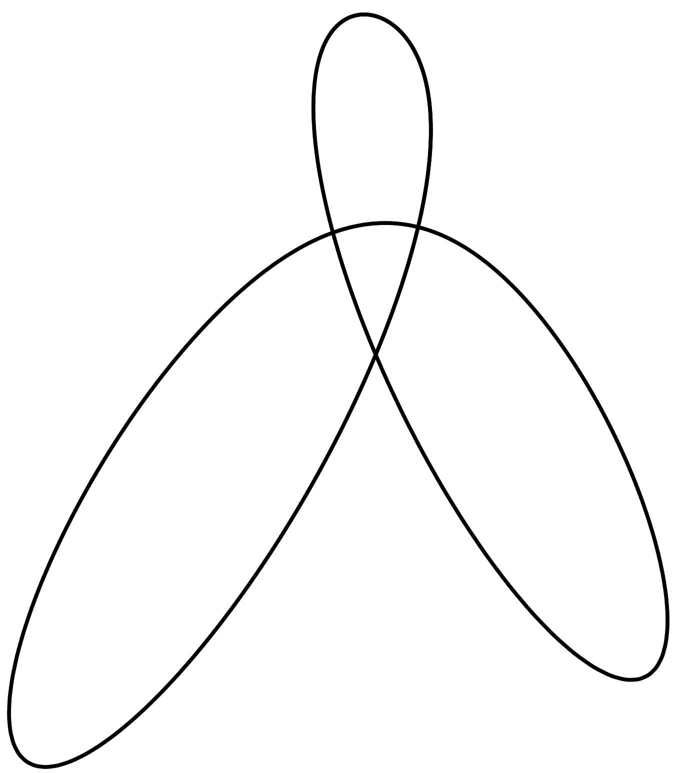
Chapter 22

Rational Curves



22.1 Rational Curves and Multiprojective Maps

In this chapter, rational curves are investigated. After a quick review of the traditional parametric definition in terms of homogeneous polynomials, we explore the possibility of defining rational curves in terms of polar forms. Because the polynomials involved are homogeneous, the polar forms are actually multilinear, and rational curves are defined in terms of multiprojective maps. Then, using the contruction at the end of the previous chapter, we define rational curves in BR-form. This allows us to handle rational curves in terms of control points in the hat space $\widehat{\mathcal{E}}$ obtained from $\mathcal{E}$. We present two versions of the de Casteljau algorithm for rational curves and show how subdivision can be easily extended from the polynomial case. We also present a very simple method for approximating closed rational curves. This method only uses two control polygons obtained from the original control polygon. We conclude this chapter with a quick tour in a gallery of rational curves.

We begin by defining rational polynomial curves in a traditional manner, and then, we show how this definition can be advantageously recast in terms of symmetric multiprojective maps. Keep in mind that rational polynomial curves really live in the projective completion $\widetilde{\mathcal{E}}$ of the affine space $\mathcal{E}$. This is the reason why homogeneous polynomials are involved. We will assume that the homogenization $\widehat{\mathbb{A}}$ of the affine line $\mathbb{A}$ is identified with the direct sum $\mathbb{R} \oplus \mathbb{R}\overline{1}$. Then, every element of $\widehat{\mathbb{A}}$ is of the form $(t, z) \in \mathbb{R}^2$.

A rational curve of degree m in an affine space $\mathcal{E}$ of dimension n is specified by n fractions, say

$$x_1(t) = \frac{F_1(t)}{F_{n+1}(t)}, \quad x_2(t) = \frac{F_2(t)}{F_{n+1}(t)}, \quad \ldots, \quad x_n(t) = \frac{F_n(t)}{F_{n+1}(t)},$$

where $F_1(X), \ldots, F_{n+1}(X)$ are polynomials of degree at most m. In order to deal with the case where the denominator $F_{n+1}(X)$ is null, we view the rational curve as the projection of the polynomial curve defined by the polynomials $F_1(X), \ldots, F_{n+1}(X)$. To make this rigorous, we can view the rational curve as a map $F\colon \widetilde{\mathbb{A}} \to \widetilde{\mathcal{E}}$. For this, we homogenize the polynomials $F_1(X), \ldots, F_{n+1}(X)$ as polynomials of the same total degree (replacing X by X/Z), so that after polarizing w.r.t. (X, Z), we get a (symmetric) multilinear map which induces a (symmetric) multiprojective map. Following this approach leads to the following definition.

Definition 22.1.1 Given some affine space $\mathcal{E}$ of finite dimension $n \geq 2$, an affine frame $(\Omega_1, (e_1, \ldots, e_n))$ for $\mathcal{E}$, and the corresponding basis $(e_1, \ldots, e_n, \langle\Omega_1, 1\rangle)$ for $\widehat{\mathcal{E}}$, a *rational curve of degree* m is a function $F\colon \widehat{\mathbb{A}} \to \widehat{\mathcal{E}}$, such that, for all $(t, z) \in \mathbb{R}^2$, we have

$$F(t, z) = F_1(t, z)e_1 \mathbin{\widehat{+}} \cdots \mathbin{\widehat{+}} F_n(t, z)e_n \mathbin{\widehat{+}} F_{n+1}(t, z)\langle\Omega_1, 1\rangle,$$

where $F_1(X, Z), \ldots, F_{n+1}(X, Z)$ are homogeneous polynomials in $\mathbb{R}[X, Z]$, each of total degree m.

It is important to require that all of the polynomials $F_1(X, Z), \ldots, F_{n+1}(X, Z)$ are homogeneous and of the *same total degree* m. Otherwise, we would not be able to polarize these polynomials and obtain a symmetric *multilinear map* (as opposed to a multiaffine map).

For example, the following homogeneous polynomials define a (unit) circle:

$$\begin{aligned} F_1(X, Z) &= Z^2 - X^2 \\ F_2(X, Z) &= 2XZ \\ F_3(X, Z) &= Z^2 + X^2. \end{aligned}$$

Note that we recover the standard parameterization of the circle, by setting $Z = 1$. But in general, it is also possible to get points at infinity in the projective space $\widetilde{\mathcal{E}}$, or a point corresponding to the point at infinity $(1, 0)$ in $\widehat{\mathbb{A}}$. For example, for $(X, Z) = (1, 0)$, we get the point of homogeneous coordinates $(-1, 0, 1)$.

Using a Lemma in Gallier [70] (see Lemma 27.2.1), for each homogeneous polynomial $F_i(X, Z)$ of total degree m, there is a unique symmetric multilinear map $f_i\colon (\mathbb{R}^2)^m \to \mathbb{R}$, such that

$$F_i(t, z) = f_i(\underbrace{(t, z), \ldots, (t, z)}_{m}),$$

for all $(t, z) \in \mathbb{R}^2$, and together, these $n + 1$ maps define a symmetric multilinear map

$$f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}},$$

such that

$$F(t, z) = f(\underbrace{(t, z), \ldots, (t, z)}_{m}),$$

for all $(t, z) \in \mathbb{R}^2$. But then, we get a symmetric multiprojective map

$$\mathbf{P}(f)\colon (\widetilde{\mathbb{A}})^m \to \widetilde{\mathcal{E}},$$

and we define the projective rational curve

$$\widetilde{F}\colon \widetilde{\mathbb{A}} \to \widetilde{\mathcal{E}},$$

such that

$$\widetilde{F}([t, z]) = \mathbf{P}(f)(\underbrace{[t, z], \ldots, [t, z]}_{m})$$

for all $(t, z) \in \mathbb{R}^2$, where we view (t, z) as homogeneous coordinates, that is, where $(t, z) \neq (0, 0)$. This leads us to the following definition.

Definition 22.1.2 Given an affine space $\mathcal{E}$ of dimension ≥ 2, a *projective rational curve F of degree m* is a (partial) projective map $F\colon \widetilde{\mathbb{A}} \to \widetilde{\mathcal{E}}$ such that there is some symmetric multilinear map $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$, with

$$F([t, z]) = \mathbf{P}(f)(\underbrace{[t, z], \ldots, [t, z]}_{m}),$$

for all $[t, z] \in \widetilde{\mathbb{A}}$ (i.e., all homogeneous coordinates $(t, z) \in \mathbb{R}^2$). The symmetric multilinear map f is also called a *polar form* of F. The *trace of F* is the subset $F(\widetilde{\mathbb{A}})$ of the projective completion $\widetilde{\mathcal{E}}$ of the affine space $\mathcal{E}$.

Since $\widetilde{\mathbb{A}} = \mathbb{RP}^1$, a projective rational curve F is a rational map $F\colon \mathbb{RP}^1 \to \widetilde{\mathcal{E}}$ defined as the diagonal of a symmetric multiprojective map (namely, $\mathbf{P}(f)$).

One should be aware that in general, the symmetric multinear map $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$ inducing the projective map $F\colon \widetilde{\mathbb{A}} \to \widetilde{\mathcal{E}}$ via

$$F([t, z]) = \mathbf{P}(f)(\underbrace{[t, z], \ldots, [t, z]}_{m})$$

for all $[t, z] \in \widetilde{\mathbb{A}}$, is **not unique**.

Let $f\colon (\widehat{\mathbb{A}})^2 \to \widehat{\mathbb{A}}$ be the bilinear map defined such that

$$f((x_1, z_1), (x_2, z_2)) = \left(ax_1x_2 + b\frac{x_1z_2 + x_2z_1}{2},\ a\frac{x_1z_2 + x_2z_1}{2} + bz_1z_2\right).$$

Then, we have

$$\mathbf{P}(f)([x_1, z_1], [x_2, z_2]) = \left[a x_1 x_2 + b \frac{x_1 z_2 + x_2 z_1}{2},\ a \frac{x_1 z_2 + x_2 z_1}{2} + b z_1 z_2 \right],$$

and thus,

$$F([x, z]) = \mathbf{P}(f)([x, z], [x, z]) = [x(ax + bz),\ z(ax + bz)] = [x, z],$$

which shows that $F\colon \widetilde{\mathbb{A}} \to \widetilde{\mathbb{A}}$ is the identity. However, by choosing different pairs (a, b), we get distinct bilinear maps f, all inducing the identity.

From the discussion just before definition 22.1.2, we note that every rational curve F as in definition 22.1.1 defines a projective rational curve $\widetilde{F}$.

The converse is also true when $\mathcal{E}$ is of finite dimension. Let G be a projective rational curve defined by the symmetric multilinear map $g\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$. Given an affine frame $(\Omega_1, (e_1, \ldots, e_n))$ for $\mathcal{E}$, and the corresponding basis $(e_1, \ldots, e_n, \langle \Omega_1, 1 \rangle)$ for $\widehat{\mathcal{E}}$, a Lemma from Gallier [70] (see Lemma 27.1.5) shows that a symmetric multilinear map $g\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$ defines a homogeneous polynomial map F of degree m in two variables, with coefficients in $\widehat{\mathcal{E}}$. Expressing these coefficients over the basis $(e_1, \ldots, e_n, \langle \Omega_1, 1 \rangle)$, we see that the polynomial map F is defined by some homogeneous polynomials $F_1(X, Z), \ldots, F_{n+1}(X, Z)$ in $\mathbb{R}[X, Z]$, each of total degree m, such that, for all $(t, z) \in \mathbb{R}^2$, we have

$$F(t, z) = F_1(t, z) e_1 \mathbin{\widehat{+}} \cdots \mathbin{\widehat{+}} F_n(t, z) e_n \mathbin{\widehat{+}} F_{n+1}(t, z) \langle \Omega_1, 1 \rangle.$$

Thus F is a rational curve, and clearly, $\widetilde{F} = G$. Therefore, when $\mathcal{E}$ is of finite dimension, definition 22.1.2 and definition 22.1.1 define the same class of projective rational curves. Definition 22.1.2 will be preferred, since it leads to a treatment of rational curves in terms of control points rather similar to the treatment of Bézier curves (it also works for any dimension, even infinite).

Remark: If the homogeneous polynomials $F_1(X, Z), \ldots, F_{n+1}(X, Z)$ have some common zero (t, z), then $\mathbf{P}(F)([t, z])$ is undefined. Such points are called *base points*. This may appear to be a problem, but in the case of curves, this problem can be fixed easily. Indeed, we can use greatest common divisors and continuity to define $\mathbf{P}(F)([t, z])$ on the common zeros of the F_i. If $(t, 0)$ is a common zero of the polynomials $F_1(X, Z), \ldots, F_{n+1}(X, Z)$, since $F_1(X, 0), \ldots, F_{n+1}(X, 0)$ are polynomials in the single variable X, they have the common root $X = t$, and thus, they are all divisible by $X - t$. If $G(X)$ is the greatest common divisor of the polynomials $F_1(X, 0), \ldots, F_{n+1}(X, 0)$, then by continuity, we let

$$\left(\frac{F_1(t, 0)}{G(t)}, \ldots, \frac{F_{n+1}(t, 0)}{G(t)} \right),$$

be the homogeneous coordinates of $F(t, 0)$.

