic spiral:* $\left(ae^{bt}\cos t, ae^{bt}\sin t\right)$.

(d) *cycloid:* $(t+\sin t, 1-\cos t)$.

(e) $(\cos t + t\sin t, \sin t - t\cos t)$.

(f) *catenary:* $(c\cosh(t/c), t)$.

5.7. Show that a translation of a curve has no effect on the speed of a curve.

5.8. Show that a rotation has no effect on the speed of a curve.

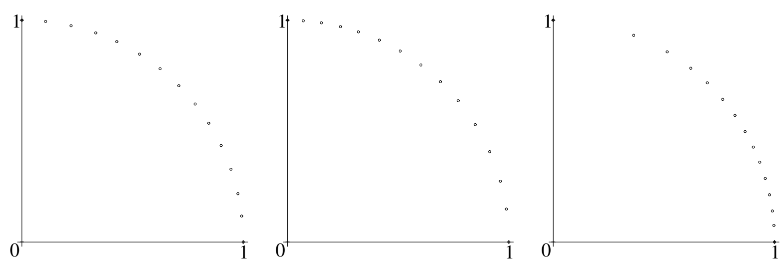
5.4 Arclength and Reparametrization

Consider the following three parametrizations of the unit quarter circle (in the first quadrant) centred at the origin.

(1) $(\cos\frac{\pi}{2}t, \sin\frac{\pi}{2}t), t \in [0,1]$,

(2) $\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right), t \in [0,1]$, and

(3) $\left(\sqrt{1-t^2}, t\right), t \in [0,1]$.



(a) (b) (c)

Figure 5.5 Different parametrizations of the quarter circle: (a) $(\cos\frac{\pi}{2}t, \sin\frac{\pi}{2}t)$, (b) $\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$, and (c) $\left(\sqrt{1-t^2}, t\right)$

Figure 5.5 shows the three parametrizations of the quarter circle evaluated at 15 equal parameter increments $t_i = i/14$, for $i = 0, \ldots, 14$. For parametrization (1) the points are equally spaced along the arc, for (2) the points are quite evenly spaced, and for (3) the points are unevenly spaced. The difference in the behaviour of the parametrizations corresponds to the fact that in (1) each parameter interval $[t_i, t_{i+1}]$ maps to a circular arc of equal length, whereas in (2) and (3) the parameter intervals map to circular arcs of varying lengths.

To explore this further a method to calculate the length of a curve is required. Consider a regular curve $\mathbf{C}(t) = (x(t), y(t))$, for $t \in [a, b]$, and a sequence of equally spaced parameter values $t_i = a + \frac{i}{n}(b - a)$ (for $i = 0, \ldots, n$) with $t_0 = a$ and $t_n = b$. The line segment from $(x(t_i), y(t_i))$ to $(x(t_{i+1}), y(t_{i+1}))$ approximates the curve on the interval $[t_i, t_{i+1}]$ and has length

$$\sqrt{(x(t_{i+1}) - x(t_i))^2 + (y(t_{i+1}) - y(t_i))^2} \; .$$

Thus the length $L(\mathbf{C})$ of the curve $\mathbf{C}$ on the interval $[a, b]$ is approximately

$$\sum_{i=0}^{n-1} \sqrt{(x(t_{i+1}) - x(t_i))^2 + (y(t_{i+1}) - y(t_i))^2} \; . \tag{5.2}$$

If the parameter increments $\delta t = t_{i+1} - t_i = (b - a)/n$ are sufficiently small, then $x'(t_i)$ is approximately equal to $(x(t_{i+1}) - x(t_i))/ (t_{i+1} - t_i)$, $y'(t_i)$ is approximately equal to $(y(t_{i+1}) - y(t_i))/ (t_{i+1} - t_i)$, and substitution into (5.2) yields that $L(\mathbf{C})$ is approximately

$$\sum_{i=0}^{n-1} \sqrt{(x'(t_i))^2 + (y'(t_i))^2}\, \delta t \; . \tag{5.3}$$

The true length of the curve over $[a, b]$ is realized by letting n tend to infinity. As n increases the line segments fit the curve more closely, and (5.3) becomes a better approximation to the length of the curve. The limit of (5.3) as n tends to infinity is

$$L(\mathbf{C}) = \int_a^b \sqrt{(x'(t))^2 + (y'(t))^2}\; dt = \int_a^b \nu(t)\; dt \; .$$

$L(\mathbf{C})$ is called the *arclength* of $\mathbf{C}(t)$ from $t = a$ to $t = b$. The *arclength function* $L_{\mathbf{C}}(t) = \int_{t_0}^{t} \sqrt{x'(u)^2 + y'(u)^2}\;\, du$, for $a \le t_0 \le b$, measures the length of the curve segment from an initial point $(x(t_0), y(t_0))$ to the point $(x(t), y(t))$. Then $L(\mathbf{C}) = L_{\mathbf{C}}(b) - L_{\mathbf{C}}(a)$.

Example 5.7

1. The speed of the quarter circle $\mathbf{C}(t) = (\cos t, \sin t)$, for $t \in \left[0, \frac{\pi}{2}\right]$, is $\nu(t) = \sqrt{(-\sin t)^2 + (\cos t)^2} = 1$. The parametrization is unit speed and the arclength function from $t_0 = 0$ is $L_{\mathbf{C}}(t) = \int_0^t 1\; du = t$. The curve has length $L_{\mathbf{C}}\left(\frac{\pi}{2}\right) - L_{\mathbf{C}}(0) = \frac{\pi}{2}$.

2. The speed of $\mathbf{C}(t) = (\cos \frac{\pi}{2}t, \sin \frac{\pi}{2}t)$, for $t \in [0,1]$, is

$$\nu(t) = \sqrt{\left(-\frac{\pi}{2}\sin\frac{\pi}{2}t\right)^2 + \left(\frac{\pi}{2}\cos\frac{\pi}{2}t\right)^2} = \frac{\pi}{2}\ .$$

   The parametrization has constant speed and the arclength function from $t_0 = 0$ is $L_{\mathbf{C}}(t) = \int_0^t \frac{\pi}{2}\, du = \frac{\pi}{2}t$. The curve has length $L_{\mathbf{C}}(1) - L_{\mathbf{C}}(0) = \frac{\pi}{2}$.

3. The speed of $\mathbf{C}(t) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$, for $t \in [0,1]$, is

$$\nu(t) = \sqrt{\left(\frac{-4t}{(1+t^2)^2}\right)^2 + \left(\frac{2(1-t^2)}{(1+t^2)^2}\right)^2} = \frac{2}{1+t^2}\ ,$$

   and the arclength function from $t_0 = 0$ is $L_{\mathbf{C}}(t) = \int_0^t \frac{2}{1+u^2}\, du = 2\arctan(t)$. Thus, with this parametrization, the unit quarter circle starts at point $(1,0)$ with speed $\nu(0) = 2$. As t increases the speed decreases until the curve reaches the point $(0,1)$ when the curve has speed $\nu(1) = 1$. The curve has length $L_{\mathbf{C}}(1) - L_{\mathbf{C}}(0) = \frac{\pi}{2}$.

4. Let the unit quarter circle be parametrized by $\mathbf{C}(t) = \left(\sqrt{1-t^2}, t\right)$, for $t \in [0,1]$. Then $(x'(t), y'(t)) = (-t(1-t^2)^{-1/2}, 1)$, and

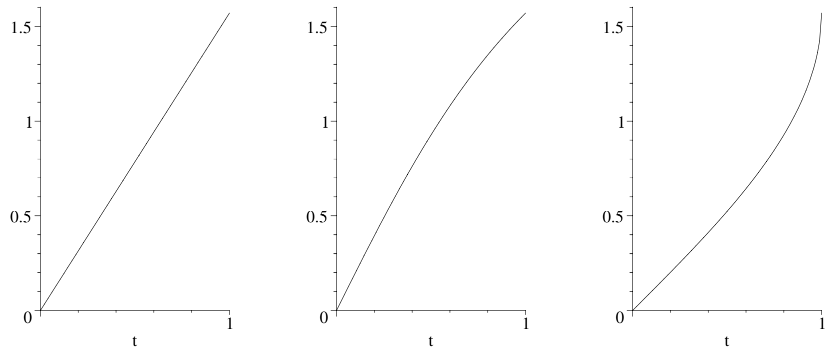
$$\nu(t) = \sqrt{\left(-t(1-t^2)^{-1/2}\right)^2 + 1} = (1-t^2)^{-1/2}\ .$$

   Thus, with this parametrization, the unit quarter circle starts at point $(1,0)$ with speed $\nu(0) = 1$. As t increases the speed increases until the curve reaches the point $(0,1)$ when the curve has infinite speed. The arclength is computed in Exercise 5.9.

The arclength functions of the three parametrizations of the quarter circle are illustrated in Figure 5.6. In each plot the vertical axis shows the length of the curve traced from $(x(0), y(0))$ to $(x(t), y(t))$. Naturally, the total curve length in each case is $\pi/2$. Parametrization (1) traces the curve uniformly. The speed of parametrization (2) is decreasing, so the curve is traced more quickly in the beginning than at the end. The speed of parametrization (3) is increasing so the curve is traced more quickly at the end than at the beginning.

Definition 5.8

Let $\mathbf{C}(t)$ and $\mathbf{D}(t)$ be curves defined on intervals I and J respectively. Then $\mathbf{D}$ is said to be a *reparametrization* of $\mathbf{C}$ if there exists a differentiable function $h : J \to I$ such that $h'(t) \neq 0$ and $\mathbf{D}(t) = \mathbf{C}(h(t))$ for all $t \in J$. The function $h(t)$ is referred to as a *reparametrization*.



(a) (b) (c)

Figure 5.6 Comparison of the arclength functions of three parametrizations of the quarter circle

Example 5.7(1) is a unit-speed parametrization of the unit quarter circle. Parametrization (1) can be obtained from the unit speed parametrization by a reparametrization with $h(t) = \frac{\pi}{2}t$ and $J = [0, 1]$. Parametrizations (2) and (3) are also reparametrizations of the unit speed quarter circle. The next theorem shows that the arclength function can be used to reparametrize a curve to give a unit speed curve with the same trace.