If (t, z) is a common zero of the polynomials $F_1(X, Z), \ldots, F_{n+1}(X, Z)$, where $z \neq 0$, then $(u, 1) = \left(\frac{t}{z}, 1 \right)$ is also a common zero, since the polynomials are homogeneous. Then, by setting $Z = 1$, we obtain polynomials $F_1(X, 1), \ldots, F_{n+1}(X, 1)$ in the single variable X, and they are all divisible by $X - u$. As before, if $G(X)$ is the greatest common divisor of the polynomials $F_1(X, 1), \ldots, F_{n+1}(X, 1)$, then by continuity, we let

$$\left(\frac{F_1(u, 1)}{G(u)}, \ldots, \frac{F_{n+1}(u, 1)}{G(u)} \right),$$

be the homogeneous coordinates of $F(t, z)$.

It should be noted that such a method does not always work in the case of surfaces, because for polynomials in two or more variables, having a common zero does not imply the existence of a greatest common divisor. Continuity does not work either.

However, as long as (t,z) is not a common zero of all the $F_i(X,Z)$, note that we can now deal with the case where (t,z) is a zero of the "denominator $F_{n+1}(X,Z)$". In this case, we get a point at infinity of homogeneous coordinates

$$(F_1(t,z),\ldots,F_n(t,z),0).$$

Remarks: A way to avoid that $\mathbf{P}(F)([t,z])$ be undefined when $z \neq 0$, is to assume that the polynomials $F_1(X,1),\ldots,F_{n+1}(X,1)$ are relatively prime. Then if $(t,0)$ is a common zero of the polynomials $F_1(X,Z),\ldots,F_{n+1}(X,Z)$, we determine $\mathbf{P}(F)([t,0])$ by continuity, as we just explained. A way to deal with improperly parameterized rational curves is proposed in Sederberg [153].

Given a projective rational curve $F\colon \widetilde{\mathbb{A}} \to \widetilde{\mathcal{E}}$ defined by some symmetric multilinear map $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$, with

$$F([t,z]) = \mathbf{P}(f)(\underbrace{[t,z],\ldots,[t,z]}_{m}),$$

for all $[t,z] \in \widetilde{\mathbb{A}}$, for any projectivity $h\colon \widetilde{\mathbb{A}} \to \widetilde{\mathbb{A}}$, the map $F \circ h$ is also a projective rational curve. Indeed, as we saw earlier, a projectivity $h\colon \widetilde{\mathbb{A}} \to \widetilde{\mathbb{A}}$ is defined by a linear map $g\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by an invertible matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

with $ad - bc \neq 0$. The linear map g can be viewed as a linear map from $\widehat{\mathbb{A}}$ to itself, and the map

$$f \circ (g,\ldots,g)\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$$

is clearly a symmetric multilinear map inducing the map $F \circ h$. The curve $F \circ h$ is said to be obtained from F by the *change of parameter h*.

If $F\colon \widetilde{\mathbb{A}} \to \widetilde{\mathcal{E}}$ is a projective curve defined by some symmetric multilinear map $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$, with

$$F([t,z]) = \mathbf{P}(f)(\underbrace{[t,z],\ldots,[t,z]}_{m}),$$

for all $[t,z] \in \widetilde{\mathbb{A}}$, by lemma 21.3.2, the restriction $g\colon \mathbb{A}^m \to \widehat{\mathcal{E}}$ of $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$ to $\mathbb{A}^m$ is a multiaffine map such that $\mathbf{P}(f) = \Pi\Omega \circ \widetilde{g}$. The multiaffine map $g\colon \mathbb{A}^m \to \widehat{\mathcal{E}}$ defines a polynomial curve $G\colon \mathbb{A} \to \widehat{\mathcal{E}}$, and the multilinear map $\widehat{g}\colon (\widehat{\mathbb{A}})^m \to \widehat{\widehat{\mathcal{E}}}$ induces a projective rational curve $\widetilde{G}\colon \widetilde{\mathbb{A}} \to \widetilde{\widehat{\mathcal{E}}}$, and $\mathbf{P}(f) = \Pi\Omega \circ \widetilde{g}$ shows that the projective rational curve F is the projection under $\Pi\Omega$ of the projective rational curve $\widetilde{G}$, which itself, can be considered as the projective completion of the polynomial curve G. This suggests the following definition.

Definition 22.1.3 Given a polynomial curve $F\colon \mathbb{A} \to \mathcal{E}$ defined by the polar form $f\colon \mathbb{A}^m \to \mathcal{E}$, the symmetric multilinear map $\widehat{f}\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$ induces a projective rational curve $\widetilde{F}\colon \widetilde{\mathbb{A}} \to \widetilde{\mathcal{E}}$, called the *projective completion* of the polynomial curve F.

Clearly the restriction of the curve $\widetilde{F}$ to $\mathbb{A}$ agrees with the polynomial curve F. The projective completion $\widetilde{F}$ of F can still be considered as a polynomial curve, except that it is also defined for the point at infinity in $\widetilde{\mathbb{A}}$.

Now, given a projective rational curve F with polar form $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$, the discussion just before definition 22.1.3 shows that F is completely determined by the polynomial curve G of polar form $g\colon \mathbb{A}^m \to \widehat{\mathcal{E}}$, the restriction of $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$ to $\mathbb{A}^m$, and the projection $\Pi\Omega$. We also know from section 18.1 that the polynomial curve G is determined by its $m+1$ control points, which are vectors $\theta_0,\ldots,\theta_m$ in $\widehat{\mathcal{E}}$. Thus, we are led to the definition of a BR-curve.

Definition 22.1.4 Given any affine space $\mathcal{E}$ of dimension $n \geq 2$, a *rational curve in Bézier form of degree m, or BR-curve*, is a map $F\colon \widetilde{\mathbb{A}} \to \widetilde{\mathcal{E}}$ such that there is a polynomial curve $G\colon \mathbb{A} \to \widehat{\mathcal{E}}$ of degree m defined by $m+1$ control points $\theta_0, \ldots, \theta_m$ in $\widehat{\mathcal{E}}$ and such that if $g\colon \mathbb{A}^m \to \widehat{\mathcal{E}}$ is the (unique) polar form of G, then

$$F(\overline{t}) = \begin{cases} \Pi(G(\overline{t})) & \text{if } \overline{t} \neq \infty; \\ \Pi\Omega([\widehat{g}((1,0), \ldots, (1,0))]_{\approx}) & \text{when } \overline{t} = \infty. \end{cases}$$

The *trace of* F is the subset $F(\widetilde{\mathbb{A}})$ of the projective completion $\widetilde{\mathcal{E}}$ of the affine space $\mathcal{E}$.

Note that in practice, we can assume that F is only defined on $\mathbb{A}$.

The following lemma shows that the class of projective rational curves and the class of BR-curves are identical.

Lemma 22.1.5 *Every rational projective curve $F\colon \widetilde{\mathbb{A}} \to \widetilde{\mathcal{E}}$ with polar form $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$ is the image under $\Pi\Omega$ of the projective completion $\widetilde{G}\colon \widetilde{\mathbb{A}} \to \widetilde{\widehat{\mathcal{E}}}$ of the polynomial curve $G\colon \mathbb{A} \to \widehat{\mathcal{E}}$ associated with the restriction $g\colon \mathbb{A}^m \to \widehat{\mathcal{E}}$ of f to $\mathbb{A}^m$. Conversely, given a polynomial curve $G\colon \mathbb{A} \to \widehat{\mathcal{E}}$ with polar form $g\colon \mathbb{A}^m \to \widehat{\mathcal{E}}$, the image under $\Pi\Omega$ of the projective completion $\widetilde{G}\colon \widetilde{\mathbb{A}} \to \widetilde{\widehat{\mathcal{E}}}$ of G is a projective rational curve. As a consequence, the class of projective rational curves and the class of BR-curves are identical.*

Proof. The first and the second part follow from lemma 21.3.2. If F is a BR-curve, and $g\colon \mathbb{A}^m \to \widehat{\mathcal{E}}$ is the polar form associated with the polynomial curve G determined by $\theta_0, \ldots, \theta_m$ in $\widehat{\mathcal{E}}$, by lemma 21.3.2, there is a multilinear map $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$ whose restriction to $\mathbb{A}^m$ is g, and we have $\mathbf{P}(f) = \Pi\Omega \circ \widetilde{g}$. By lemma 21.2.6 applied to the special case where $\overrightarrow{F} = \widehat{\mathcal{E}}$, the natural projection $\Pi\colon (\widehat{\mathcal{E}} - \{\Omega\}) \to \widetilde{\mathcal{E}}$ is the restriction of the central projection $\Pi\Omega\colon (\widetilde{\widehat{\mathcal{E}}} - \{\Omega\}) \to \widetilde{\mathcal{E}}$ to $(\widehat{\mathcal{E}} - \{\Omega\})$ (where Ω is the origin 0 of $\widehat{\mathcal{E}}$, where $\widehat{\mathcal{E}}$ is viewed as an affine space). Since $\widehat{G}(t,1) = \widehat{g}((t,1), \ldots, (t,1))$ is not a point at infinity for $\overline{t} \in \mathbb{A}$, we have

$$\begin{aligned} \Pi\Omega(\widetilde{G}(t,1)) &= \Pi\Omega(\widetilde{g}([t,1], \ldots, [t,1])) \\ &= \Pi\Omega([\widehat{g}((t,1), \ldots, (t,1))]_{\approx}) \\ &= \Pi(g(\overline{t}, \ldots, \overline{t})) \\ &= \Pi(G(\overline{t})). \end{aligned}$$

By definition 22.1.4, this shows that the BR-curve F is the projective rational curve defined by $\mathbf{P}(f)$. Conversely, a projective rational curve $F\colon \widetilde{\mathbb{A}} \to \widetilde{\mathcal{E}}$ defined by a polar form $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$ is the projection under $\Pi\Omega$ of the projective completion $\widetilde{G}\colon \widetilde{\mathbb{A}} \to \widetilde{\widehat{\mathcal{E}}}$ of the polynomial curve $G\colon \mathbb{A} \to \widehat{\mathcal{E}}$ associated with the restriction $g\colon \mathbb{A}^m \to \widehat{\mathcal{E}}$ of f to $\mathbb{A}^m$. But then, using the same argument as above, $\Pi\Omega(\widehat{G}(t,1)) = \Pi(G(\overline{t}))$, for $\overline{t} \in \mathbb{A}$, and thus, F is BR-curve. $\square$

Assuming that $\mathcal{E}$ is of dimension $n \geq 2$, when a projective rational curve F is defined by a symmetric multilinear map $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$ which is obtained by polarizing some homogeneous polynomials $F_1(X,Z), \ldots, F_{n+1}(X,Z)$ in $\mathbb{R}[X,Z]$, each of total degree m, the polynomial curve $G\colon \mathbb{A} \to \widehat{\mathcal{E}}$ associated with the restriction $g\colon \mathbb{A}^m \to \widehat{\mathcal{E}}$ of f to $\mathbb{A}^m$ is simply determined by the polynomials $F_1(X,1), \ldots, F_{n+1}(X,1)$ in $\mathbb{R}[X]$ obtained by setting $Z = 1$. If we homogenize the polynomials $F_1(X,1), \ldots, F_{n+1}(X,1)$, getting homogeneous polynomials of total degree m, we get the original homogeneous polynomials $F_1(X,Z), \ldots, F_{n+1}(X,Z)$ back.

When the polynomial curve $G\colon \mathbb{A} \to \widehat{\mathcal{E}}$ inducing the projective rational curve $F\colon \widetilde{\mathbb{A}} \to \widetilde{\mathcal{E}}$ is given by some polynomials $F_1(X), \ldots, F_{n+1}(X)$, since a projectivity $h\colon \widetilde{\mathbb{A}} \to \widetilde{\mathbb{A}}$ is defined by a linear map $g\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by an invertible matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

with $ad - bc \neq 0$, a change of parameter h consists in substituting

$$\frac{aZ + b}{cZ + d}$$

for X in the polynomials, and readjusting the polynomials so that denominators disapear. The curve resulting from the change of parameter has the same trace as the original curve (since the change of parameter is a projectivity).