Theorem 5.9

Let $\mathbf{C}(t) = (x(t), y(t))$ be a regular curve defined on an interval I with arclength function $s(t) = L_{\hat{\mathbf{C}}}(t)$. Then $\mathbf{C}(s) = \left(x\left(L_{\mathbf{C}}^{-1}(s)\right), y\left(L_{\mathbf{C}}^{-1}(s)\right)\right)$ is a unit speed curve.

Proof

Let $s = L_{\mathbf{C}}(t)$. Differentiating with respect to t gives $L'_{\mathbf{C}}(t) = |\mathbf{C}'(t)| = \nu(t)$. Since $\mathbf{C}(t)$ is regular, $L'_{\mathbf{C}}(t) \neq 0$ for all $u \in I$, and the inverse function theorem implies that the inverse $t = L_{\mathbf{C}}^{-1}(s)$ exists. Let $h(s) = L_{\mathbf{C}}^{-1}(s)$. Then $\frac{dh}{ds} = 1/L'_{\mathbf{C}} \neq 0$, and $t = h(s)$ may be used to reparametrize $\mathbf{C}(t)$ to give the curve $\hat{\mathbf{C}}(s) = \mathbf{C}(L_{\mathbf{C}}^{-1}(s)) = \left(x\left(L_{\mathbf{C}}^{-1}(s)\right), y\left(L_{\mathbf{C}}^{-1}(s)\right)\right)$. Then the chain rule gives

$$\frac{d\hat{\mathbf{C}}}{ds}(s) = \frac{dL_{\mathbf{C}}^{-1}(s)}{ds}\left(x'\left(L_{\mathbf{C}}^{-1}(s)\right), y'\left(L_{\mathbf{C}}^{-1}(s)\right)\right) ,$$

and

$$\begin{aligned}
\left|\frac{d\hat{\mathbf{C}}}{ds}(s)\right| &= \left|\frac{dL_{\mathbf{C}}^{-1}(s)}{ds}\right| \left|\left(x'\left(L_{\mathbf{C}}^{-1}(s)\right), y'\left(L_{\mathbf{C}}^{-1}(s)\right)\right)\right| \\
&= \left|\frac{1}{L'_{\mathbf{C}}(t)}\right| |\mathbf{C}'(t)| = \frac{1}{|\mathbf{C}'(t)|} |\mathbf{C}'(t)| = 1 \,.
\end{aligned}$$

Hence $\hat{\mathbf{C}}(s)$ is unit speed, and the proof is complete.

□

Example 5.10

Consider parametrization (3) of the unit quarter circle. Exercise 5.9 will determine the arclength function to be $s = \arcsin(u)$. Substituting $t = \sin(s)$ into $\left(\sqrt{1-t^2}, t\right)$ gives the unit speed parametrization of the circle $(\cos(s), \sin(s))$.

EXERCISES

5.9. Show that the arclength function of $\mathbf{C}(t) = \left(\sqrt{1-t^2}, t\right), t \in [0,1]$ is $L_{\mathbf{C}}(t) = \arcsin t$ (assume arcsin has range $\left[-\frac{1}{2}\pi, \frac{1}{2}\pi\right]$).

5.10. Determine the speed and arclength function for each of the following curves

(a) *cycloid:* $(t + \sin t, 1 - \cos t)$, $t \in [-\pi, \pi]$.

(b) $(\cos t + t \sin t, \sin t - t \cos t)$, $t > 0$.

(c) *catenary:* $(c \cosh(t/c), t)$, $t \in \mathbb{R}$.

(d) *astroid:* $\left(\cos^3 t, \sin^3 t\right)$, $t \in [0, \pi/2]$.

(e) *logarithmic spiral:* $\left(ae^{bt} \cos t, ae^{bt} \sin t\right)$.

5.11. Determine the length of the cycloid and the astroid of the previous exercise.

5.12. Write a program to determine the arclength of a curve using the summation formula (5.2) to within a user specified accuracy $\epsilon > 0$. This is achieved by increasing the number of increments n until the difference between successive computed approximate lengths is less than ϵ.

5.13. Determine the unit speed parametrization of the unit quarter circle from parametrization (2) by reparametrizing with the arclength function.

5.14. Obtain a unit speed reparametrization of

(a) *cycloid:* $(t + \sin t, 1 - \cos t)$.

(b) $(\cos t + t \sin t, \sin t - t \cos t)$.

(c) *catenary:* $(c \cosh(t/c), t)$.

(d) *logarithmic spiral:* $\left(ae^{bt} \cos t, ae^{bt} \sin t\right)$.

(e) $y = \cosh(x - 1)$.

5.5 Application: Numerical Controlled Machining and Offsets

Numerically controlled (NC) milling machines are used to make products and parts, or the moulds and dies from which the parts are manufactured. A CAD definition of a curve, describing the shape of a part, can be converted into a sequence of commands which are used to drive the milling machine cutting tool. NC machines can be programmed to move the tool in various ways. For instance, a five-axis machine can perform both translations and orientations of the tool, whereas a two-axis machine can translate the tool freely in the x- and y-directions, but a fixed orientation of the tool is maintained. The NC machine is programmed to move the cutter along a path so that the unwanted portion of the material is removed, and the remaining material has the desired shape.

In many applications the tool is a *ball-end* or *ball-nose cutter*. For a two-axis machine, cutting in a specified plane, the ball-end cutter can be considered a circular disk of fixed radius d. Suppose the shape to be cut is given by a regular curve $\mathbf{C}(t) = (x(t), y(t))$, with unit normal $\mathbf{n}(t)$. Referring to Figure 5.7, the cutter disk is required to be perpendicular to the curve, which implies that the disk centre is a distance d along the curve normal. Therefore, as the shape is cut, the disk centre follows the path of the curve $\mathbf{O}_d(t) = \mathbf{C}(t) + d \cdot \mathbf{n}(t)$ called the *offset* or *parallel* of $\mathbf{C}(t)$ at a distance d. The sign of d determines which side of the curve the cutter lies. Offset curves generalise to offset surfaces which are discussed in Section 9.2.1.

Example 5.11

Consider the curve $\mathbf{C}(t) = (x(t), y(t)) = (t, t^2)$. Then $(x'(t), y'(t)) = (1, 2t)$.

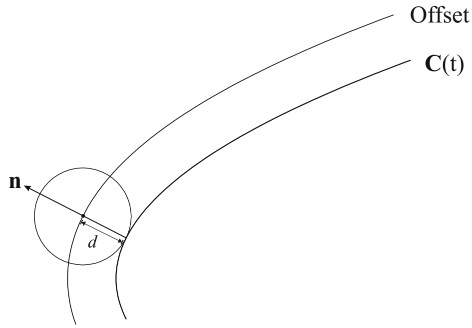


Figure 5.7 Path of centre of ball-end cutter along offset

Hence, $\mathbf{n}(t) = \left(-\frac{2t}{\sqrt{1+4t^2}}, \frac{1}{\sqrt{1+4t^2}}\right)$, and the offset at a distance d is

$$\mathbf{O}_d(t) = \left(t - d\frac{2t}{\sqrt{1+4t^2}}, t^2 + d\frac{1}{\sqrt{1+4t^2}}\right) .$$

Figure 5.8 shows the offsets at distances $d = -2, -1, 0.5, 1$. Note that the offsets are not parabolas. The $d = 1$ offset exhibits *cusp singularities*. If the cutting is to be executed on the same side as the normal to the parabola, then the cutting disk must have a radius less than 1 in order to avoid singularities of the offset. Such singularities indicate that the cutting tool is too big to cut the desired shape. (See Exercise 10.7.)

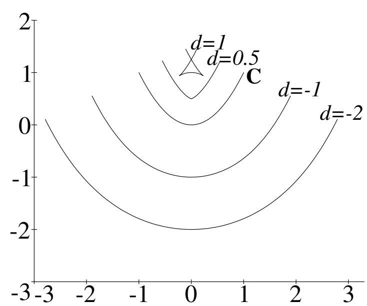


Figure 5.8 Offsets of the parabola (t, t^2)

EXERCISES

5.15. Determine the offset of $\mathbf{C}(t) = \left(1 - 3t + 3t^2, 3t^2 - 2t^3\right)$ at a distance d. Plot the curve and its offset at a distance $d = 1$.

5.16. Determine the offsets at a distance d for the following curves:

(a) $(c \cosh(t/c), t)$.

(b) $(e^{bt} \cos t, e^{bt} \sin t)$.

(c) $(\cos t + t \sin t, \sin t - t \cos t)$.

5.17. Show that the offset at a distance d of a circle of radius r is a circle of radius $r + d$.

5.6 Conics

The simplest implicitly defined planar curve is a straight line given by a linear equation $ax + by + c = 0$. Curves defined implicitly by a quadratic polynomial equation

$$ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0 \tag{5.4}$$

are called *conics*. Circles, ellipses, hyperbolas, and parabolas are all types of conic. "Conics" or "conic sections" receive their name from a classical geometrical method of construction, namely, as the curve of intersection of a plane with a cone.

A *cone* is the surface formed by rotating a line **L** through a fixed point **O** about a fixed axis **OA** so that **L** maintains a constant angle $\alpha < \pi/2$ with the axis. The point **O** is called the *vertex* of the cone. The cone consists of two parts called *sheets* which meet at the vertex. Consider a plane, not passing through **O**, making an angle β with the axis. When $\beta > \alpha$, the intersection curve of the plane and the cone is an *ellipse* lying entirely in one sheet. When $\beta = \pi/2$ (so the axis is perpendicular to the plane) the intersection is a *circle*, a special case of the ellipse. When $\beta < \alpha$, the plane intersects both sheets of the cone resulting in a curve of two separate branches called a *hyperbola*. When $\beta = \alpha$, the plane intersects the cone in one sheet to give a curve called a *parabola*. The ellipse, parabola, and hyperbola are illustrated in Figure 5.9. There are also degenerate conics which arise when the plane passes through the vertex. The degenerate cases are a union of two lines when $\beta > \alpha$, two coincident lines when $\beta = \alpha$, and the point **O** when $\beta < \alpha$.

If **L** is a line parallel to the axis, then the resulting surface is a cylinder which may be considered a cone with its vertex at infinity. A plane intersects

(a) (b) (c)

Figure 5.9 Conic sections

the cylinder in an ellipse, or in the degenerate cases of two distinct parallel lines, two coincident lines, or no intersection. The reader is referred to [26] and [5] for a historical account of conics and a proof that the sections of a cone are expressible by quadratic equations.

There is a second geometric construction for conics called the focus–directrix construction [26]. Given a fixed line $\mathbf{D}$ in the plane, called the *directrix*, and a fixed point $\mathbf{F}$, called the *focus*, the locus of all points $\mathbf{P}$ such that the distance $\mathbf{PF}$ from $\mathbf{P}$ to $\mathbf{F}$ is proportional to the distance $\mathbf{PD}$ from $\mathbf{P}$ to the directrix, is a conic. Thus there exists a constant ϵ, called the *eccentricity*, such that $\mathbf{PF} = \epsilon\mathbf{PD}$.

PF = ePD

Example 5.12

Let a conic have directrix the x-axis, focus $\mathbf{F}(2, 3)$, and eccentricity $\epsilon = 4$. Let $\mathbf{P}(x, y)$ be a general point on the conic. Then

$$\mathbf{PD} = y, \;\; \mathbf{PF} = \sqrt{(x-2)^2 + (y-3)^2}\ .$$

Hence $\mathbf{PF} = \epsilon\mathbf{PD}$ implies

$$\sqrt{(x-2)^2 + (y-3)^2} = 4y\ ,$$

giving the conic with the equation $x^2 - 15y^2 - 4x - 6y + 13 = 0$ shown in Figure 5.10.

To prove that the focus–directrix construction gives a conic it is sufficient to show that the curve satisfies the general equation (5.4). Suppose the directrix is the line $lx + my + n = 0$, and the focus is $(x_{\mathbf{F}}, y_{\mathbf{F}})$. Then

$$\mathbf{PD} = (lx_{\mathbf{F}} + my_{\mathbf{F}} + n) \Big/ \sqrt{l^2 + m^2}$$

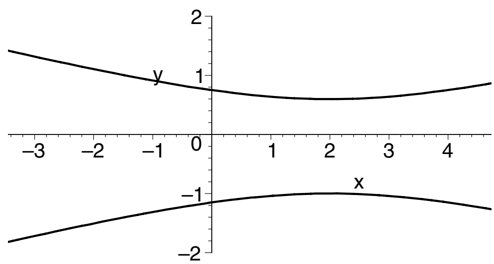


Figure 5.10 Conic $x^2 - 15y^2 - 4x - 6y + 13 = 0$

and

$$\mathbf{PF} = \sqrt{(x - x_{\mathbf{F}})^2 + (y - y_{\mathbf{F}})^2}\ .$$

Hence,

$$\epsilon(lx + my + n) \Big/ \sqrt{l^2 + m^2} = \sqrt{(x - x_{\mathbf{F}})^2 + (y - y_{\mathbf{F}})^2}\ .$$

Squaring both sides and multiplying through by $l^2 + m^2$ yields

$$\epsilon^2(lx + my + n)^2 = (l^2 + m^2)((x - x_{\mathbf{F}})^2 + (y - y_{\mathbf{F}})^2)$$

which is a quadratic equation in x and y of the form (5.4) where $a = \epsilon^2 l^2 - l^2 - m^2$, $b = \epsilon^2 lm$, $c = \epsilon^2 m^2 - \left(l^2 + m^2\right)$, $d = \epsilon^2 nl + x_{\mathbf{F}}(l^2 + m^2)$, $e = \epsilon^2 mn + y_{\mathbf{F}}\left(l^2 + m^2\right)$, $f = \epsilon^2 n^2 - \left(l^2 + m^2\right)\left(x_{\mathbf{F}}^2 + y_{\mathbf{F}}^2\right)$.

The converse, that any non-degenerate conic can be obtained by a focus–directrix construction, can be proved in two steps: (i) computation of the eccentricity of a conic expressed in implicit form, and (ii) computation of the focus and directrix. The first step is proved in Exercise 5.18, while the remainder of the proof can be found in [26].

Exercise 5.18

Let a conic have focus $(x_{\mathbf{F}}, y_{\mathbf{F}})$, eccentricity ϵ, and directrix with equation $x\cos\theta + y\sin\theta - p = 0$. Expand the expression

$$\left((x - x_{\mathbf{F}})^2 + (y - y_{\mathbf{F}})^2\right) = \epsilon^2(x\cos\theta + y\sin\theta - p)^2$$

and compare the coefficients with a scalar multiple of the coefficients of (5.4). Show that

$$\frac{\left(2 - \epsilon^2\right)^2}{1 - \epsilon^2} = \frac{(a + c)^2}{ac - b^2}\ .$$

5.6.1 Classification of Conics

Consider a conic defined by Equation (5.4). If (5.4) is a product of two linear factors, then the conic is a union of two lines and it is said to be a *reducible* conic. Otherwise, the conic is said to be *irreducible*. A condition on the coefficients of (5.4) for the conic to be reducible is determined as follows. Suppose that $a \neq 0$. Then multiply (5.4) through by a and complete the square to give

$$(ax + by + d)^2 - ((b^2 - ac)y^2 + 2(bd - ae)y + (d^2 - af)) = 0 \, . \tag{5.5}$$

Let $A = b^2 - ac$, $B = 2(bd - ae)$, and $C = (d^2 - af)$. Then (5.5) can be written

$$(ax + by + d)^2 - \big(Ay^2 + By + C\big) = 0 \, . \tag{5.6}$$

The expression (5.6) has two linear factors if and only if it can be written as the difference of two squares. Thus $Ay^2 + By + C$ must be a perfect square, which is possible if and only if $B^2 - 4AC = 0$. Hence the condition for the conic to be reducible is

$$B^2 - 4AC = 4(bd - ae)^2 - 4(b^2 - ac)(d^2 - af) = 0 \, .$$

Dividing through by $-4a$, the condition for reducibility can be expressed as the following determinant $\mathit{\Delta}$ which is called the *discriminant* of the conic.

$$\mathit{\Delta} = \begin{vmatrix} a & b & d \\ b & c & e \\ d & e & f \end{vmatrix} = 0 \, .$$

When $\mathit{\Delta} = 0$, $Ay^2 + By + C = A(y + B/2A)^2$ and two cases can be distinguished. (1) When $A = b^2 - ac \geq 0$, (5.6) has two real linear factors and the conic is a pair of lines. (2) When $A = b^2 - ac < 0$, (5.6) has two imaginary linear factors and the conic is an isolated point. The reader is left the exercise of showing that $\mathit{\Delta} = 0$ is also the condition for reducibility in the case when $a = 0$.

Next, suppose that (5.4) has two real linear factors ($a \neq 0$, $b^2 - ac \geq 0$)

$$a(x - \alpha_1 y + \beta_1)(x - \alpha_2 y + \beta_2) = 0 \, .$$

Expanding the brackets and comparing the coefficients of the resulting expression with (5.4) gives $\alpha_1\alpha_2 = c/a$, $\alpha_1 + \alpha_2 = -2b/a$. A simple computation yields that the angle θ between the two lines is given by $\tan\theta = 2\sqrt{b^2 - ac}\big/(a + c)$. It follows that the conic is a pair of perpendicular lines when $\mathit{\Delta} = 0$ and $a + c = 0$, and a pair of parallel lines whenever $\mathit{\Delta} = 0$ and $b^2 - ac = 0$. This concludes the study of the reducible conics.

The irreducible conics are as follows: (1) hyperbolas when $b^2 - ac > 0$, (2) ellipses when $b^2 - ac < 0$, and (3) parabolas when $b^2 - ac = 0$. The distinction can be explained by the conic's behaviour at infinity. Let (X, Y, W) be homogeneous coordinates of a point (x, y), so that $x = X/W$ and $y = Y/W$. Substituting into (5.4) and multiplying through by W^2 yields that the homogeneous coordinates of any point on the conic satisfies the homogeneous equation

$$\mathbf{C}(X, Y, W) = aX^2 + 2bXY + cY^2 + 2dXW + 2eYW + fW^2 = 0\,. \tag{5.7}$$

The points at infinity of the conic are obtained by setting $W = 0$ in (5.7) to give

$$aX^2 + 2bXY + cY^2 = 0\,. \tag{5.8}$$

When $b^2 - ac > 0$, (5.8) can be expressed as two distinct real linear factors $a\,(X + \mu_1 Y)\,(X + \mu_2 Y) = 0$, and it follows that the conic has two distinct real points at infinity, $(\mu_1, -1, 0)$ and $(\mu_2, -1, 0)$. Likewise, when $b^2 - ac < 0$, (5.8) can be expressed as two complex conjugate linear factors which give rise to two complex conjugate points at infinity. When $b^2 - ac = 0$, (5.8) can be expressed as a perfect square $a\,(X + \mu_1 Y)^2 = 0$, and hence the conic has a repeated real point at infinity, $(\mu_1, -1, 0)$.

The tangent lines to a curve at points at infinity are the *asymptotes* of the curve. Therefore, when $b^2 - ac > 0$ the irreducible conic has two real asymptotes and the curve is a hyperbola, and when $b^2 - ac < 0$, there are no real asymptotes and the conic is an ellipse. When $b^2 - ac = 0$, the asymptote is the line at infinity $W = 0$ and the conic is a parabola.

Definition 5.13

The *centre* of a conic $\mathbf{C}(x, y) = 0$ is the point (x, y) satisfying

$$\frac{\partial \mathbf{C}}{\partial x}(x, y) = 2ax + 2by + 2d = 0, \quad \frac{\partial \mathbf{C}}{\partial y}(x, y) = 2bx + 2cy + 2e = 0\,.$$

If $b^2 - ac \neq 0$ then the conic has centre

$$(x, y) = \left((be - cd) \,/\, \left(ac - b^2\right), (bd - ae) \,/\, \left(ac - b^2\right)\right)\,,$$

otherwise there is no centre. Conics with a centre are called *central* conics. Of the irreducible conics, the ellipse and hyperbola are central conics but the parabola is not.