Thus, in practice, when defining (projective) rational curves, it is enough to specify $n+1$ polynomials $F_1(X), \ldots, F_{n+1}(X)$ in $\mathbb{R}[X]$. Such polynomials define a polynomial curve $G\colon \mathbb{A} \to \widehat{\mathcal{E}}$. Since G is a curve in $\widehat{\mathcal{E}}$, control points $\theta \in \widehat{\mathcal{E}}$ are either *weighted points* of the form $\langle a, w\rangle$, where $w \in \mathbb{R}$ ($w \neq 0$) is called the *weight of* $a \in \mathcal{E}$, or *control vectors* $u \in \overrightarrow{\mathcal{E}}$. We can then determine the control points of the curve G by computing the polar form g of G. But we have to remember that the coordinates of these control points will be with respect to the basis $(e_1, \ldots, e_n, \langle \Omega_1, 1\rangle)$ of $\widehat{\mathcal{E}}$, where $(\Omega_1, (e_1, \ldots, e_n))$ is the affine frame for $\mathcal{E}$. Thus, for every $\theta = \langle a, \lambda\rangle \in \widehat{\mathcal{E}}$, to find the coordinates of the point $a \in \mathcal{E}$, if θ has coordinates $(x_1, \ldots, x_n, \lambda)$ over $(e_1, \ldots, e_n, \langle \Omega_1, 1\rangle)$, then $a \in E$ has coordinates

$$\left(\frac{x_1}{\lambda}, \ldots, \frac{x_n}{\lambda}\right),$$

over $(\Omega_1, (e_1, \ldots, e_n))$, as explained in lemma 4.2.1. Let us give some examples.

Example 1. Consider the circle defined by the polynomials

$$\begin{aligned} G_1(X) &= 1 - X^2 \\ G_2(X) &= 2X \\ G_3(X) &= 1 + X^2. \end{aligned}$$

We get the polar forms

$$\begin{aligned} g_1(X_1, X_2) &= 1 - X_1X_2 \\ g_2(X_1, X_2) &= X_1 + X_2 \\ g_3(X_1, X_2) &= 1 + X_1X_2. \end{aligned}$$

With respect to the affine frame $(\overline{0}, \overline{1})$ in $\mathbb{A}$, we get the following three control points $\theta_0, \theta_1, \theta_2$:

$$\begin{aligned} \theta_0 &= (1, 0, 1) \\ \theta_1 &= (1, 1, 1) \\ \theta_2 &= (0, 2, 2). \end{aligned}$$

We have to remember that the above are the coordinates of the control points with respect to $\widehat{\mathcal{E}}$. The points $a, b, c \in \mathcal{E}$, corresponding to the control points $\theta_0, \theta_1, \theta_2$, are $a = (1, 0)$, $b = (1, 1)$, and $c = (0, 1)$. Thus, we have $\theta_0 = \langle a, 1\rangle$, $\theta_1 = \langle b, 1\rangle$, and $\theta_2 = \langle c, 2\rangle$. Note that c has weight 2. When $t \in [0, 1]$, the point $F(t) = \Pi(G(t))$ is on the circle between a and c (a quarter-circle). The entire circle, excluding the point $d = (-1, 0)$, is obtained when $t \in \mathbb{A}$. The point d is obtained for $t = \infty$.

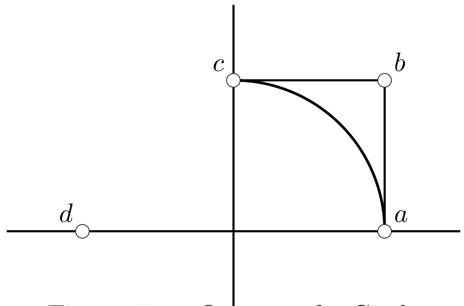


Figure 22.1: Quarter of a Circle

Let us consider another way of describing a circle.

Example 2. Consider the circle defined by the polynomials

$$\begin{aligned} G_1(X) &= 1 - 2X \\ G_2(X) &= 2X - 2X^2 \\ G_3(X) &= 2X^2 - 2X + 1. \end{aligned}$$

We get the polar forms

$$\begin{aligned} g_1(X_1, X_2) &= 1 - (X_1 + X_2) \\ g_2(X_1, X_2) &= X_1 + X_2 - 2X_1X_2 \\ g_3(X_1, X_2) &= 2X_1X_2 - (X_1 + X_2) + 1. \end{aligned}$$

With respect to the affine frame $(\overline{0}, \overline{1})$ in $\mathbb{A}$, we get the following three control points $\theta_0, \theta_1, \theta_2$:

$$\begin{aligned} \theta_0 &= (1, 0, 1) \\ \theta_1 &= (0, 1, 0) \\ \theta_2 &= (-1, 0, 1). \end{aligned}$$

This time, we note that θ_1 is the control vector e_2. The points $a, d \in \mathcal{E}$, corresponding to the control points θ_0 and θ_2, are $a = (1, 0)$, and $d = (-1, 0)$. Thus, we have $\theta_0 = \langle a, 1\rangle$, $\theta_1 = e_2$, and $\theta_2 = \langle d, 1\rangle$. When $t \in [0, 1]$, the point $F(t) = \Pi(G(t))$ is on the half circle of diameter ad. The entire circle is obtained when $t \in \widetilde{\mathbb{A}}$.

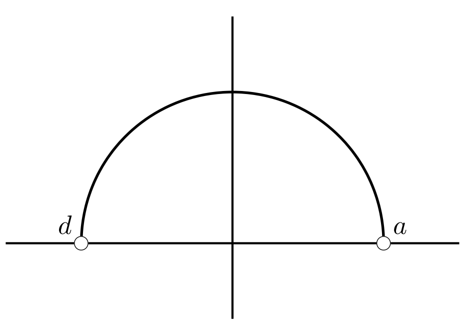


Figure 22.2: Half Circle

Example 3. Consider the hyperbola defined by the polynomials

$$\begin{aligned} G_1(X) &= X^2 \\ G_2(X) &= 1 \\ G_3(X) &= X. \end{aligned}$$

We get the polar forms

$$\begin{aligned} g_1(X_1, X_2) &= X_1X_2 \\ g_2(X_1, X_2) &= 1 \\ g_3(X_1, X_2) &= \frac{X_1 + X_2}{2}. \end{aligned}$$

With respect to the affine frame $(\overline{0}, \overline{1})$ in $\mathbb{A}$, we get the following three control points $\theta_0, \theta_1, \theta_2$:

$$\begin{aligned} \theta_0 &= (0, 1, 0) \\ \theta_1 &= \left(0, 1, \frac{1}{2}\right) \\ \theta_2 &= (1, 1, 1). \end{aligned}$$

We note that θ_0 is the control vector e_2. The points $c_2, b \in \mathcal{E}$, corresponding to the control points θ_1 and θ_2, are $c_2 = (0, 2)$, and $b = (1, 1)$. Thus, we have $\theta_0 = e_2$, $\theta_1 = \left\langle c_2, \frac{1}{2} \right\rangle$, and $\theta_2 = \langle b, 1 \rangle$. When $t \in [0, 1]$, the point $F(t) = \Pi(G(t))$ is on the arc of hyperbola from the point at infinity in the direction of e_2 along the axis $X = 0$, to the point b. The axis $X = 0$ is an asymptote. The entire hyperbola is obtained when $t \in \widetilde{\mathbb{A}}$.

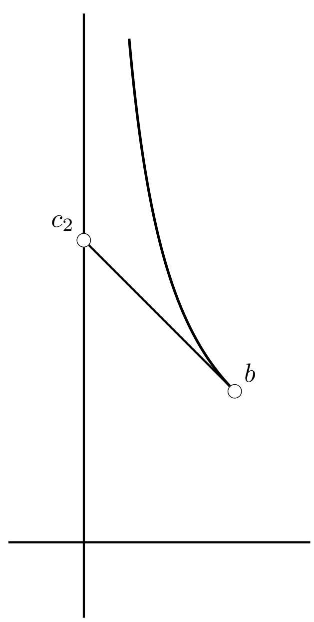


Figure 22.3: An arc of Hyperbola

Example 4. Consider the ellipse defined by the polynomials

$$\begin{aligned} G_1(X) &= 2 - 2X^2 \\ G_2(X) &= 2X \\ G_3(X) &= 1 + X^2. \end{aligned}$$

We get the polar forms

$$
\begin{aligned}
g_1(X_1, X_2) &= 2 - 2X_1X_2 \\
g_2(X_1, X_2) &= X_1 + X_2 \\
g_3(X_1, X_2) &= 1 + X_1X_2.
\end{aligned}
$$

With respect to the affine frame $(\overline{0}, \overline{1})$ in $\mathbb{A}$, we get the following three control points $\theta_0, \theta_1, \theta_2$:

$$
\begin{aligned}
\theta_0 &= (2, 0, 1) \\
\theta_1 &= (2, 1, 1) \\
\theta_2 &= (0, 2, 2).
\end{aligned}
$$

The points $a_2, b_2, c \in \mathcal{E}$, corresponding to the control points $\theta_0, \theta_1, \theta_2$, are $a_2 = (2, 0)$, $b_2 = (2, 1)$, and $c = (0, 1)$. Thus, we have $\theta_0 = \langle a_2, 1\rangle$, $\theta_1 = \langle b_2, 1\rangle$, and $\theta_2 = \langle c, 2\rangle$. Note that c has weight 2. When $t \in [0, 1]$, the point $F(t) = \Pi(G(t))$ is on the ellipse between a_2 and c (a quarter-ellipse). The entire ellipse, excluding the point $d_2 = (-2, 0)$, is obtained when $t \in \mathbb{A}$. The point d_2 is obtained for $t = \infty$.

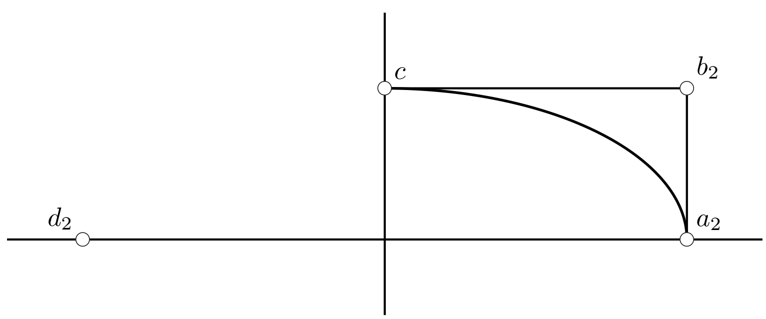


Figure 22.4: Quarter of Ellipse

Example 5. Consider the singular cubic defined by the polynomials

$$
\begin{aligned}
G_1(X) &= X^2 - X^3 \\
G_2(X) &= X^3 \\
G_3(X) &= (1 - X)^3.
\end{aligned}
$$

We get the polar forms

$$
\begin{aligned}
g_1(X_1, X_2, X_3) &= \frac{X_1X_2 + X_1X_3 + X_2X_3}{3} - X_1X_2X_3 \\
g_2(X_1, X_2, X_3) &= X_1X_2X_3 \\
g_3(X_1, X_2, X_3) &= 1 - (X_1 + X_2 + X_3) + X_1X_2 + X_1X_3 + X_2X_3 - X_1X_2X_3.
\end{aligned}
$$

With respect to the affine frame $(\overline{0}, \overline{1})$ in $\mathbb{A}$, we get the following four control points $\theta_0, \theta_1, \theta_2$, θ_3:

$$
\begin{aligned}
\theta_0 &= (0, 0, 1) \\
\theta_1 &= (0, 0, 0) \\
\theta_2 &= \left(\frac{1}{3}, 0, 0\right) \\
\theta_3 &= (0, 1, 0).
\end{aligned}
$$

This example is particularly remarkable, because we get the control point $\theta_0 = \langle \Omega_1, 1\rangle$, where Ω_1 is the origin of the affine plane, and the three control vectors $\theta_1 = 0 = \Omega$, $\theta_2 = \frac{e_1}{3}$, and $\theta_3 = e_2$. The projection $\Pi(\Omega)$ of Ω (the origin $\Omega = 0$ of $\widehat{\mathcal{E}}$) is not even defined! The other two control vectors are projected to points at infinity. In fact, the curve F of this example has a cusp at the origin. It is the "pointy" singular cubic of implicit equation $Y^2 = X^3$, and when $t \in [0, 1]$, the point $F(t) = \Pi(G(t))$ travels on the branch from the origin to the point at infinity in the direction e_2. As we shall see shortly, a version of the de Casteljau algorithm can be used to construct F. The cubic $Y^2 = X^3$ can be parameterized in a simpler way, for example as $X = t^2$, $Y = t^3$. But then, over the interval $[0, 1]$, only the portion of the curve between the origin and the point $(1, 1)$ is represented.