In addition to the implicit form (5.4), conics have a parametric form

$$(x\,(t)\,, y\,(t)) = \left(\frac{a_0 + a_1 t + a_2 t^2}{c_0 + c_1 t + c_2 t^2}, \frac{b_0 + b_1 t + b_2 t^2}{c_0 + c_1 t + c_2 t^2}\right)\,, \tag{5.9}$$

where the coefficients c_0, c_1, c_2 are not all zero. Conics can also be defined parametrically by other functions such as trigonometric and hyperbolic functions.

Sections 5.6.4 and 5.6.5 show how to convert a non-degenerate conic from implicit to parametric form and vice versa. In particular, Theorem 5.26 shows that a parametric curve of the form (5.9) can be expressed in the implicit form (5.4), and is therefore a conic.

Recall that the irreducible types hyperbola/parabola/ellipse are distinguished by the fact that they have two real/one real/two complex conjugate points at infinity. For a parametric conic (5.9), the points at infinity occur at parameter values for which the denominator of the coordinate functions vanishes, that is, when $c_2t^2 + c_1t + c_0 = 0$. When $c_1^2 - 4c_0c_2 > 0$ the denominator vanishes at two real values of t which give rise to two real points at infinity. The conic is therefore a hyperbola. Similarly, it can be shown that the conic is an ellipse when $c_1^2 - 4c_0c_2 < 0$, and a parabola when $c_1^2 - 4c_0c_2 = 0$. In particular, if a conic is parametrized by quadratic polynomials, then $c_1 = c_2 = 0$, $c_0 = 1$, and the conic is a parabola.

Summary. A conic is either *irreducible* when $\Delta \neq 0$, or *reducible* when $\Delta = 0$. The irreducible conics have three distinct types, namely, ellipse, parabola, and hyperbola. Reducible conics are a union of two lines (real or imaginary) which may be distinct and non-parallel, distinct and parallel, or coincident. The types of conic are summarized in Table 5.1.

Table 5.1 Summary of conic types

$b^2 - ac$	Δ	**Central**	**Conic type**	$c_1^2 - 4c_0c_2$
> 0	$\neq 0$	Yes	hyperbola	> 0
< 0	$\neq 0$	Yes	ellipse	< 0
$= 0$	$\neq 0$	No	parabola	$= 0$
> 0	$= 0$	Yes	two real distinct intersecting lines	
< 0	$= 0$	Yes	two complex conjugate lines intersecting in a real point	
$= 0$	$= 0$	No	two real distinct parallel lines	
$= 0$	$= 0$	No	two real coincident lines	

Example 5.14

The following examples show how to determine whether a conic is irreducible or reducible, and whether an irreducible conic is an ellipse, a hyperbola, or a parabola.

1. Consider the conic given by $x^2 + 2xy - 3y^2 + 4x - 5 = 0$. Then $a = 1$, $b = 1$, $c = -3$, $d = 2$, $e = 0$, $f = -5$,

$$\Delta = \begin{vmatrix} 1 & 1 & 2 \\ 1 & -3 & 0 \\ 2 & 0 & -5 \end{vmatrix} = 32 \, .$$

$a x^2$
$2b xy$
$c y^2$
$2d x$
$2e y$
f

$$D \to \begin{vmatrix} a & b & d \\ b & c & e \\ d & e & f \end{vmatrix}$$

Since $\Delta \neq 0$ the conic is irreducible. Further, $b^2 - ac = 4 > 0$, hence the conic is a hyperbola.

2. Consider the conic given by $-2x^2 + xy - x - y + 3 = 0$. Then $a = -2$, $b = 1/2$, $c = 0$, $d = -1/2$, $e = -1/2$, $f = 3$,

$$\Delta = \begin{vmatrix} -2 & 1/2 & -1/2 \\ 1/2 & 0 & -1/2 \\ -1/2 & -1/2 & 3 \end{vmatrix} = 0 \, .$$

Since $\Delta = 0$ the conic is reducible. Completing the square of the conic yields

$$-\frac{1}{2}\left(-2x + \frac{1}{2}y - \frac{1}{2}\right)^2 + \left(-\frac{5}{4}y + \frac{25}{8} + \frac{1}{8}y^2\right) = 0 \, .$$

Factorize the quadratic in y to give the difference of two squares

$$-\frac{1}{2}\left(-2x + \frac{1}{2}y - \frac{1}{2}\right)^2 + \frac{1}{8}(y-5)^2 = 0 \, .$$

Factorizing,

$$(x-1)(-2x+y-3) = 0 \, .$$

Hence, the conic is the union of the two lines $x - 1 = 0$ and $2x - y + 3 = 0$.

EXERCISES

5.19. For each of the conics below determine whether the conic is irreducible or reducible. If it is irreducible then determine whether it is an ellipse, a hyperbola, or a parabola. If the conic is reducible then determine the linear factors.

(a) $4x^2 - 3xy + y^2 - x + 2y + 7 = 0$.

(b) $-2x^2 + y^2 + 3x - 4y + 1 = 0$.

(c) $3x^2 - 5xy - x - 2y^2 + 9y - 4 = 0$.

(d) $x^2 - 3xy + 5y^2 - 2x + 6 = 0$.

(e) $2x^2 + 2xy - 5x - 3y + 3 = 0$.

(f) $2x^2 - 2y^2 + 3x + 4y + 7 = 0$.

(g) $2xy + 3y = 5$.

(h) $3x^2 - xy - 2y^2 + 6x + 4y = 0$.

(i) $2x^2 + 2xy - 3x - 3y + 1 = 0$.

(j) $2x^2 - 4xy + 2y^2 - 9 = 0$.

5.20. Let $\mathbf{C}(x, y) = 0$ be a central conic. Any line through the centre is called a *diameter*. The centre is the midpoint of the two points of intersection of $\mathbf{C}$ with any diameter. Verify this fact for

(a) the hyperbola $4x^2 - 9y^2 = 16$, and

(b) the ellipse $4x^2 + 9y^2 + x - 6y = 0$.

5.21. There are conics which have no real points. For example, $x^2 + y^2 = -1$. Determine others.

5.6.2 Conics in Standard Form

It can be shown that hyperbolas and ellipses have two lines of reflectional symmetry, and that a parabola has one. A conic for which the lines of symmetry coincide with the coordinate axes is said to be in *standard form*. It will be shown that any irreducible conic can be obtained by applying a composite transformation consisting of a rotation and a translation to a conic in standard form. Conversely, any conic can be obtained by applying a transformation to a conic in standard form. The implicit and parametric standard forms of the irreducible conics are given in Table 5.2.

Recall that in homogeneous coordinates, the conic (5.4) is given by (5.7). Let $\mathbf{x} = (X, Y, W)$ then (5.7) can be expressed in the matrix form

$$\mathbf{C}(X, Y, W) = \mathbf{x}\mathsf{M}\mathbf{x}^T = \begin{pmatrix} X & Y & W \end{pmatrix} \begin{pmatrix} a & b & d \\ b & c & e \\ d & e & f \end{pmatrix} \begin{pmatrix} X \\ Y \\ W \end{pmatrix}.$$

The discriminant of $\mathbf{C}$ is denoted $\Delta_{\mathbf{C}} = \det(\mathsf{M})$.

Table 5.2 Standard forms for the irreducible conics

Conic	**Implicit forms**	**Parametric forms**
Hyperbola	$x^2/a^2 - y^2/b^2 = 1$	$\left(\frac{a(b^2+a^2t^2)}{a^2t^2-b^2}, \frac{2ab^2t}{a^2t^2-b^2}\right)$, $t \in \mathbb{R}$, $t \neq \frac{b}{a}$; $\pm(a\cosh\theta, b\sinh\theta)$, $\theta \in \mathbb{R}$; $(a\sec\theta, b\tan\theta)$, $\theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, and $\theta \in \left(\frac{\pi}{2}, \frac{3\pi}{2}\right)$.
Parabola	$y = mx^2$ $x = my^2$	(t, mt^2) , $t \in \mathbb{R}$; (mt^2, t) , $t \in \mathbb{R}$.
Ellipse	$x^2/a^2 + y^2/b^2 = 1$	$\left(\frac{a(1-t^2)}{1+t^2}, \frac{2bt}{1+t^2}\right)$, $t \in \mathbb{R}$; $(a\cos\theta, b\sin\theta)$, $\theta \in [0, 2\pi]$.

Theorem 5.15

Let $\hat{\mathbf{C}}$ be the image of the conic $\mathbf{C} = \mathbf{x}\mathsf{M}\mathbf{x}^T$ following the application of a non-singular planar transformation with transformation matrix A. Then

$$\Delta_{\hat{\mathbf{C}}} = \det(\mathsf{A})^2 \Delta_{\mathbf{C}} \; .$$

So a non-singular transformation does not affect the irreducibility of a conic.

Proof

Let the conic be $\mathbf{C} = \mathbf{x}\mathsf{M}\mathbf{x}^T$. Then the transformation $\mathbf{x} = \mathbf{y}\mathsf{A}$ yields a conic $\hat{\mathbf{C}} = (\mathbf{y}\mathsf{A})\mathsf{M}(\mathbf{y}\mathsf{A})^T = \mathbf{y}\mathsf{AMA}^T\mathbf{y}^T = \mathbf{y}\hat{\mathsf{M}}\mathbf{y}^T$, where $\hat{\mathsf{M}} = \mathsf{AMA}^T$. Hence

$$\Delta_{\hat{\mathbf{C}}} = \det(\mathsf{AMA}^T) = \det(\mathsf{A})\det(\mathsf{M})\det(\mathsf{A}^T) = \det(\mathsf{A})^2\det(\mathsf{M}) = \det(\mathsf{A})^2\Delta_{\mathbf{C}}.$$

Since A is non-singular, $\det(\mathsf{A}) \neq 0$. Therefore $\Delta_{\hat{\mathbf{C}}} = 0$ if and only if $\Delta_{\mathbf{C}} = 0$. Hence $\mathbf{C}$ is irreducible if and only if $\hat{\mathbf{C}}$ is irreducible.

$\square$

The distinctions of hyperbola, parabola, and ellipse apply with respect to a particular Cartesian coordinate system. The effect on a conic of an application of an orthogonal change of coordinates (see Exercise 2.21) is expressed by the following theorem.