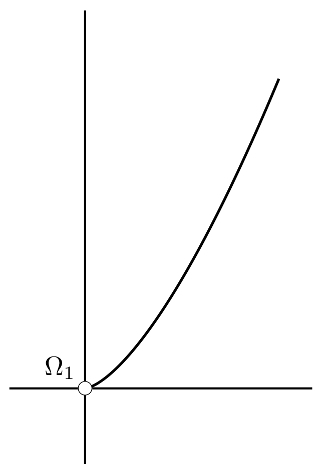


Figure 22.5: Cuspidal Cubic

Example 6. Consider the following cubic, known as the "Folium of Descartes", and defined by the polynomials

$$\begin{aligned} G_1(X) &= 3X \\ G_2(X) &= 3X^2 \\ G_3(X) &= 1 + X^3. \end{aligned}$$

We get the polar forms

$$\begin{aligned} g_1(X_1, X_2, X_3) &= X_1 + X_2 + X_3 \\ g_2(X_1, X_2, X_3) &= X_1X_2 + X_1X_3 + X_2X_3 \\ g_3(X_1, X_2, X_3) &= 1 + X_1X_2X_3. \end{aligned}$$

With respect to the affine frame $(\overline{0}, \overline{1})$ in $\mathbb{A}$, we get the following four control points $\theta_0, \theta_1, \theta_2, \theta_3$:

$$\begin{aligned} \theta_0 &= (0, 0, 1) \\ \theta_1 &= (1, 0, 1) \\ \theta_2 &= (2, 1, 1) \\ \theta_3 &= (3, 3, 2). \end{aligned}$$

Letting $a = (1, 0)$, $b_2 = (2, 1)$, and $c_3 = \left(\frac{3}{2}, \frac{3}{2}\right)$, we have the control points $\theta_0 = \langle \Omega_1, 1\rangle$, $\theta_1 = \langle a, 1\rangle$, $\theta_2 = \langle b_2, 1\rangle$, and $\theta_3 = \langle c_3, 2\rangle$. One should figure out how the curve is traversed when t ranges from $-\infty$ to $+\infty$. It is rather strange.

It can be shown that the line $y + x + 1 = 0$ is an asymptote, and that the Folium of Descartes is also defined by the implicit equation

$$x^3 + y^3 - 3xy = 0.$$

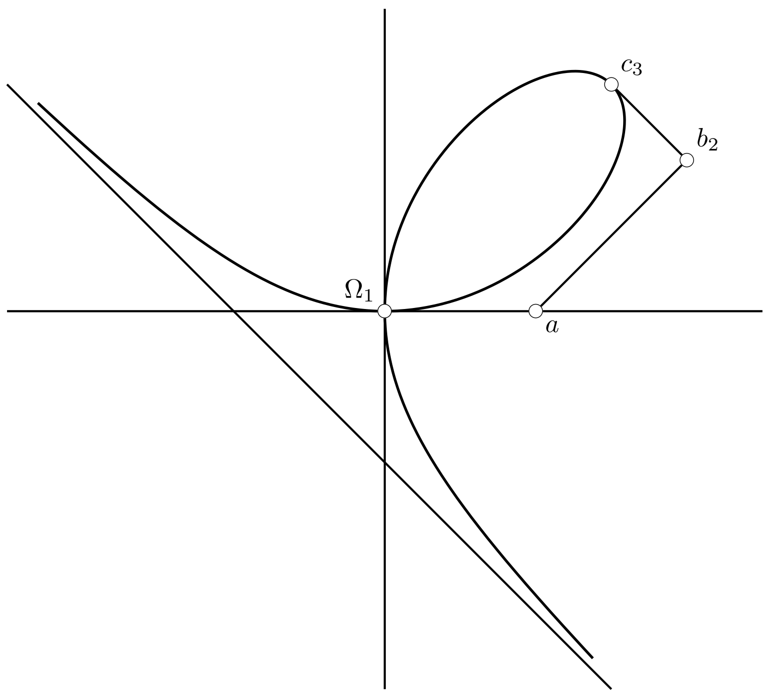


Figure 22.6: Folium of Descartes

It is sometimes convenient to have an explicit formula giving the current point $F(\bar{t})$ on a rational curve specified as a BR-curve.

22.2 Rational Curves and Bernstein Polynomials

Let us assume that F is specified by the $m+1$ control points $(\theta_0, \ldots, \theta_m)$, and that I and J are the sets of indices such that, $I \cup J = \{0, \ldots, m\}$, $I \cap J = \emptyset$, and for every $i \in I$,

$$\theta_i = \langle a_i, w_i \rangle$$

is a weighted point, where $a_i \in \mathcal{E}$ and $w_i \neq 0$, and for every $j \in J$,

$$\theta_j = u_j$$

is a control vector. Then, with respect to the affine frame $(\bar{0}, \bar{1})$ in $\mathbb{A}$, the polynomial curve

$$\mathcal{B}[\theta_0, \ldots, \theta_m] \colon \mathbb{A} \to \widehat{\mathcal{E}},$$

defined (in $\widehat{\mathcal{E}}$) by the control points $(\theta_0, \ldots, \theta_m)$, also abbreviated as $\mathcal{B}[\theta]$, is given in terms of the Bernstein polynomials as

$$\mathcal{B}[\theta](\bar{t}) = \sum_{0 \leq i \leq m} B_i^m(t)\, \theta_i.$$

Now, the above formula gives $\mathcal{B}[\theta](\overline{t})$ as a barycenter in $\widehat{\mathcal{E}}$, which, by lemma 4.1.2, can be expressed more explicitly. The explicit form of the barycenter depends on the quantity

$$w(t) = \sum_{i \in I} w_i B_i^m(t).$$

If $w(t) \neq 0$, then

$$\mathcal{B}[\theta](\overline{t}) = \left\langle \sum_{i \in I} \frac{w_i B_i^m(t)}{w(t)}\, a_i + \sum_{j \in J} \frac{B_j^m(t)}{w(t)}\, u_j, w(t) \right\rangle,$$

else if $w(t) = 0$, then

$$\mathcal{B}[\theta](\overline{t}) = \sum_{i \in I} w_i B_i^m(t)\, a_i + \sum_{j \in J} B_j^m(t) u_j,$$

where

$$\sum_{i \in I} w_i B_i^m(t)\, a_i = \sum_{i \in I} w_i B_i^m(t) \mathbf{ba_i}$$

for any $b \in \mathcal{E}$, which, by lemma 2.4.1, is a vector independent of b.

Then, since $F(\overline{t}) = \Pi(\mathcal{B}[\theta](\overline{t}))$, we have the following:

Letting

$$w(t) = \sum_{i \in I} w_i B_i^m(t),$$

if $w(t) \neq 0$, then

$$F(\overline{t}) = \sum_{i \in I} \frac{w_i B_i^m(t)}{w(t)}\, a_i + \sum_{j \in J} \frac{B_j^m(t)}{w(t)}\, u_j,$$

else if $w(t) = 0$ and $\sum_{i \in I} w_i B_i^m(t)\, a_i + \sum_{j \in J} B_j^m(t) u_j \neq 0$, then

$$F(\overline{t}) = \left(\sum_{i \in I} w_i B_i^m(t)\, a_i + \sum_{j \in J} B_j^m(t) u_j \right)_{\infty},$$

else if $w(t) = 0$ and $\sum_{i \in I} w_i B_i^m(t)\, a_i + \sum_{j \in J} B_j^m(t) u_j = 0$, then

$$F(\overline{t}) = \textit{undefined}.$$

In the third case, we can determine the value of $F(\overline{t})$ by continuity.

Remark: If $J = \emptyset$ and $w_i \geq 0$ for all $i, 0 \leq i \leq m$, then the trace $F([0,1])$ of F belongs to the convex hull of the points $a_0, \ldots, a_m$.

We close this section by discussing briefly the representation of the conics as rational curves, and some convenient changes of parameters.

An ellipse of implicit equation

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

has the parametric representation

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

A hyperbola of implicit equation

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

has the parametric representation

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

Of course, the parabola is a polynomial curve of degree 2. More generally, given a real conic defined by an implicit equation, if this conic is not empty (has some real point), taking any point on the conic as an origin, we can write its equation in the form

$$ax^2 + bxy + cy^2 + dx + ey = 0,$$

where a, b, c are not all null. Provided that d and e are not both null, we can find a rational parameterization by intersecting the conic with the line of equation $y = mx$, namely

$$x = \frac{-em-d}{cm^2+bm+a},$$
$$y = \frac{m(-em-d)}{cm^2+bm+a}.$$

The case $d = e = 0$ is easily handled (we get a single point, a double line, or two intersecting lines). For details on the conics as rational curves, we refer the reader to Fiorot and Jeannin [60], or Farin [58, 57].

Given a rational curve defined over $[0, +\infty[$, or the entire real line $\mathbb{A}$, it is sometimes useful to make a change of parameter to obtain a curve with the same trace, but parameterized over $[0, 1]$. In the first case, we can use the bijection

$$\psi(u) = \frac{\alpha u}{1-u},$$

with $\alpha > 0$, which is a strictly increasing function. By continuity, we let $\psi(1) = \infty$.

In the case of a curve parameterized over $\mathbb{A}$, we can use one of the maps

$$\varphi(u) = \frac{\alpha(1-t)^2 + 2\beta(1-t)t + \gamma t^2}{2t(1-t)},$$

where $\alpha\beta < 0$. Such a map is a bijection, strictly increasing when $\alpha < 0$, and strictly decreasing when $\alpha > 0$. Indeed,

$$\varphi(u) = \frac{(\alpha - 2\beta + \gamma)t^2 + 2(\beta - \alpha)t + \alpha}{2t(1-t)},$$

and after some calculations, we can show that

$$\varphi'(u) = \frac{(\gamma - 2\alpha)t^2 + 2\alpha t - \alpha}{4t^2(1-t)^2}.$$

To determine whether we can have

$$(\gamma - 2\alpha)t^2 + 2\alpha t - \alpha = 0$$

we compute the discriminant

$$\Delta = 4\alpha^2 + 4\alpha(\gamma - 2\alpha) = 4(\alpha\gamma - \alpha^2).$$

Thus, when $\alpha\gamma < 0$, the equation has no real roots, and we observe that $\varphi'(u) > 0$ when $\alpha < 0$, and $\varphi'(u) < 0$ when $\alpha > 0$, which means that φ is strictly increasing when $\alpha < 0$, and strictly decreasing when $\alpha > 0$.

When $\alpha < 0$, it is easily verified that $\lim_{u \to 0} \varphi(u) = -\infty$, and that $\lim_{u \to 1} \varphi(u) = +\infty$. We can extend φ so that it becomes a map $\varphi\colon [0,1] \to \widetilde{\mathbb{A}}$, by letting $\varphi(0) = \varphi(1) = \infty$. Then, φ is surjective, and a bijection on $]0,1[$.

By some appropriate changes of parameter, it is possible to parameterize an entire rational curve over the interval $[0,1]$, but this may raise the degree of the curve. For example, the full circle can be obtained as a curve of degree four, see Fiorot and Jeannin [60]. A very simple an effective alternative not involving any change of parameter will be presented in section 22.4.

We now consider generalizations of the de Casteljau algorithm to rational curves. This is actually quite easy.

22.3 Subdivision Algorithms for Rational Curves

There are two natural ways to extend the de Casteljau algorithm for polynomial curves to rational curves. The first method is to apply the de Casteljau algorithm in the homogenization $\widehat{\mathcal{E}}$ of the affine space $\mathcal{E}$, to find the point $G(\overline{t})$ on the polynomial curve $G\colon \mathbb{A} \to \widehat{\mathcal{E}}$ specified by the sequence of control points $(\theta_0, \ldots, \theta_m)$ in $\widehat{\mathcal{E}}$, and then project $G(\overline{t})$ onto $\widetilde{\mathcal{E}}$ using Π, getting $F(\overline{t}) = \Pi(G(\overline{t}))$. We just have to remember that when we perform linear interpolations, we must use the operations $\cdot$ and $\widehat{+}$ of the vector space $\widehat{\mathcal{E}}$. It is also possible that we get points at infinity when $G(\overline{t})$ is a vector in $\overrightarrow{\mathcal{E}}$, or even that $\Pi(G(\overline{t}))$ is undefined, when $G(\overline{t}) = \Omega$, the origin of $\widehat{\mathcal{E}}$. In this latter case, we can determine $F(\overline{t})$ by continuity.

The second method is to use the isomorphism $\widehat{\Omega}\colon \widehat{E} \to \mathcal{F}$ of lemma 4.3.2. We first map the sequence of control points $(\theta_0, \ldots, \theta_m)$ in $\widehat{\mathcal{E}}$ to the sequence of control points $(T_1, \ldots, T_m)$ in $\mathcal{F}$, where $T_i = \widehat{\Omega}(\theta_i)$, compute the point $P(\overline{t})$ on the polynomial curve $P = \widehat{\Omega}(G)$ in $\mathcal{F}$ from these control points, using the de Casteljau algorithm in $\mathcal{F}$, and then project $P(\overline{t})$ onto $\widetilde{\mathcal{E}}$, using $\Pi \circ \widehat{\Omega}^{-1}$. The advantage of the second method is that we only perform one division at the end. Although quite simple and effective in many cases, Farin [58] claims that this method is numericaly less stable when the weights of the control points vary significantly in magnitude. This seems to contradict the statement made by Fiorot and Jeannin [60, 61], that the second method is more stable! Clearly, more experimentation is needed!