Theorem 5.16

Let $\hat{\mathbf{C}}$ be the image of the conic $\mathbf{C}$ (given by (5.7)) following the application of an orthogonal change of coordinates. Then

$$\begin{aligned} b^2 - ac &= \hat{b}^2 - \hat{a}\hat{c} \,, \\ a + c &= \hat{a} + \hat{c} \,, \end{aligned}$$

where $\hat{a}$, $\hat{b}$ and $\hat{c}$ denote the corresponding coefficients of $\hat{\mathbf{C}}$.

Proof

Let the orthogonal change of coordinates be $X = \hat{X}\cos\theta - \hat{Y}\sin\theta + g\hat{W}$, $Y = \hat{X}\sin\theta + \hat{Y}\cos\theta + h\hat{W}$, $W = \hat{W}$ (expressed in homogeneous coordinates). Then substituting for X and Y in (5.7) yields

$$\hat{a}\hat{X}^2 + 2\hat{b}\hat{X}\hat{Y} + \hat{c}\hat{Y}^2 + \hat{d}\hat{X}\hat{W} + \hat{e}\hat{Y}\hat{W} + \hat{f}\hat{W}^2 = 0$$

where

$$\begin{aligned} \hat{a} &= a\cos^2\theta + 2b\cos\theta\sin\theta + c\sin^2\theta \,, \\ \hat{b} &= b\left(\cos^2\theta - \sin^2\theta\right) + (c-a)\sin\theta\cos\theta \,, \\ \hat{c} &= a\sin^2\theta - 2b\cos\theta\sin\theta + c\cos^2\theta \,, \\ \hat{d} &= bg\sin\theta + e\sin\theta + bh\cos\theta + d\cos\theta + ag\cos\theta + ch\sin\theta \,, \\ \hat{e} &= ch\cos\theta - ag\sin\theta + bg\cos\theta - bh\sin\theta - d\sin\theta + e\cos\theta \,, \text{ and} \\ \hat{f} &= ag^2 + ch^2 + 2dg + 2bgh + f + 2eh \,. \end{aligned}$$

Then

$$\begin{aligned} \hat{b}^2 - \hat{a}\hat{c} &= \left(b\left(\cos^2\theta - \sin^2\theta\right) + (c-a)\sin\theta\cos\theta\right)^2 \\ &\quad - \left(a\cos^2\theta + 2b\cos\theta\sin\theta + c\sin^2\theta\right) \\ &\quad \times \left(a\sin^2\theta - 2b\cos\theta\sin\theta + c\cos^2\theta\right) \\ &= b^2 - ac \,, \end{aligned}$$

and

$$\begin{aligned} \hat{a} + \hat{c} &= \left(a\cos^2\theta + 2b\cos\theta\sin\theta + c\sin^2\theta\right) \\ &\quad + \left(a\sin^2\theta - 2b\cos\theta\sin\theta + c\cos^2\theta\right) \\ &= a + c \,. \end{aligned}$$

□

Let A be the transformation matrix of an orthogonal change of coordinates. Then, by Exercise 2.21, $\det(\mathsf{A}) = 1$. Since $\hat{b}^2 - \hat{a}\hat{c} = b^2 - ac$ and $\Delta_{\hat{\mathbf{C}}} = \det(\mathsf{A})^2 \Delta_{\mathbf{C}} = \Delta_{\mathbf{C}}$ the type of conic is unaffected by a change of coordinates. If (5.4) is multiplied through by a constant μ, then the quantities $a+c$, $b^2 - ac$, and Δ become $\mu\,(a+c)$, $\mu^2\left(b^2-ac\right)$, and $\mu^3\Delta$. Thus the ratios $(a+c) : \left(b^2 - ac\right)^{1/2} : \Delta^{1/3}$ are *absolute invariants*. A conic expressed in any rectangular Cartesian coordinate system has the same absolute invariants.

Theorem 5.17

An irreducible conic can be mapped to a conic in standard form by applying an orthogonal change of coordinates.

Proof

Let the conic be given by (5.7). First, apply a rotation $X = \hat{X}\cos\theta - \hat{Y}\sin\theta$, and $Y = \hat{X}\sin\theta + \hat{Y}\cos\theta$. Then (5.7) has the form

$$\hat{a}\hat{X}^2 + 2\hat{b}\hat{X}\hat{Y} + \hat{c}\hat{Y}^2 + 2\hat{d}\hat{X}\hat{W} + 2\hat{e}\hat{Y}\hat{W} + \hat{f}\hat{W}^2 = 0\ , \tag{5.10}$$

where the coefficients are given by the expressions in Theorem 5.16, but with g and h set equal to zero. The $\hat{X}\hat{Y}$ term can be eliminated by choosing the angle θ so that the coefficient $\hat{b}$ vanishes: the required angle satisfies $\tan 2\theta = 2b/(a-c)$ if $a \neq c$, and $\theta = \pi/4$ or $3\pi/4$ if $a = c$. Provided $\hat{a} \neq 0$ and $\hat{c} \neq 0$, then (5.10) has the form

$$\hat{a}\left(\hat{X} + \frac{\hat{d}}{\hat{a}}\hat{W}\right)^2 + \hat{c}\left(\hat{Y} + \frac{\hat{e}}{\hat{c}}\hat{W}\right)^2 + \left(\hat{f} - \frac{\hat{d}^2}{\hat{a}} - \frac{\hat{e}^2}{\hat{c}}\right)\hat{W}^2 = 0\ .$$

Applying the translation $\mathsf{T}\left(-\frac{\hat{d}}{\hat{a}}, -\frac{\hat{e}}{\hat{c}}\right)$, yields a standard form for the hyperbola or ellipse given by

$$\hat{a}\hat{X}^2 + \hat{c}\hat{Y}^2 + \left(\hat{f} - \frac{\hat{d}^2}{\hat{a}} - \frac{\hat{e}^2}{\hat{c}}\right)\hat{W}^2 = 0\ .$$

If $\hat{a} = 0$, $\hat{c} \neq 0$, then (5.10) has the form

$$\hat{c}\left(\hat{Y} + \frac{\hat{e}}{\hat{c}}\hat{W}\right)^2 + 2\hat{d}\hat{X}\hat{W} + \hat{f}\hat{W}^2 = 0\ .$$

Applying the translation $\mathsf{T}\left(-\frac{\hat{f}}{2\hat{d}}, -\frac{\hat{e}}{\hat{c}}\right)$ gives

$$\hat{c}\hat{Y}^2 + 2\hat{d}\hat{X}\hat{W}\ ,$$

a standard form for the parabola. Similarly, a standard form for the parabola is obtained when $\hat{a} \neq 0$, $\hat{c} = 0$. The case $\hat{a} = 0$, $\hat{c} = 0$ is not considered since the conic is reducible.

□

Example 5.18

To determine the standard form for the ellipse

$$2x^2 - 2\sqrt{3}xy + 4y^2 + 5x + 6y - 1 = 0 \,.$$

Then $a = 2$, $b = -\sqrt{3}$, $c = 4$, and the required rotation angle is given by $\tan 2\theta = \sqrt{3}$, yielding $\theta = \pi/6$. Then $\cos(\pi/6) = \sqrt{3}/2$, $\sin(\pi/6) = \frac{1}{2}$, and the required rotation is $X = \frac{\sqrt{3}}{2}\hat{X} - \frac{1}{2}\hat{Y}$, and $Y = \frac{1}{2}\hat{X} + \frac{\sqrt{3}}{2}\hat{Y}$, giving the conic

$$\begin{aligned} &2\left(\frac{\sqrt{3}}{2}\hat{X} - \frac{1}{2}\hat{Y}\right)^2 - 2\sqrt{3}\left(\frac{\sqrt{3}}{2}\hat{X} - \frac{1}{2}\hat{Y}\right)\left(\frac{1}{2}\hat{X} + \frac{\sqrt{3}}{2}\hat{Y}\right) \\ &+ 4\left(\frac{1}{2}\hat{X} + \frac{\sqrt{3}}{2}\hat{Y}\right)^2 + 5\left(\frac{\sqrt{3}}{2}\hat{X} - \frac{1}{2}\hat{Y}\right)\hat{W} \\ &+ 6\left(\frac{1}{2}\hat{X} + \frac{\sqrt{3}}{2}\hat{Y}\right)\hat{W} - \hat{W}^2 = 0 \,. \end{aligned}$$

Simplifying yields

$$\hat{X}^2 + 5\hat{Y}^2 + \left(\frac{5}{2}\sqrt{3} + 3\right)\hat{X}\hat{W} + \left(3\sqrt{3} - \frac{5}{2}\right)\hat{Y}\hat{W} - \hat{W}^2 = 0 \,.$$

Then, completing the squares in $\hat{X}$ and $\hat{Y}$,

$$\begin{aligned} &\left(\hat{X} + \left(\frac{5}{4}\sqrt{3} + \frac{3}{2}\right)\hat{W}\right)^2 + 5\left(\hat{Y} + \left(\frac{3}{10}\sqrt{3} - \frac{1}{4}\right)\right)^2 \\ &- \left(1 + \left(\frac{5}{4}\sqrt{3} + \frac{3}{2}\right)^2 + \left(\frac{3}{10}\sqrt{3} - \frac{1}{4}\right)^2\right)\hat{W}^2 = 0 \,, \end{aligned}$$

and making the translation $\bar{X} = \hat{X} + \left(\frac{5}{4}\sqrt{3} + \frac{3}{2}\right)\hat{W}$, $\bar{Y} = \hat{Y} + \left(\frac{3}{10}\sqrt{3} - \frac{1}{4}\right)\hat{W}$, $\bar{W} = \hat{W}$ gives

$$\bar{X}^2 + 5\bar{Y}^2 - \left(\frac{827}{100} + \frac{18}{5}\sqrt{3}\right)\bar{W}^2 = 0 \,.$$

In Cartesian coordinates, the standard form of the conic is $x^2 + 5y^2 = \left(\frac{827}{100} + \frac{18}{5}\sqrt{3}\right)$. Figure 5.11 shows the original conic and the computed standard form conic.

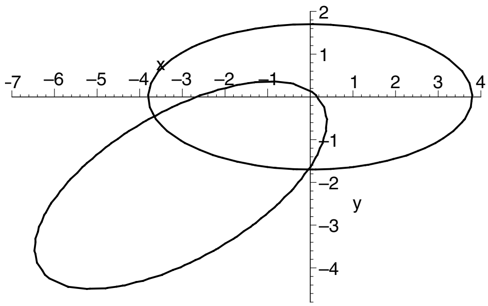


Figure 5.11

EXERCISES

5.22. Determine the standard form of the following conics:

(a) $13x^2 - 10xy + 13y^2 - 12\sqrt{2}x + 60\sqrt{2}y + 72 = 0$.

(b) $6x^2 + 12xy + 6y^2 - 35\sqrt{2}x - 37\sqrt{2}y + 118 = 0$.

(c) $11x^2 - 6x\sqrt{3}y - 6x\sqrt{3} + y^2 + 2y - 63 = 0$.

5.23. Determine the absolute invariants of the standard forms. Compute the absolute invariants for each of the conics of the previous exercise, and verify that the computed standard form has the same invariants.

5.24. Use Exercise 5.18 to show that a conic with eccentricity $\epsilon > 0$ is a hyperbola if $\epsilon > 1$, an ellipse if $\epsilon < 1$, or a parabola if $\epsilon = 1$.

5.25. Show that a translation leaves the values of a, b, c unaltered. These quantities are called *translational invariants*.

5.6.3 Intersections of a Conic with a Line

The points of intersection of a conic and a line are found by a process of elimination of variables to give a quadratic polynomial equation in one of the variables. The equation is solved, and backward substitution of the solutions is used to determine the points of intersection. The derivation of the quadratic depends on whether the conic and the line are given in implicit or in parametric form. The procedure is explained by means of examples.

Example 5.19

To find the intersections of the conic $x^2 - 3xy + y^2 + 4x - 2 = 0$ and the line parametrized by $(x(t), y(t)) = (2t + 1, t - 1)$, shown in Figure 5.12, substitute $x = 2t + 1$ and $y = t - 1$ into the conic equation to give

$$(2t + 1)^2 + (t - 1)^2 - 3(2t + 1)(t - 1) + 4(2t + 1) - 2 = -t^2 + 13t + 7 = 0 .$$

The solutions are approximately $t = -0.518$ and $t = 13.518$. Substituting the values of t into the parametric equation $(x(t), y(t)) = (2t + 1, t - 1)$ gives the two points of intersection $(-0.036, -1.518)$ and $(28.036, 12.518)$.

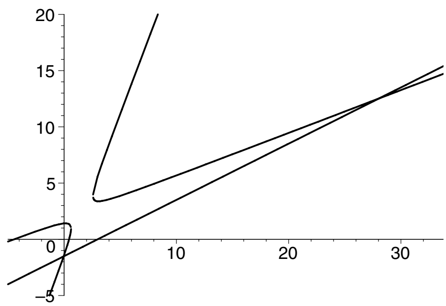


Figure 5.12 Intersection of conic $x^2 - 3xy + y^2 + 4x - 2 = 0$ and line $(2t + 1, t - 1)$

Example 5.20

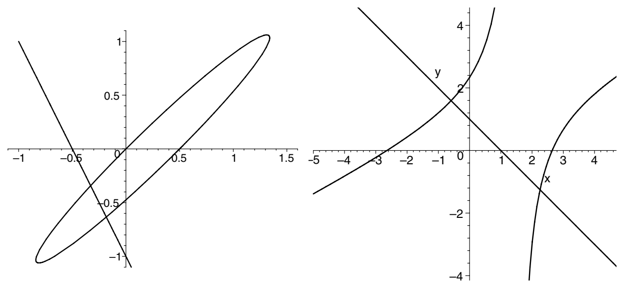
To find the intersection of the conic $(x(t), y(t)) = \left(\frac{3t+t^2}{1+2t^2}, \frac{3t}{1+2t^2}\right)$ with the line $2x + y + 1 = 0$, shown in Figure 5.13(a), substitute $x = \frac{3t+t^2}{1+2t^2}$ and $y = \frac{3t}{1+2t^2}$ into the equation of the line to give

$$2\left(\frac{3t + t^2}{1 + 2t^2}\right) + \left(\frac{3t}{1 + 2t^2}\right) + 1 = \frac{4t^2 + 9t + 1}{1 + 2t^2} = 0 .$$

The solutions of $4t^2 + 9t + 1 = 0$ are approximately $t = -2.133$ and $t = -0.117$. Substituting for t in $\left(\frac{3t+t^2}{1+2t^2}, \frac{3t}{1+2t^2}\right)$ yields the two points of intersection $(-0.183, -0.634)$ and $(-0.329, -0.342)$.

Example 5.21

To determine the intersection of the line $x + y - 1 = 0$ and the conic



(a) (b)

Figure 5.13 (a) Intersection of conic $\left(\frac{3t+t^2}{1+2t^2}, \frac{3t}{1+2t^2}\right)$ and line $2x+y+1=0$, and (b) intersection of conic $x^2-2xy+3y-7=0$ and line $x+y-1=0$

$x^2 - 2xy + 3y - 7 = 0$, shown in Figure 5.13(b), substitute $y = 1 - x$ into the conic equation to give

$$x^2 - 2x(1-x) + 3(1-x) - 7 = 3x^2 - 5x - 4 = 0 \ .$$

Solving yields $x = -0.591$ and $x = 2.257$. Substituting the solutions into $y = 1 - x$ yields the points $(-0.591, 1.591)$ and $(2.257, -1.257)$.

If both the line and the conic are expressed in parametric form, then the line is converted to implicit form, and the method of Example 5.20 is applied.

EXERCISES

5.26. Find the points of intersection of the following conics and lines:

(a) conic $9x^2 - xy + y^2 - 4x + 2y + 1 = 0$, line $(x(t), y(t)) = (2t - 3, -3t + 4)$.

(b) line $x + 3y - 6 = 0$, conic $3x^2 - 2xy + y^2 - 5x + 6y - 16 = 0$.

(c) line $-2x + 5y + 7 = 0$, conic $(x(t), y(t)) = (3t^2 - 4t + 1, 2t^2 - 9t)$.

(d) line $(t+1, t-1)$, conic $x^2 + 2xy + x - y - 1 = 0$.

(e) line $-3x - 2y + 4 = 0$, conic $(t^2 + 1, t - 1)$.

5.27. The conic segment $(x(t), y(t)) = (3t^2 - 4t - 1, 2t^2 - 9t + 10)$, $t \in [-1, 4]$ is to be clipped by the rectangle with bottom left corner at $(0, 0)$ and upper right corner $(20, 20)$ as shown in Figure 5.14. The clipping operation removes the parts of the conic contained outside the rectangle. Determine the parameter values where the conic intersects each side of the rectangle. For each conic segment inside the rectangle, determine the parameter interval on which it is defined. For any parametrized conic, what is the maximum number of segments that arise following a rectangular clipping operation?

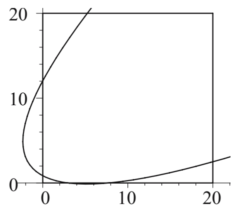


Figure 5.14

5.6.4 Parametrization of an Irreducible Conic

An irreducible conic $\mathbf{C}(x, y) = 0$ can be parametrized by performing the following steps.

1. Determine a point $\mathbf{P}(p_1, p_2)$ on the conic.
2. Consider the family of lines $y = (x-p_1)t+p_2$, parametrized by t, consisting of all lines in the plane through $\mathbf{P}$. Each parameter value t corresponds to a line in the family.
3. Determine the points of intersection of the line $y = (x - p_1)t + p_2$ and the conic as follows. Substitute $y = (x - p_1)t + p_2$ in $\mathbf{C}(x, y) = 0$ to give a quadratic polynomial (dependent on t) $q_t(x) = 0$. The roots of $q_t(x) = 0$ are the x–coordinates of the intersection points. Since $\mathbf{P}(p_1, p_2)$ is known to be an intersection point of the conic and every line in the family, it follows that $x = p_1$ is a root of $q_t(x)$. Hence $x - p_1$ is a factor of $q_t(x)$.
4. Factorise $q_t(x)$ as $(x - p_1)\,(\beta(t)x - \alpha(t))$ for some quadratic polynomials $\alpha(t)$ and $\beta(t)$. Then the second root of $q_t(x) = 0$ is $x = \alpha(t)/\beta(t)$, giving the x-coordinate of the other point of intersection as a function of t: $x = x(t)$.
5. Substitute $x = \alpha(t)/\beta(t)$ in $y = (x - p_1)t + p_2$ to give $y = y(t)$.
6. It follows that $(x(t), y(t))$ parametrizes the conic, since every point $\mathbf{Q}$ on the conic, distinct from $\mathbf{P}$, is the intersection of the conic and the line $\overline{\mathbf{PQ}}$ through $\mathbf{P}$.

Different choices of the point $\mathbf{P}$ will give rise to alternative parametrizations of the conic.

Example 5.22

Find a parametrization of the hyperbola $-2x^2 - 5xy + 4y^2 + x - 5y + 15 = 0$ by considering lines through the point $(2, 3)$.

It is easily checked that $(2, 3)$ is a point on the conic. The family of lines through $(2, 3)$ is given by $y = t\,(x - 2) + 3$. Substituting for y in the equation of the conic gives

$$-2x^2 - 5x\,(t\,(x-2)+3) + 4\,(t\,(x-2)+3)^2 + x - 5\,(t\,(x-2)+3) + 15 = 0\ ,$$

which factorizes as

$$(x-2)\left(-2x - 5tx + 4t^2x - 8t^2 + 19t - 18\right) = 0\ .$$

Solving for x yields $x = 2$ and

$$x = \frac{18 - 19t + 8t^2}{-2 - 5t + 4t^2}\ . \tag{5.11}$$

The solution $x = 2$ corresponds to the known intersection $(2, 3)$. Using (5.11) to substitute for x in $y = t\,(x - 2) + 3$ gives

$$y = t\left(\frac{18 - 19t + 8t^2}{-2 - 5t + 4t^2} - 2\right) + 3 = \frac{-6 + 7t + 3t^2}{-2 - 5t + 4t^2}\ .$$

It follows that

$$(x(t), y(t)) = \left(\frac{18 - 19t + 8t^2}{-2 - 5t + 4t^2}, \frac{-6 + 7t + 3t^2}{-2 - 5t + 4t^2}\right)$$

is a parametrization of the conic. The values of t for which the denominator vanishes are the solutions of $4t^2 - 5t - 2 = 0$, that is, $t = -0.319$ and $t = 1.569$. Therefore the curve is defined on the parameter intervals $(-\infty, -0.319)$, $(-0.319, 1.569)$, and $(1.569, \infty)$. Each interval corresponds to a branch or a part of a branch of the conic.