We now present the first approach. The input is the sequence $(\theta_0, \ldots, \theta_m)$ of control points in $\widehat{\mathcal{E}}$. We assume that we have initialized the variables $\theta_{i,0}$, such that $\theta_{i,0} = \theta_i$, for all i, $0 \le i \le m$. The rational version of the de Casteljau algorithm is given below in pseudo-code.

begin
  for $j := 1$ **to** m **do**
    for $i := 0$ **to** $m - j$ **do**
      if $\theta_{i,j-1} = u_{i,j-1}$ **then**
        if $\theta_{i+1,j-1} = u_{i+1,j-1}$ **then**
          $\theta_{i,j} := (1-t)u_{i,j-1} + t u_{i+1,j-1}$
        else $\{\theta_{i+1,j-1} = \langle b_{i+1,j-1}, w_{i+1,j-1}\rangle\}$
          $w_{i,j} := t w_{i+1,j-1}$;
          $\theta_{i,j} := \left\langle b_{i+1,j-1} + \frac{(1-t)}{w_{i,j}} u_{i,j-1}, w_{i,j} \right\rangle$
        endif
      else $\{\theta_{i,j-1} = \langle b_{i,j-1}, w_{i,j-1}\rangle\}$
        if $\theta_{i+1,j-1} = u_{i+1,j-1}$ **then**
          $w_{i,j} := (1-t) w_{i,j-1}$;
          $\theta_{i,j} := \left\langle b_{i,j-1} + \frac{t}{w_{i,j}} u_{i+1,j-1}, w_{i,j} \right\rangle$
        else $\{\theta_{i+1,j-1} = \langle b_{i+1,j-1}, w_{i+1,j-1}\rangle\}$

$w_{i,\,j} = (1-t)w_{i,\,j-1} + tw_{i+1,\,j-1}$;
if $w_{i,\,j} = 0$ **then**
$\theta_{i,\,j} := (1-t)w_{i,\,j-1}b_{i,\,j-1} + tw_{i+1,\,j-1}b_{i+1,\,j-1}$ {a vector}
else
$\theta_{i,\,j} := \left\langle \frac{(1-t)w_{i,\,j-1}}{w_{i,\,j}} b_{i,\,j-1} + \frac{tw_{i+1,\,j-1}}{w_{i,\,j}} b_{i+1,\,j-1}, w_{i,\,j} \right\rangle$
endif
endif
endif
endfor
endfor ;
$F(\overline{t}) := \Pi(\theta_{0,\,m})$
end

To be more specific, the result of the algorithm, $F(\overline{t}) = \Pi(\theta_{0,\,m}) \in \widetilde{\mathcal{E}}$, is defined such that

$$\Pi(\theta_{0,\,m}) = \begin{cases} \textit{undefined} & \text{if } \theta_{0,\,m} = 0; \\ (u_{0,\,m})_\infty & \text{if } \theta_{0,\,m} = u_{0,\,m}, \text{ where } u_{0,\,m} \in (\overrightarrow{\mathcal{E}} - \{0\}); \\ b_{0,\,m} & \text{if } \theta_{0,\,m} = \langle b_{0,\,m}, w_{0,\,m} \rangle, \text{ where } b_{0,\,m} \in \mathcal{E}, \text{ and } w_{0,\,m} \neq 0. \end{cases}$$

Traditionally, most authors don't bother dealing with points at infinity, which happens when $\theta_{0,\,m} = u_{0,\,m}$ is a vector in $\overrightarrow{\mathcal{E}}$. When $\theta_{0,\,m} = 0$, it is possible to find $F(\overline{t})$ by continuity.

We now give the second version of the de Casteljau algorithm, given an affine frame $(\Omega_1, (e_1, \ldots, e_n))$ for $\mathcal{E}$, and the corresponding basis $(e_1, \ldots, e_n, \langle \Omega_1, 1 \rangle)$ for $\widehat{\mathcal{E}}$.

begin
for $i := 0$ **to** m **do**
if $\theta_i = \langle (x_1, \ldots, x_n), w \rangle$ **then**
$b_{i,\,0} := (wx_1, \ldots, wx_n, w)$
else $\{\theta_i = (x_1, \ldots, x_n)\}$
$b_{i,\,0} := (x_1, \ldots, x_n, 0)$
endif
endfor;
for $j := 1$ **to** m **do**
for $i := 0$ **to** $m - j$ **do**
$b_{i,\,j} := (1-t)b_{i,\,j-1} + tb_{i+1,\,j-1}$
endfor
endfor ;
let $(y_1, \ldots, y_n, w) = b_{0,\,m}$;
if $w = 0$ **then**
if $(y_1, \ldots, y_n) \neq 0$ **then**
$F(\overline{t}) := (y_1, \ldots, y_n)_\infty$
else
$F(\overline{t})$ *undefined*
endif
else
$F(\overline{t}) := \left(\frac{y_1}{w}, \ldots, \frac{y_n}{w} \right)$
endif
end

A weighted point $\langle a, w\rangle$, where a is a point in $\mathcal{E}$, is represented as $\langle (x_1, \ldots, x_n), w\rangle$, where $(x_1, \ldots, x_n)$ are the coordinates of $a \in \mathcal{E}$, over the affine frame $(\Omega_1, (e_1, \ldots, e_n))$ of $\mathcal{E}$, and a vector $u \in \overrightarrow{\mathcal{E}}$, is represented as $(x_1, \ldots, x_n)$, where $(x_1, \ldots, x_n)$ are the coordinates of the vector u, over the basis $(e_1, \ldots, e_n)$ of $\overrightarrow{\mathcal{E}}$. Again, the point $F(\bar{t})$ on the rational curve may be a point at infinity, when $w = 0$.

Remark: Since it can be shown that projections preserve tangents, when θ_0 and θ_1 are weighted points $\theta_0 = \langle a_0, w_0\rangle$, $\theta_1 = \langle a_1, w_1\rangle$, if $a_0 \neq a_1$, since (θ_0, θ_1) is the tangent to $\mathcal{B}[\theta]$ at θ_0, then (a_0, a_1) is the tangent to F at a_0. A similar property holds for θ_{m-1} and θ_m, when they are weighted points. Furthermore, when $\theta_0 = u_0$ is a control vector, and $\theta_1 = \langle a_1, w_1\rangle$ is a weighted point, then the line parallel to u_0 passing through a_1 is an asymptote to the curve. Also, the de Casteljau algorithm gives the tangent to F at $\bar{t}$, when it exists. For example, considering the first version of the de Casteljau algorithm, when $\theta_{0,\,m-1} = \langle b_{0,\,m-1}, w_{0,\,m-1}\rangle$ and $\theta_{1,\,m-1} = \langle b_{1,\,m-1}, w_{1,\,m-1}\rangle$ are weighted points and $b_{0,\,m-1} \neq b_{1,\,m-1}$, then $(b_{0,\,m-1}, b_{1,\,m-1})$ is the tangent to F at $\bar{t}$. The reader can easily figure out how the second algorithm yields the tangent to F at $\bar{t}$ when it exists.

Using the observation about affine interpolations in $\widehat{\mathcal{E}}$ and cross-ratios in $\widetilde{\mathcal{E}}$ made after definition 5.8.4, we can give a geometric interpretation of the above versions of the de Casteljau algorithm. We give such an interpretation for the first algorithm, but a similar interpretation is easily given for the second version (using the projection $\Pi \circ \widehat{\Omega}^{-1}$). The algorithm computes the weighted points $\theta_{i,\,j}$ in $\widehat{\mathcal{E}}$, according to the inductive formula

$$\theta_{i,\,j} = (1-t)\cdot \theta_{i,\,j-1} \,\widehat{+}\, t\cdot \theta_{i+1,j-1}.$$

Thus, letting

$$\gamma_{i,\,j} = \theta_{i,\,j-1} \,\widehat{+}\, \theta_{i+1,\,j-1},$$

we have the following cross-ratio

$$[p(\theta_{i,\,j-1}), p(\theta_{i+1,\,j-1}), p(\gamma_{i,\,j}), p(\theta_{i,\,j})] = \frac{1-t}{t},$$

for all j, $0 \leq j \leq m$, and all i, $0 \leq i \leq m-j$. When $\theta_{i+1,\,j-1} = \langle b_{i+1,\,j-1}, w_{i+1,\,j-1}\rangle$ and $\theta_{i,\,j-1} = \langle b_{i,\,j-1}, w_{i,\,j-1}\rangle$, assuming that $w_{i,\,j} = (1-t)w_{i,\,j-1} + tw_{i+1,\,j-1} \neq 0$, and $v_{i,\,j} = w_{i,\,j-1} + w_{i+1,\,j-1} \neq 0$, we have $\gamma_{i,\,j} = \langle g_{i,\,j}, v_{i,\,j}\rangle$ and $\theta_{i,\,j} = \langle b_{i,\,j}, w_{i,\,j}\rangle$, where

$$g_{i,\,j} = \frac{w_{i,\,j-1}}{w_{i,\,j-1} + w_{i+1,\,j-1}} b_{i,\,j-1} + \frac{w_{i+1,\,j-1}}{w_{i,\,j-1} + w_{i+1,\,j-1}} b_{i+1,\,j-1},$$

and

$$b_{i,\,j} = \frac{(1-t)w_{i,\,j-1}}{w_{i,\,j}} b_{i,\,j-1} + \frac{tw_{i+1,\,j-1}}{w_{i,\,j}} b_{i+1,\,j-1}.$$

Then, we have

$$[b_{i,\,j-1}, b_{i+1,\,j-1}, g_{i,\,j}, b_{i,\,j}] = \frac{1-t}{t},$$

for all j, $0 \leq j \leq m$, and all i, $0 \leq i \leq m-j$. This shows that $b_{i,\,j}$ is the unique point such that the cross-ratio of the points $(b_{i,\,j-1}, b_{i+1,\,j-1}, g_{i,\,j}, b_{i,\,j})$ is $\frac{1-t}{t}$.

Thus, we can say that the de Casteljau algorithm for polynomial curves computes points using ratios, whereas the de Casteljau algorithm for rational curves computes points using cross-ratios. When $j = 1$ and $0 \leq i \leq m-1$, the points

$$g_i = g_{i,\,1} = \frac{w_i}{w_i + w_{i+1}} b_i + \frac{w_{i+1}}{w_i + w_{i+1}} b_{i+1}$$

can be thought of as "shape parameters". Indeed (up to a scalar), they determine the weights w_i.

Figure 22.7 illustrates the computation of a point on a rational curve. We consider the quarter-circle C of example 1, and compute the point $C(1/2)$. Recall that the control points are $\theta_0 = \langle a, 1\rangle$, $\theta_1 = \langle b, 1\rangle$, and $\theta_2 = \langle c, 2\rangle$, where $a = (1, 0)$, $b = (1, 1)$, and $c = (0, 1)$.

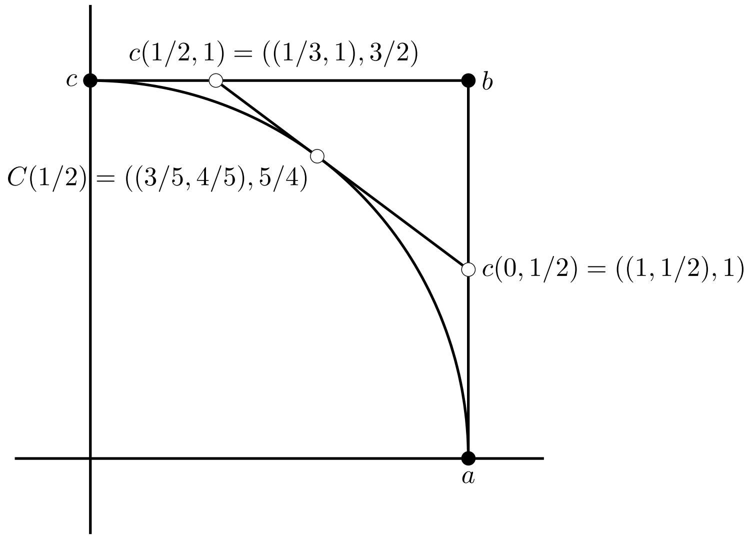


Figure 22.7: Construction of the point $C(1/2)$ on a quarter-circle

Using algorithm 2, we first determine $c(0,1/2)$ and $c(1/2,1)$, getting

$$c(0,1/2) = \left(1, \frac{1}{2}, 1\right), \qquad c(1/2,1) = \left(\frac{1}{2}, \frac{3}{2}, \frac{3}{2}\right),$$

which by projection, yields the points $\left(1, \frac{1}{2}\right)$ and $\left(\frac{1}{3}, 1\right)$. Finally, by interpolation between $c(0,1/2)$ and $c(1/2,1)$, we get $C(1/2) = \left(\frac{3}{4}, 1, \frac{5}{4}\right)$, which by projection, yields the point $\left(\frac{3}{5}, \frac{4}{5}\right)$. Note that the line segment from $c(0,1/2)$ to $c(1/2,1)$ is the tangent to the circle at $C(1/2)$.