Example 5.23

Find a parametrization of the conic $\mathbf{C}(x, y) = x^2 - 2xy + 4y^2 + 2x + y + 1 = 0$.

First determine a point on the conic. One way to do this is to find an intersection of the conic with one of the axes. If this fails one can try intersecting with other lines parallel to one of the axes. Set $y = 0$ in $\mathbf{C}(x, y) = 0$. Then $x^2 + 2x + 1 = 0$. Thus $x = -1$. Hence $(-1, 0)$ is a point on the conic. Next consider the family of lines $y = (x+1)t$ through $(-1,0)$. Substituting $y = (x+1)t$ in $\mathbf{C}(x,y) = 0$, gives

$$x^2 - 2x(x+1)t + 4((x+1)t)^2 + 2x + (x+1)t + 1 = 0 ,$$

and factorizing gives

$$(x+1)\left(-2xt + 4t^2x + x + 1 + 4t^2 + t\right) = 0 .$$

Each line in the family intersects the conic in two points: $(-1,0)$ and one other. Setting the second factor equal to zero, and solving for x, gives the x-coordinate of the unknown intersection point

$$x = -\frac{4t^2+t+1}{4t^2-2t+1} .$$

The y-coordinate is obtained by substituting for x in $y = (x+1)t$ giving

$$y = \left(\left(-\frac{4t^2+t+1}{4t^2-2t+1}\right)+1\right)t = -\frac{3t^2}{4t^2-2t+1} .$$

Thus the conic is defined parametrically by

$$(x(t), y(t)) = \left(-\frac{4t^2+t+1}{4t^2-2t+1}, -\frac{3t^2}{4t^2-2t+1}\right) .$$

The denominator $4t^2 - 2t + 1$ does not vanish for real values of t, and hence the parametrization is defined for all t.

EXERCISES

5.28. Determine another parametrization of the conic $x^2 - 2xy + 4y^2 + 2x + y + 1 = 0$ of Example 5.23 by considering lines through the point $(-2,-1)$.

5.29. Convert the following conics from implicit to parametric form:

(a) $x^2 + 2y^2 - 2xy + 2y = 0$; consider lines through $(0,0)$.

(b) $x^2 - 2xy + 5y^2 - 2x + 3y + 1 = 0$.

(c) $x^2 + 2xy - y^2 - 1 = 0$.

(d) $2x^2 - y^2 + 4x - 2y = 0$.

5.6.5 Converting from Parametric Form to Implicit Form

Conics defined by polynomial coordinate functions are easily converted to implicit form.

Example 5.24

Consider the conic defined by $x = 2t^2 - 3t$ and $y = t^2 + t - 2$. Add scalar multiples of the equations to eliminate the quadratic terms

$$\begin{aligned} x - 2y &= \left(2t^2 - 3t\right) - 2\left(t^2 + t - 2\right) \\ &= -5t + 4 \ . \end{aligned}$$

Solving for t in terms of x and y gives $t = (-x + 2y + 4)/\,5$. Substituting for t in $x = 2t^2 - 3t$ (or alternatively, in $y = t^2 + t - 2$) gives

$$x = 2\left(\frac{-x + 2y + 4}{5}\right)^2 - 3\left(\frac{-x + 2y + 4}{5}\right) \ .$$

Expanding and simplifying gives an implicit equation for the conic

$$x^2 - 4xy - 13x + 4y^2 + y - 14 = 0 \ .$$

In general, a conic is parametrized by rational functions and the approach indicated in Example 5.24 is tedious. A more general method follows from the following result.

Theorem 5.25

A necessary and sufficient condition that two quadratics

$$a_0 + a_1 t + a_2 t^2 = 0, \text{ and} \tag{5.12}$$
$$b_0 + b_1 t + b_2 t^2 = 0, \tag{5.13}$$

have a common solution is

$$\begin{vmatrix} a_0 b_2 - a_2 b_0 & a_1 b_2 - a_2 b_1 \\ a_0 b_1 - a_1 b_0 & a_0 b_2 - a_2 b_0 \end{vmatrix} = 0 \ . \tag{5.14}$$

Proof

Suppose (5.12) and (5.13) have a common solution. Then $b_2 \times (5.12) - a_2 \times (5.13)$ yields

$$(a_0 b_2 - a_2 b_0) + (a_1 b_2 - a_2 b_1)\, t = 0 \ . \tag{5.15}$$

Similarly, $b_0 \times (5.12) - a_0 \times (5.13)$ yields

$$((a_1 b_0 - a_0 b_1) + (a_2 b_0 - a_0 b_2)\, t)\, t = 0 \, . \tag{5.16}$$

Eliminating t from (5.15) and (5.16) gives

$$(a_0 b_2 - a_2 b_0)\,(a_2 b_0 - a_0 b_2) - (a_1 b_0 - a_0 b_1)\,(a_1 b_2 - a_2 b_1) = 0 \, ,$$

and hence (5.14). The proof of the converse is left as an exercise for the reader.

□

Theorem 5.26

The conic with parametrization

$$x = \frac{a_0 + a_1 t + a_2 t^2}{c_0 + c_1 t + c_2 t^2} \, , \; y = \frac{b_0 + b_1 t + b_2 t^2}{c_0 + c_1 t + c_2 t^2} \tag{5.17}$$

has an implicit equation of the form

$$(A_1 x + B_1 y + C_1)^2 - (A_0 x + B_0 y + C_0)(A_2 x + B_2 y + C_2) = 0$$

where the coefficients A_i, B_i, C_i are the signed 2×2 minors of the matrix

$$\mathsf{Q} = \begin{pmatrix} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \\ c_0 & c_1 & c_2 \end{pmatrix} \, .$$

Proof

For all t at which the conic is defined, (5.17) can be multiplied through by the denominator to give

$$\begin{aligned} a_0 - c_0 x + (-c_1 x + a_1)\, t + (a_2 - c_2 x)\, t^2 &= 0 \, , \text{ and} && (5.18) \\ b_0 - c_0 y + (-c_1 y + b_1)\, t + (b_2 - c_2 y)\, t^2 &= 0 \, . && (5.19) \end{aligned}$$

Applying Theorem 5.25 to (5.18) and (5.19) gives the necessary and sufficient condition $D_1^2 - D_2 D_3 = 0$ where

$$\begin{aligned} D_1 &= (b_0 c_2 - b_2 c_0)\, x + (a_2 c_0 - a_0 c_2)\, y + a_0 b_2 - a_2 b_0 \, , \\ D_2 &= (b_1 c_2 - b_2 c_1)\, x + (a_2 c_1 - a_1 c_2)\, y + a_1 b_2 - a_2 b_1 \, , \text{ and} \\ D_3 &= (b_0 c_1 - b_1 c_0)\, x + (a_1 c_0 - a_0 c_1)\, y + a_0 b_1 - a_1 b_0 \, . \end{aligned}$$

The proof is now complete since every point (x, y) of the conic satisfies $D_1^2 - D_2 D_3 = 0$, a quadratic polynomial in x and y of the form

$$(A_1 x + B_1 y + C_1)^2 - (A_0 x + B_0 y + C_0)(A_2 x + B_2 y + C_2) = 0$$

where the coefficients A_i, B_i, C_i are the signed minors of the matrix Q.

□

Example 5.27

To determine an implicit equation for the conic

$$(x(t), y(t)) = \left(\frac{18 - 19t + 8t^2}{-2 - 5t + 4t^2}, \frac{-6 + 7t + 3t^2}{-2 - 5t + 4t^2}\right) .$$

Apply Theorem 5.26 to

$$\mathsf{Q} = \begin{pmatrix} 18 & -19 & 8 \\ -6 & 7 & 3 \\ -2 & -5 & 4 \end{pmatrix} .$$

The required minors are

$$\begin{aligned} A_0 &= (7)(4) - (-5)(3) = 43 , \\ A_1 &= -((-6)(4) - (-2)(3)) = 18 \quad \text{etc.} \end{aligned}$$

Alternatively, compute the transpose of the adjugate matrix of Q, to give

$$\begin{pmatrix} A_0 & A_1 & A_2 \\ B_0 & B_1 & B_2 \\ C_0 & C_1 & C_2 \end{pmatrix} = \begin{pmatrix} 43 & 18 & 44 \\ 36 & 88 & 128 \\ -113 & -102 & 12 \end{pmatrix} .$$

Therefore

$$\begin{aligned} &(A_1 x + B_1 y + C_1)^2 - (A_0 x + B_0 y + C_0)(A_2 x + B_2 y + C_2) \\ &= (18x + 88y - 102)^2 - (43x + 36y - 113)(44x + 128y + 12) . \end{aligned}$$

Expanding and simplifying gives $784\left(-2x^2 - 5xy + x + 4y^2 - 5y + 15\right) = 0$. The solution reverses the computation of Example 5.22.

Exercise 5.30

Convert the following conics from parametric to implicit form:

(a) $\left(t^2 - 1, t + 2\right)$,

(b) $\left(2t^2 - 1, t + 3\right)$,

(c) $\left(2t^2 + t - 1, t^2 - 3t + 3\right)$,

(d) $\left(\frac{t^2+1}{t}, 2t\right)$,

(e) $\left(-\frac{4t^2+t+1}{4t^2-2t+1}, -\frac{3t^2}{4t^2-2t+1}\right)$.

Theorem 5.26 can be generalized to planar rational curves of any degree (see [23], [13]):

Theorem 5.28

Let

$$(x(t), y(t)) = \left(\frac{\sum_{i=0}^{n} a_i t^i}{\sum_{i=0}^{n} c_i t^i}, \frac{\sum_{i=0}^{n} b_i t^i}{\sum_{i=0}^{n} c_i t^i} \right)$$

be a rational curve of degree of degree n. Then an implicit form is obtained from the *Bezout resultant*

$$\begin{vmatrix} D_{0,0} & \cdots & D_{0,n} \\ \vdots & & \vdots \\ D_{n,0} & \cdots & D_{n,n} \end{vmatrix} = 0 ,$$

where $D_{i,j} = \sum_{\substack{k \le \min(i,j) \\ m=i+j-k+1}} (b_m c_k - c_m b_k)\, x + (a_k c_m - a_m c_k)\, y + (a_m b_k - a_k b_m)$.