It is easy to adapt the subdivision algorithm described for polynomial curves to the rational case. This can be done in two ways. Either we subdivide in $\widehat{\mathcal{E}}$, getting polylines determined by weighted points in $\widehat{\mathcal{E}}$, and project these control points down on $\mathcal{E}$ using Π. Some care must be exercised to avoid problems with points at infinity. For a fairly complete treatment of this method, the reader is referred to Fiorot and Jeannin [60, 61]. The second method, which we advocate since it seems simpler, is to use the isomorphism $\widehat{\Omega}\colon \widehat{\mathcal{E}} \to \mathcal{F}$ of lemma 4.3.2. Practically, what this means is that given a vector in $\widehat{\mathcal{E}}$ expressed as

$$u = (x_1, \ldots, x_n, w),$$

we have

$$\widehat{\Omega}(u) = (wx_1, \ldots, wx_n, w),$$

if $w \neq 0$, and

$$\widehat{\Omega}(u) = (x_1, \ldots, x_n, 0),$$

if $w = 0$.

We can then apply the subdivision method of section 18.2, but in $\mathcal{F}$. The following function send a control polygon consisting of points in $\widehat{\mathcal{E}}$ to a control polygon of points in $\mathcal{F}$.

```
(*  Maps a polygon to the hat space  by multiplying all coords except  *)
(*  the weight, by the weight  (for each point)                        *)
```

```
prepare[{poly__}] :=
 Block[
 {lpoly = {poly}, newp = {},  pt, h, w, i, l1},
 l1 = Length[lpoly];
 Do[
     pt = lpoly[[i]]; w = Last[pt]; h = Drop[pt, -1];
     If[w =!= 0, h = w * h; pt = Append[h, w]
       ];
     newp = Append[newp,pt],    {i, 1, l1}
   ];
 newp
];
```

Once a list of control polygons (in $\mathcal{F}$) has been computed using subdivision, it is necessary to map these control polygons down to the original affine space $\mathcal{E}$. This is performed by the function *proj* such that,

$$proj[(x_1, \ldots, x_n, w)] = (x_1/w, \ldots, x_n/w),$$

if $w \neq 0$,

$$proj[(x_1, \ldots, x_n, 0)] = (x_1, \ldots, x_n),$$

if some $x_i \neq 0$, which corresponds to a point at infinity, and

$$proj[(0, \ldots, 0, 0)] = \text{undefined}.$$

Concretely, we have to be careful in dividing by w when $|w|$ is very close to zero, since due to limited numerical precision, this may cause a division by zero. Thus, we only perform division when $|w| > 10^{-20}$, and we give a warning if $|w| \leq 10^{-16}$. Another problem is that a vector $(x_1, \ldots, x_n, w)$ may be very close to the zero vector. Thus, we use a function to test whether a vector is near zero, in which case this vector is ignored. This situation does occur in practice. For example, a torus is a surface of degree 4, but it contains circles which are of degree 2. When we use our algorithm to compute the control net of a circle on a torus, we get a control polygon of degree 4 with two degenerate zero entries. Actually, in the case of curves, it is not a problem to discard zero vectors (due to continuity), except that it may be necessary to iterate subdivision to get a better approximation. The following *Mathematica* functions implement these ideas.

```
(*   checks to see whether a vector is almost a zero vector     *)

  nearzero[a_] :=    Block[
  {stop, i, d, l},
   l = Length[a]; i = 1; stop = 1;
   While[stop === 1 && i <= l,
      d = a[[i]];
      If[Abs[d] < 10^(-16), i = i + 1, stop = 0];
        ];
    If[stop === 1, Print["*** Warning, Near Zero Vector: ", a, "  ***"]];
    stop
  ];

(* To project a point in the hat space back onto
   the affine space   *)
```

```
 aproj[{poly__}] :=
 Block[
 {pt = {poly}, res, h, w},
  w = Last[pt]; h = Drop[pt, -1];
      If[Abs[w] > 10^(-20),
         If[Abs[w] <= 10^(-16),
            Print["Warning: Point with very small
                   weight: ", pt]];  h = h/w,
         h = 10^(20) * h;  Print["*** Warning: Point at
         infinity!: ", pt, " ***"]
        ];
   h
 ];


(* To project a list of points in the hat space back onto
    the affine space   *)
(* points in the hat space which are almost the zero vector
   are discarded    *)

proj[{poly__}] :=
 Block[
 {sl = {poly}, np = {}, pt, h,  j, l2, flag},
     l2 = Length[sl];
     Do[
        pt = sl[[j]]; flag = nearzero[pt];
        (* If pt is almost the zero vector, discard   *)
        If[flag =!= 1, h = aproj[pt];  np = Append[np,h]],  {j, 1, l2}
       ];
 np
];


(* To project a list of control polygons in the hat space  *)
(* back onto the affine space                              *)

 projlis[{netlis__}] :=
 Block[
 {slis = {netlis}, newlis = {},
  anet, newnet, j, l2},
     l2 = Length[slis];
     Do[
        anet = slis[[j]]; newnet = proj[anet];
        newlis = Append[newlis,newnet],  {j, 1, l2}
       ];
 newlis
 ];
```

In order to display a curve segment using the subdivision function `subdiv`, use the following functions:

```
(*   Projects down the list of control polygons from the hat space *)
(* onto the affine space,    and creates the polyline.             *)

 tolineseg[{poly__}] :=
 Block[
 {pol = {poly}, res, newsl},
 (
  newsl = projlis[pol];
  res = makedgelis[newsl];
  res
 )
 ];


(*  To display the result of n steps of subdivision  *)
(*  curve segment over [r, s]                        *)


showsub2D[{poly__},r_,s_,lx_,mx_,ly_,my_,n_] :=    Block[
            {pol1, curv},
                       pol1 = prepare[{poly}];
                       curv = subdiv[pol1, r, s, n];
                       curv = tolineseg[curv];
                       Show[
                       Graphics[curv],
                       AspectRatio -> Automatic,
                       PlotRange -> {{lx, mx}, {ly, my}},
                       DisplayFunction -> $DisplayFunction]
   ];
```

The purpose of the parameters `lx, mx, ly, my`, is to define the size of the viewing window, which is $[lx, mx] \times [ly, my]$. We will now present a new method for approximating closed rational curves.

22.4 Approximating Closed Rational Curves

A common problem is to draw a closed rational curve. As usual, we assume that our curves are defined in some ambient affine space $\mathcal{E}$ of dimension at least 2. In order to explain the basic intuition behind the method presented in this Chapter as clearly as possible, we first consider rational curves specified explictly in terms of rational functions (as opposed to control points). For example, consider the following rational curve of degree 8 (a rose, see section 22.5) specified by the fractions

$$x = \frac{t(7 - 35t^2 + 21t^4 - t^6)}{(t^2 + 1)^4},$$
$$y = \frac{t^2(7 - 35t^2 + 21t^4 - t^6)}{(t^2 + 1)^4}.$$

The problem is that no matter how large the interval $[r, s]$ is, the trace $F([r, s])$ of F over $[r, s]$ is not the trace of the entire curve. In this particular example we could take advantage of symmetries, but in general, this may not be possible. There are rational bijections between $]-1, 1[$ and $\mathbb{R}$, for example, the map

$$t \mapsto \frac{t}{1 - t^2},$$

but they are at least quadratic, and cause the degree of the curve to be doubled, leading to inefficiency. Furthermore, if the curve is specified in terms of control points, it is rather complicated and expensive to compute the control points of the curve obtained after the change of variable.

A nice way to get around these problems is to observe that the function $t \mapsto \frac{1}{t}$ maps $]0,1]$ bijectively onto $[1,+\infty[$, and maps $[-1,0[$ bijectively over $]-\infty,-1]$. Thus, if we perform the change of variable $t = 1/u$, we get

$$x = \frac{u(7u^6 - 35u^4 + 21u^2 - 1)}{(u^2+1)^4},$$
$$y = \frac{7u^6 - 35u^4 + 21u^2 - 1}{(u^2+1)^4},$$

whose trace is identical to the trace of the curve F, but whose trace over $[-1,1]$ is the complement of the trace of F over $[-1,1]$. In particular, note that $G(0)$ corresponds to $F(\infty) = (0,-1)$. The method is general: given the fractions $Q_1(t)$ and $Q_2(t)$ defining F, we obtain the curve G by substituting $1/t$ for t in Q_1 and Q_2, getting the fractions $R_1(t) = Q_1(1/t)$ and $R_2(t) = Q_2(1/t)$, and we render F and G over $[-1,1]$ to render the entire trace of the curve F.

However, the above method assumes that the fractions defining the curve F are given explicitly. If the rational curve F is given by control points, it is necessary to first compute the fractions defining F, perform the substitution of $1/t$ for t, and then compute the control points for G. This is obviously a lot of work, and it may be computationally expensive.

Fortunately, there is a very simple (and cheap) way of getting the control points of G from the control points of F. Indeed, if $(\beta_0,\ldots,\beta_m)$ are the control points (in $\widehat{\mathcal{E}}$) of the BR-curve F w.r.t. the affine frame $(-1,1)$, the control points $(\theta_0,\ldots,\theta_m)$ (in $\widehat{\mathcal{E}}$) of the BR-curve G w.r.t. $(-1,1)$ are given by the equations

$$\theta_i = (-1)^i\, \beta_i.$$

Actually, it turns out that the above formula is valid for every affine frame (r,s)!

The upshot is that in order to render the entire trace of the curve F (over $[-\infty,+\infty]$), we just have to render both F and G over $[r,s]$. The frame $(-1,1)$ is just a special case (and so is $(0,1)$, a frame often used).

We will prove the correctness of the above formula using a simple geometric argument about ways of partitioning the real projective line into two disjoint segments. We will also show that other known methods for drawing rational curves can be justified using the above geometric argument.

The fact that the "complementary part" of a conic specified by three control points $((b_0,w_0),(b_1,w_1),(b_2,w_2))$ is defined by the control points $((b_0,w_0),(b_1,-w_1),(b_2,w_2))$ where the sign of the middle weight is flipped, has been shown by Lee [111] and Patterson [135]. Our result is a natural generalization to rational curves of arbitrary degree. Other methods for drawing closed rational curves have been investigated by Bajaj and Royappa [7], and by DeRose [45], who credits Patterson [135] for the original idea behind the method. We will compare these methods with ours after proving the correctness of our method.

We now show how the formula

$$\theta_i = (-1)^i\, \beta_i$$

is derived.

Recall that if $F\colon \widetilde{\mathbb{A}} \to \widetilde{\mathcal{E}}$ is a rational curve of degree m specified by some symmetric multilinear map $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$ such that

$$F([t,z]) = \mathbf{P}(f)(\underbrace{[t,z],\ldots,[t,z]}_{m})$$

for all $[t, z] \in \widetilde{\mathbb{A}}$, the control points $(\beta_0, \ldots, \beta_m)$ of the BR-curve G w.r.t. the affine frame (r, s) are given by the equations

$$\beta_i = f(\underbrace{r, \ldots, r}_{m-i}, \underbrace{s, \ldots, s}_{i}).$$

Since $\widetilde{\mathbb{A}} = \mathbb{RP}^1$, a rational curve is a map $F\colon \mathbb{RP}^1 \to \widetilde{\mathcal{E}}$ with domain the real projective line $\mathbb{RP}^1$, and to draw the entire trace of the curve F over $\mathbb{RP}^1$, we can partition $\mathbb{RP}^1$ into closed intervals $I_0, \ldots, I_p$ (intersecting only at boundary points) and find some simple projectivities $\varphi_1, \ldots, \varphi_p$ of $\mathbb{RP}^1$ such that $\varphi_i(I_0) = I_i$ $(1 \leq i \leq p)$. The simplest case arises when $p = 2$, and we described a way of partitioning $\mathbb{RP}^1$ where $I_0 = [-1, 1]$, $I_1 = \mathbb{RP}^1 -]-1, 1[$, and the projectivity $\varphi_1\colon \mathbb{RP}^1 \to \mathbb{RP}^1$

$$\varphi_1\colon t \mapsto \frac{1}{t}$$

is induced by the linear map $(u, v) \mapsto (v, u)$.

Now, recall from Chapter 5, Section 5.2, that the real projective line $\mathbb{RP}^1$ is obtained by identifying antipodal points on the circle S^1. However, up to homeomorphism, we can view $\mathbb{RP}^1$ as the result of identifying antipodal points on any closed convex polygon with central symmetry inscribed in the circle. Requiring central symmetry is not indispensible, but makes life easier since pairs of antipodal vertices are identified. The simplest convex polygons of this type are rectangles and squares. Thus, we obtain all the partitions of $\mathbb{RP}^1$ into two disjoint segments by projecting an inscribed rectangle (or square) onto some line not passing through the center, from the center of the circle. For instance, the case where $I_0 = [-1, 1]$ corresponds to a square, as illustrated in figure 22.8.

Case 1: $I_0 = [-1, 1]$.

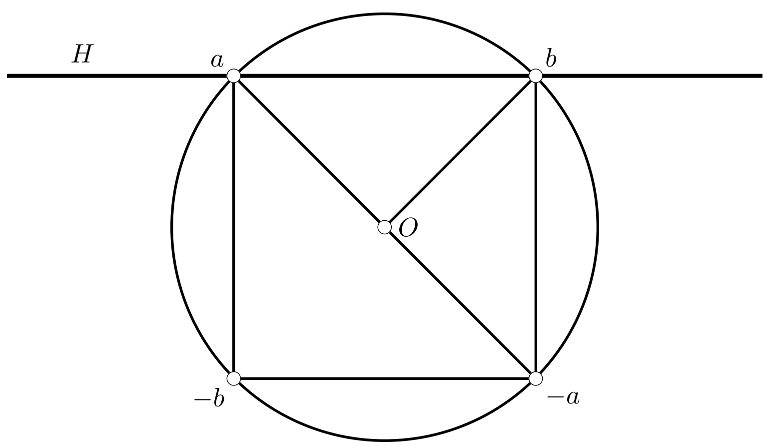


Figure 22.8: Model of the projective line, Case 1: $I_0 = [-1, 1]$

In this case, the line of projection H (of equation $y = 1$) contains an edge of the square (here (a, b)). Actually, this case applies to any affine frame $(-s, s)$, where $s \neq 0$. Letting $a = (-s, 1)$ and $b = (s, 1)$, it is trivial to verify that the linear map

$$(u, v) \mapsto \left(sv, \frac{u}{s}\right)$$

inducing the projectivity φ_1 is the unique linear map such that

$$\varphi_1(a) = -a, \quad \varphi_1(b) = b.$$

The points $(a, b, -a, -b)$ are the vertices of the inscribed square, and φ_1 maps the top edge (a, b) of the square onto the right edge $(-a, b)$. When a line L through the origin and passing through a point of the edge $(-a, b)$ varies, the intersection of L with the line H varies in $\varphi_1([-s, s])$.

We are now ready to tackle the general case.

Case 2: $I_0 = [r,\ s]$.

Given any affine frame $(r,\ s)$ (where $r < s$), we let the line H of equation $y = 1$ be the line of projection, we let $a = (r, 1)$ and $b = (s, 1)$ (points on H), and we define a rectangle $(c, b, -c, -b)$ inscribed in the circle C of center $O = (0, 0)$ and of radius $R = \sqrt{s^2 + 1}$ (so that b is on C), and a projectivity $\varphi\colon \mathbb{RP}^1 \to \mathbb{RP}^1$, as follows: c is the point on the upper half-circle defined as the intersection of the line (O, a) with C, and φ is the projectivity induced by the unique linear map such that

$$\varphi(a) = -a, \quad \varphi(b) = b.$$

Figure 22.9 illustrates the case where a is inside the circle.

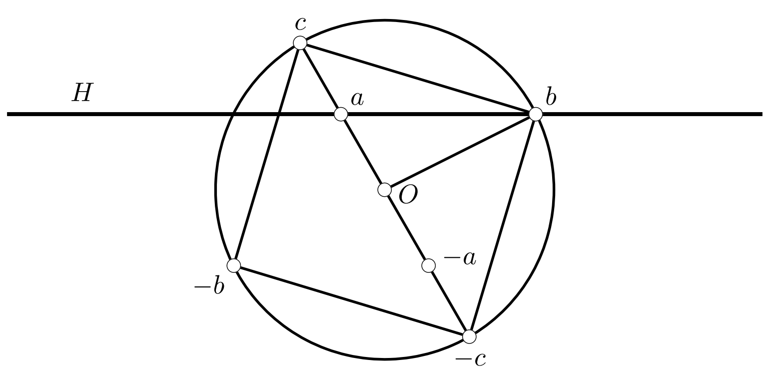


Figure 22.9: Model of the projective line, Case 2: $I_0 = [r, s]$

The rational curve G whose trace over $[r,\ s]$ is equal to the trace of F over $\varphi([r,\ s])$ is defined as follows.

Definition 22.4.1 For every affine frame (r, s) $(r < s)$ and every rational curve $F\colon \mathbb{RP}^1 \to \widetilde{\mathcal{E}}$ of degree m specified by some symmetric multilinear map $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$, the rational curve G is specified by the symmetric multilinear map $g\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$ such that

$$g((t_1, z_1), \ldots, (t_m, z_m)) = f(\varphi(t_1, z_1), \ldots, \varphi(t_m, z_m)),$$

where φ is the projectivity defined earlier.

Note that the point $F(\infty)$ corresponding to the point at infinity in $\mathbb{RP}^1$, is given by $G((0, 1))$. It should also be noted that it is quite possible that

$$f(\underbrace{(t, z), \ldots, (t, z)}_{m}) = 0$$

for some $(t, z) \neq (0, 0)$. In such a case, we have what is called a *base point*. This corresponds to the situation where the polynomials $F_1(t), \ldots, F_{n+1}(t)$ vanish simultaneously. Such situations arise in practice, for example after degree-raising. Another more devious situation where base points arise is when computing control nets of curves on rational surfaces, for example, a torus. We discovered this situation in drawing a torus in terms of u-curves and v-curves. Fortunately, in the case of curves, there is a simple remedy. Indeed, it is easy to justify using continuity and the fact that if polynomials in one variable vanish simultaneously, then they have a greatest common divisor, that bad points of the form $0 = (\underbrace{0, \ldots, 0}_{n+1})$ can simply be discarded. The price to pay is that it may be necessary to subdivide more in order to retain a proper level of visual smoothness.

Lemma 22.4.2 *For every affine frame (r, s) $(r < s)$ and every rational curve $F\colon \mathbb{RP}^1 \to \widetilde{\mathcal{E}}$ of degree m specified by some symmetric multilinear map $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$, if $g\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$ is the symmetric multilinear map of definition 22.4.1, except for base points, F and G have the same trace. In particular, the trace $G([r, s])$ is the union of the traces $F([-\infty, r])$ and $F([s, +\infty])$. Furthermore, if $(\beta_0, \ldots, \beta_m)$ are the control points (in $\widehat{\mathcal{E}}$) of F w.r.t. the affine frame (r, s), the control points $(\theta_0, \ldots, \theta_m)$ (in $\widehat{\mathcal{E}}$) of the curve G w.r.t. (r, s) are given by the equations*

$$\theta_i = (-1)^i\, \beta_i.$$

Proof. We have

$$g((t_1, z_1), \ldots, (t_m, z_m)) = f(\varphi(t_1, z_1), \ldots, \varphi(t_m, z_m)),$$

and thus

$$\mathbf{P}(g)([t_1, z_1], \ldots, [t_m, z_m]) = \mathbf{P}(f)([\varphi(t_1, z_1)], \ldots, [\varphi(t_m, z_m)]).$$

Now, since φ is a bijection from $[r, s]$ to $\mathbb{RP}^1 - \,]r, s[$, F and G have the same trace, and the trace $G([r, s])$ is the union of the traces $F([-\infty, r])$ and $F([s, +\infty])$. Finally, the control points θ_i of G w.r.t. (r, s) are given by

$$\theta_i = g(\underbrace{a, \ldots, a}_{m-i}, \underbrace{b, \ldots, b}_{i}),$$

and since

$$g((t_1, z_1), \ldots, (t_m, z_m)) = f(\varphi(t_1, z_1), \ldots, \varphi(t_m, z_m)),$$

$\varphi(a) = -a$, and $\varphi(b) = b$, we get $\theta_i = f(\underbrace{-a, \ldots, -a}_{m-i}, \underbrace{b, \ldots, b}_{i})$, that is

$$\theta_i = (-1)^{m-i} f(\underbrace{a, \ldots, a}_{m-i}, \underbrace{b, \ldots, b}_{i}) = (-1)^{m-i}\, \beta_i.$$

However, the points $(-1)^{m-i}\, \beta_i$ and $(-1)^i\, \beta_i$ have the same projection under Π, so we might as well use the simpler expression $(-1)^i\, \beta_i$. $\square$

Remark. The above proof has the advantage that it does not require an explicit computation of the projectivity φ, but of course, an explicit formula for φ can be found. The projectivity φ is induced by the linear map

$$(t, z) \mapsto \left(\frac{(s + r)t}{s - r} - \frac{2rsz}{s - r}, \frac{2t}{s - r} - \frac{(s + r)z}{s - r} \right),$$

so that

$$\varphi(t) = \frac{(s + r)t - 2rs}{2t - (s + r)}.$$

If we recall that the Bernstein polynomials of degree m over $[r, s]$ are given by

$$B_i^m[r, s](t) = \binom{m}{i} \left(\frac{s - t}{s - r} \right)^{m-i} \left(\frac{t - r}{s - r} \right)^{i},$$

it is easily shown that under the change of variable φ, we get

$$\frac{s-\varphi(t)}{s-r} = -\frac{s-t}{2t-(s+r)}, \quad \text{and} \quad \frac{\varphi(t)-r}{s-r} = \frac{t-r}{2t-(s+r)}.$$

Noting the presence of the minus sign in the first expression, we get

$$B_i^m[r,s](\varphi(t)) = \frac{(-1)^{m-i}(s-r)^m}{(2t-(s+r))^m} \binom{m}{i} \left(\frac{s-t}{s-r}\right)^{m-i} \left(\frac{t-r}{s-r}\right)^i,$$

that is

$$B_i^m[r,s](\varphi(t)) = \frac{(-1)^{m-i}(s-r)^m}{(2t-(s+r))^m} B_i^m[r,s](t).$$

From this and the fact that a rational curve F can be expressed as

$$F(t) = \sum_{i=0}^m \frac{w_i B_i^m[r,s](t)}{w(t)}\, b_i,$$

where we assumed that the control points are of the form $\langle b_i, w_i\rangle$ (with $w_i \neq 0$) and where

$$w(t) = \sum_{i=0}^m w_i B_i^m[r,s](t),$$

we get

$$F(\varphi(t)) = \sum_{i=0}^m \frac{(-1)^{m-i} w_i B_i^m[r,s](t)}{w(\varphi(t))}\, b_i,$$

where

$$w(\varphi(t)) = \sum_{i=0}^m (-1)^{m-i} w_i B_i^m[r,s](t),$$

and we obtain another proof our result (since we can multiply both the numerator and the denominator by $(-1)^{m+2i} = (-1)^m$, getting weights of the form $(-1)^i w_i$). Note that the above proof does not account for control vectors, but this can be done too. Of course, we prefer the geometric proof of Lemma 22.4.2 to this more computational proof, which, in our opinion, obsures what's really going on!

Lemma 22.4.2 shows that F and G have the same trace. It also shows that if the control points (in $\widehat{\mathcal{E}}$) of F w.r.t. (r,s) are $(\beta_0,\ldots,\beta_m)$, if β_i is of the form (a,w), where a is a point in $\mathcal{E}$ and $w \neq 0$ is a weight, then

$$\theta_i = (a,\ (-1)^i\, w),$$

and if β_i is a control vector $u \in \overrightarrow{\mathcal{E}}$, then

$$\theta_i = (-1)^i\, u.$$

The upshot is that in order to render the entire trace of the curve F (over $[-\infty,+\infty]$), it is enough to render both F and G over $[r,s]$, and the computation of the control points of G from those of F over (r,s) is very simple. For example, in the case of the ellipse F specified by the fractions

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

the control points w.r.t. $(-1, 1)$ are

$$\beta_0 = ((-2, 3), 2),\ \beta_1 = ((0, 1), 0),\ \beta_2 = ((2, 0), 2),$$

where β_1 is a control vector, and the control points $\theta_0, \theta_1, \theta_2$ are

$$\theta_0 = ((-2, 3), 2),\ \theta_1 = ((0, -1), 0),\ \theta_2 = ((2, 0), 2).$$

Bajaj and Royappa have investigated another method for drawing closed rational curves (and more generally, rational varieties) [7]. Their method is based on the observation that the maps $\psi_1 : t \mapsto \frac{t}{1-t}$ and $\psi_2 : t \mapsto \frac{-t}{1-t}$ map $[0, 1[$ bijectively onto $[0, +\infty[$ and $[0, -\infty[$ respectively. There is a simple geometric explanation for the choice of their projectivities. This case corresponds to a square and to the choice where the line of projection H (of equation $y = 1$) passes through a vertex and is perpendicular to one of the main diagonals of the square as shown in figure 22.10.