□

5.7 Conics in Space

A conic in three-dimensional space is given parametrically by

$$\left(\frac{a_0 + a_1 t + a_2 t^2}{d_0 + d_1 t + d_2 t^2}, \frac{b_0 + b_1 t + b_2 t^2}{d_0 + d_1 t + d_2 t^2}, \frac{c_0 + c_1 t + c_2 t^2}{d_0 + d_1 t + d_2 t^2} \right) .$$

Any conic in space is contained in a plane. To verify this, suppose every point of the conic lies in the plane $Ax + By + Cz + D = 0$. Then

$$\begin{aligned} A\left(a_0 + a_1 t + a_2 t^2\right) &+ B\left(b_0 + b_1 t + b_2 t^2\right) \\ &+ C\left(c_0 + c_1 t + c_2 t^2\right) + D\left(d_0 + d_1 t + d_2 t^2\right) = 0 , \end{aligned}$$

that is,

$$\begin{aligned} (Aa_2 + Bb_2 + Cc_2 + Dd_2)\, t^2 &+ (Aa_1 + Bb_1 + Cc_1 + Dd_1)\, t \\ &+ Aa_0 + Bb_0 + Cc_0 + Dd_0 = 0 . \quad (5.20) \end{aligned}$$

Since this holds for all t (in an interval) the coefficients of (5.20) must be identically zero, implying

$$\begin{aligned} Aa_0 + Bb_0 + Cc_0 + Dd_0 &= 0 , \\ Aa_1 + Bb_1 + Cc_1 + Dd_1 &= 0 , \text{ and} \\ Aa_2 + Bb_2 + Cc_2 + Dd_2 &= 0 . \end{aligned}$$

The equations can be interpreted as defining three planes in the three-dimensional projective space with homogeneous coordinates (A, B, C, D). Thus the coefficients A, B, C, D can be determined using the method for computing the intersection of three planes given in Section 3.4. This yields

$$(A, B, C, D) = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \\ a_0 & b_0 & c_0 & d_0 \\ a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \end{vmatrix}$$

or any multiple of the determinant.

A planar representation of the conic can be obtained by applying a viewplane coordinate mapping. The origin and the X- and Y-axis directions in the derived plane are specified, and the viewplane coordinate matrix VC is computed. Then VC is applied to $(x(t), y(t), z(t))$ to give a conic in the specified Cartesian coordinate system.

Example 5.29

Consider the conic $(x(t), y(t), z(t)) = \left(\frac{3+6t-4t^2}{-1-6t+2t^2}, \frac{9-6t^2}{-1-6t+2t^2}, \frac{1-3t+t^2}{-1-6t+2t^2}\right)$. Then

$$\begin{aligned} (A, B, C, D) &= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \\ 3 & 9 & 1 & -1 \\ 6 & 0 & -3 & -6 \\ -4 & -6 & 1 & 2 \end{vmatrix} \\ &= 54\mathbf{e}_1 - 18\mathbf{e}_2 + 36\mathbf{e}_3 + 36\mathbf{e}_4 \\ &= (54, -18, 36, 36) \ . \end{aligned}$$

After dividing (A, B, C, D) through by 18, the conic is found to lie in the plane $3x - y + 2z + 2 = 0$. This can be verified by substituting $x = x(t), y = y(t), z = z(t)$ into the equation of the plane,

$$3\left(\frac{3+6t-4t^2}{-1-6t+2t^2}\right) - \left(\frac{9-6t^2}{-1-6t+2t^2}\right) + 2\left(\frac{1-3t+t^2}{-1-6t+2t^2}\right) + 2 = 0 \ .$$

Multiplying through by the denominator gives

$$3\left(3+6t-4t^2\right) - \left(9-6t^2\right) + 2\left(1-3t+t^2\right) + 2\left(-1-6t+2t^2\right) = 0 \ .$$

Expanding the brackets yields that the left-hand side of the equation is identically zero, and hence the conic lies in the derived plane.

Let the plane have origin and axes as specified in Example 4.11. Applying the viewing coordinate mapping matrix VC to the conic gives

$$
\begin{aligned}
&\begin{pmatrix} 3+6t-4t^2 & 9-6t^2 & 1-3t+t^2 & -1-6t+2t^2 \end{pmatrix} \\
&\quad \times \begin{pmatrix} 0.385 & 0.360 & -0.333 \\ 0.642 & 0.600 & 0.111 \\ -0.706 & 0.960 & -0.222 \\ 0.449 & -1.200 & 0.778 \end{pmatrix} \\
&= \begin{pmatrix} 5.778+1.734t-5.2t^2 & 8.64+6.48t-6.48t^2 & -1-6t+2t^2 \end{pmatrix} .
\end{aligned}
$$

The planar representation of the conic is $\left(\frac{5.778+1.734t-5.2t^2}{-1-6t+2t^2}, \frac{8.64+6.48t-6.48t^2}{-1-6t+2t^2}\right)$. The vector form for the conic in space is

$$
\mathcal{O} + \left(\frac{5.778+1.734t-5.2t^2}{-1-6t+2t^2}\right)\mathbf{X} + \left(\frac{8.64+6.48t-6.48t^2}{-1-6t+2t^2}\right)\mathbf{Y} ,
$$

where $\mathcal{O}$, $\mathbf{X}$ and $\mathbf{Y}$ are as given in Example 4.11.

Exercise 5.31

By applying Theorem 5.6, show that the conic

$$
(x(t), y(t), z(t)) = \left(\frac{9+3t-4t^2}{3+4t+4t^2}, \frac{4t-3t^2}{3+4t+4t^2}, \frac{1+t^2}{3+4t+4t^2}\right)
$$

lies in the plane $-16x+55y+273z-43=0$.

5.8 Applications of Conics

Example 5.30 (Headlights and Radar)

Parabolas have the special property that rays of light, emanating from a light source positioned at the focus, are reflected in the parabola along parallel lines, as illustrated in Figure 5.15. This property is used in the design of car headlight reflectors. A reflector has the shape of a paraboloid, that is, a surface obtained by rotating a parabola about its axis of symmetry. If a headlight bulb is positioned at the focus of the parabola then it produces a beam of light consisting of the reflected parallel rays of light.

The same property is used in the design of radar or satellite dishes. Signals from a distant point travel along (nearly) parallel rays. Signals which reach a paraboloid shaped dish are reflected along linear paths which pass through the focus. The satellite receiver is positioned at the focus in order to obtain the best reception of the signals.

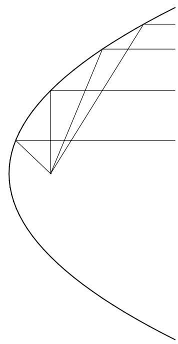


Figure 5.15

EXERCISES

5.32. Show that a ray of light emanating from the focus $\mathbf{F}(0, m)$ of the parabola $y = 4mx^2$ is reflected parallel to the axis of symmetry, that is the y-axis, as follows.

(a) Determine the tangent vector/line at a point $\mathbf{P}(x, y)$ on the parabola.

(b) Determine the angle α between the line $\mathbf{FP}$ and the tangent.

(c) Show that the angle between the tangent and y-axis is also α and deduce that the reflection of the ray is parallel to the y-axis.

5.33. Show that all rays of light emanating from one focus of an ellipse are reflected in lines containing the other focus. The reader may wish to contemplate an elliptical snooker table with the cue ball positioned at one focus and a single pocket positioned at the other focus. If the ball is struck (without spin) with sufficient strength then the ball will hit the elliptical cushion and rebound along a line containing the pocket.

Example 5.31 (Suspension Bridges)

Suspension bridges are designed so that a cable hanging from two pillars or towers carries the weight of the bridge uniformly along the cable. The resulting shape of the cable is a parabola. Suppose the pillars are $1,410$ metres apart (the span of the Humber Bridge, Hull, UK) and the height of the pillars is h metres. Let the origin be the lowest point of the parabola, and let the horizontal plane of the bridge be the x-axis. Then the parabola is symmetrical about the y-axis and passes through the points $(-705, h)$, $(705, h)$, $(0, 0)$. Let the parabola be $y = ax^2 + bx + c$. Then, clearly $c = 0$, and

$$\begin{aligned} h &= a(-705)^2 + b(-705), \text{ and} \\ h &= a(705)^2 + b(705). \end{aligned}$$

Thus $a = h/(705)^2$ and $b = 0$ giving the parabola $y = \frac{h}{705^2}x^2$.

In reality, one must account for the earth's curvature when modelling large structures. So in the example of the bridge the distance between the base of the pillars is $1,410$ metres, but the distance between the tops of the pillars is greater.

Example 5.32 (Radar)

Discovered as recently as the 1940s, the method known as hyperbolic navigation has had a considerable influence on sea and air navigation. A receiver records the radio signals transmitted from two fixed stations. Assuming that the velocity v of radio energy is constant, the distances travelled by radio energy are proportional to the time taken. Suppose the times taken to receive the signals, sent at the same time, from each station are t_1 and t_2. Then the distance from each station is vt_1 and vt_2. Hence the difference in distance of the receiver from the stations is $v(t_1 - t_2)$. The locus of all possible positions of the receiver relative to the fixed stations is a branch of a hyperbola with the stations positioned at the foci (see Exercise 5.34). There are two points on the hyperbola a given distance vt_1 from the first station and it remains to decide which is the correct location. Commercial hyperbolic navigation systems include the Decca Navigation System, LORAN, Omega, and Global Positioning Systems (GPS).

Exercise 5.34

Let $\mathbf{F}_1$ and $\mathbf{F}_2$ be fixed points, and let d_1 and d_2 be the distances of a point $\mathbf{P}$ from $\mathbf{F}_1$ and $\mathbf{F}_2$ respectively. Show that the locus of all points $\mathbf{P}$ such that $d_1 - d_2$ is constant is a hyperbola.