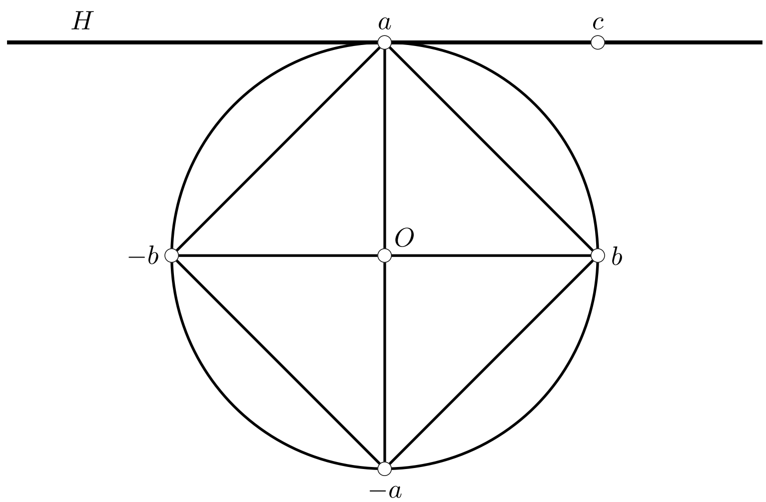


Figure 22.10: Model of the projective line, Case 3

However, Bajaj and Royapa do not consider the problem of computing the control points of the curves $F\left(\frac{t}{1-t}\right)$ and $F\left(\frac{-t}{1-t}\right)$. This can be done, but the resulting method is more complicated than ours. This is because the projectivities ψ_1 and ψ_2 do not map $c = (1, 1)$ to an already existing vertex.

Another method for drawing rational curves is due to DeRose [45]. Basically, the method consists in using the homogeneous Bernstein polynomials $\binom{m}{k} u^k v^{m-k}$ and to view a rational curve as a rational map from the projective line. Then, by using any $2D$ model of the projective line, it is possible to draw a closed rational curve in one piece. The model used by DeRose is the C^0-continuous curve in $\mathbb{A}^2$ defined such that $t \mapsto (t,\ 1 - |t|)$ over $[-1, 1]$. This curve is a model of the projective line in $\mathbb{R}^2$ (identifying the points $(-1, 0)$ and $(1, 0)$). In fact, this corresponds to Case 3 above! We can draw the closed rational curve F in a single piece as the trace of $F([u, 1 - |u|])$.

The advantage of DeRose's method is that it does not require a new control polygon, as in our method (although, computing this new control polygon is very simple, as we showed). The disadvantage is that it requires sampling some model of the projective line, and that the subdivision version of the standard De Casteljau algorithm cannot be used. Futher computer experimentation seems needed to compare the two methods.

The *Mathematica function* computing the control polygon θ (in the hat space) from the original control polygon (in the hat space) is shown below.

```
(*  computes the weights of the complementary part of a  closed *)
(* curve, i.e. the part corresponding to t outside ]r, s[ *)

  negpoly[{poly__}] :=
  Block[
  {cpoly = {poly}, l1, i, pol, pt, w, h},
                     pol = {}; l1 = Length[cpoly];
                     Do[ pt = cpoly[[i]];
                         If[OddQ[l1 - i], pt = -pt];
                         pol = Prepend[pol, pt],  {i, 1, l1}
                       ];
  pol
  ];
```

Finally, we can render a closed rational curve using the following functions. The function `rendclo2D` is given in the $2D$ case, but the $3D$ case is just as easy.

```
(* To create a list of line segments from a list of control polygons *)
 makedgelis[{poly__}] :=
 Block[
 {res, sl,  newsl = {poly},
  i, j, l1, l2},
  (l1 = Length[newsl];  res = {};
  Do[
    sl = newsl[[i]]; l2 = Length[sl];
    Do[If[j > 1, res = Append[res, Line[{sl[[j-1]], sl[[j]]}]]], {j, 1, l2}
      ], {i, 1, l1}
   ];
  res
 )
 ];

(*   Projects down the list of control polygons from the hat space *)
(* onto the affine space,    and creates the polyline.            *)

 tolineseg[{poly__}] :=
 Block[
 {pol = {poly}, res, newsl},
 (
  newsl = projlis[pol];
  res = makedgelis[newsl];
  res
 )
 ];


(* computes the control polygon for display.   *)
(*  Control vectors cause extra trouble        *)
```

```
controlfig2D[{poly__}] :=
 Block[
 {p = {poly}, polylin, cpoly, cvec, cpt, conpt, res,
  totnum, nodenum, ptnum, vecnum, w, i, n, pt, orig = 1},
  n := Length[p];
  polylin = {}; cpoly = {}; cvec = {}; cpt = {}; conpt = {};
  totnum = 0; nodenum = 0; ptnum = 0; vecnum = 0;
  Do[w := Last[p[[i]]];
     If[w === 0, pt = Drop[p[[i]],-1] + orig;
                 cvec = Append[cvec, Line[{{orig, orig}, pt}]];
                 pt = pt + {0.1,0};
                 cvec = Append[cvec, Text["v", pt, {-1,0}]];
                 vecnum = vecnum + 1;
                 If[nodenum > 1, polylin = Append[polylin, Line[cpoly]];
                            totnum = totnum + 1; nodenum = 0; cpoly = {}
                               , nodenum = 0; cpoly = {}]
                ,
                 pt = Drop[p[[i]],-1];
                 cpt = Append[cpt, Point[pt]];
                 ptnum = ptnum + 1;
                 cpoly = Append[cpoly, pt];
                 nodenum = nodenum + 1
        ], {i, 1, n}
     ];
  If[nodenum > 1, polylin = Append[polylin, Line[cpoly]];
                  totnum = totnum + 1; nodenum = 0; cpoly = {}
                , nodenum = 0; cpoly = {}];
  If[cvec =!= {}, cvec = Append[cvec, Point[{orig,orig}]]];
  conpt = Prepend[cpt, PointSize[0.01]];
  res = Join[conpt,cvec,polylin]; res = Prepend[res, RGBColor[0,1,0]];
  res
 ];

(*  rfig2 works as controlfig2D, except that it keeps track of the list  *)
(*  of weights in wll, and appends it as second argument                 *)

 rfig2D[{poly__}] :=
 Block[
 {p = {poly}, polylin, cpoly, cvec, cpt, conpt, res,
  totnum, nodenum, ptnum, vecnum, w, i, n, pt, wll, orig = 1},
  n := Length[p];
  polylin = {}; cpoly = {}; cvec = {}; cpt = {}; conpt = {}; wll = {};
  totnum = 0; nodenum = 0; ptnum = 0; vecnum = 0;
  Do[w := Last[p[[i]]]; wll = Append[wll, w];
     If[w === 0, pt = Drop[p[[i]],-1] + orig;
                 cvec = Append[cvec, Line[{{orig, orig}, pt}]];
                 pt = pt + {0.1,0};
                 cvec = Append[cvec, Text["v", pt, {-1,0}]];
                 vecnum = vecnum + 1;
                 If[nodenum > 1, polylin = Append[polylin, Line[cpoly]];
```

```
                              totnum = totnum + 1; nodenum = 0; cpoly = {}
                                , nodenum = 0; cpoly = {}]
                ,
                 pt = Drop[p[[i]],-1];
                 cpt = Append[cpt, Point[pt]];
                 ptnum = ptnum + 1;
                 cpoly = Append[cpoly, pt];
                 nodenum = nodenum + 1
        ], {i, 1, n}
     ];
  If[nodenum > 1, polylin = Append[polylin, Line[cpoly]];
                  totnum = totnum + 1; nodenum = 0; cpoly = {}
                , nodenum = 0; cpoly = {}];
  If[cvec =!= {}, cvec = Append[cvec, Point[{orig,orig}]]];
  conpt = Prepend[cpt, PointSize[0.01]];
  res = Join[conpt,cvec,polylin]; res = Prepend[res, RGBColor[1,0,0]];
  res = Append[res, wll];
  res
 ];

(*  projects  control polygon in hat space back onto affine space   *)
(*  keeps the weights                                                *)

 newpoly[{cpoly__}, r_, s_, l_, m_] :=
 Block[
 {res = {cpoly}, l1, i, npt, pt, newpt, w},
  l1 = Length[res]; newp = {};
  Do[
    w = Last[res[[i]]]; pt = Drop[res[[i]],-1];
    If[w =!= 0, pt = pt/w; npt = Append[pt, w], npt = res[[i]]];
    newp = Append[newp, npt], {i, 1, l1}
    ];
  newp
 ];

 (* to display a closed rational 2D curve.    *)
 (* If flagpol =!= 1, do not display control polygon  *)
 (* If flagpol === -1, show original arc in black,
    and dual arc in red *)

  rendclo2D[{poly__},r_,s_,lx_,mx_,ly_,my_,flagpol_,n_] :=
  Block[
  {cpoly = {poly}, np1, np2, npol1, npol2, wl1, wl2, l, m, pol,
   curv1, curv2, pol1, pol2, image},
                      l = -1; m = 1;
        (*  maps control poly to hat space                *)
                      pol1 = prepare[cpoly];
        (* computes control polygon wrt  [-1, 1]          *)
                      np1 = newcpoly[pol1, r, s, l, m];
        (* projects down to affine space           *)
                      npol1 =  newpoly[np1, r, s, l, m];
```

```
                    npol1 =  rfig2D[npol1];
                    wl1 = Last[npol1]; npol1 = Drop[npol1,-1];
                    Print["Weights: ", wl1];
                    Print["New Cont. Poly.: ", np1];
        (*  computes dual control polygon                *)
                    pol2 =  negpoly[np1];
                    np2 =  newpoly[pol2, r, s, l, m];
                    npol2 =  rfig2D[np2];
                    wl2 = Last[npol2]; npol2 = Drop[npol2,-1];
                    Print["Dual Weights: ", wl2];
                    Print["Dual Cont. Poly.: ", pol2];
        (* subdivides  *)
                    curv1 = subdiv[np1, l, m, n];
                    curv1 = tolineseg[curv1];
                    Print["First curve segment done! "];
                    curv2 = subdiv[pol2, l, m, n];
                    curv2 = tolineseg[curv2];
                    Print["Second curve segment done! "];
                    Print["Ready to display! "];
                    If[flagpol === 1,
                        image = {controlfig2D[{poly}], npol1, curv1,
                                 {RGBColor[1,0,0], curv2}},
                                 If[flagpol === -1,
                             image = {curv1, {RGBColor[1,0,0], curv2}},
                             image = {curv1, curv2}
                                   ]
                    ];
                    Show[Graphics[{image}],
                    AspectRatio -> Automatic,
                    PlotRange -> {{lx, mx}, {ly, my}},
                    DisplayFunction -> $DisplayFunction];
    ];
```

The procedure rendclo2D assumes that it is given a control polygon over the affine frame $(0, 1)$, and it computes a new control polygon over $(-1, 1)$. It is very easy to modify this procedure so that it takes as input a control polygon over any frame (r, s), omitting the computation of the new control polygon w.r.t. the frame $(-1, 1)$. It is also immediate to write output procedures for the $3D$ case.

The method is illustrated in Figure 22.11 on the quartic given by the following control polygon w.r.t $(0, 1)$:

```
inpol = {{0,0,1},{2,6,1},{6,8,2},{10,4,1},{10,0,1}}
```

In this example, we did not compute a new control polygon w.r.t. $(-1, 1)$.

A remarkable feature of our method is that it works even for nonclosed rational curves with branches going to infinity. Indeed, what happens is that *Mathematica* clips the parts of the curve outside of the viewing window! For example, Figure 22.12 shows the cubic given by the following control polygon over $(0, 1)$:

```
inpol = {{0,0,1}, {2,6,1}, {6,8,2}, {10,0,1}}
```

In the above example, a new control polygon over $(-1, 1)$ was computed. More examples are given in the next section.

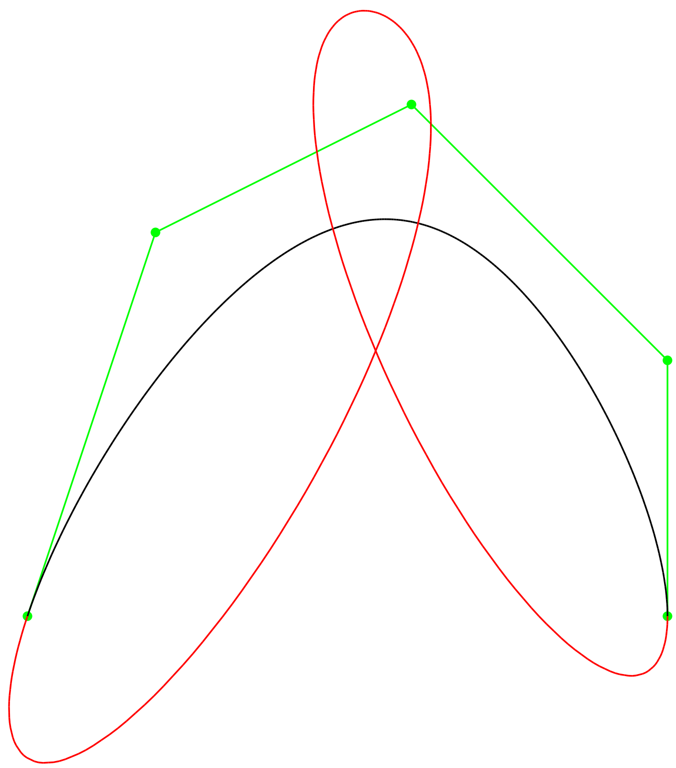


Figure 22.11: A Closed Quartic

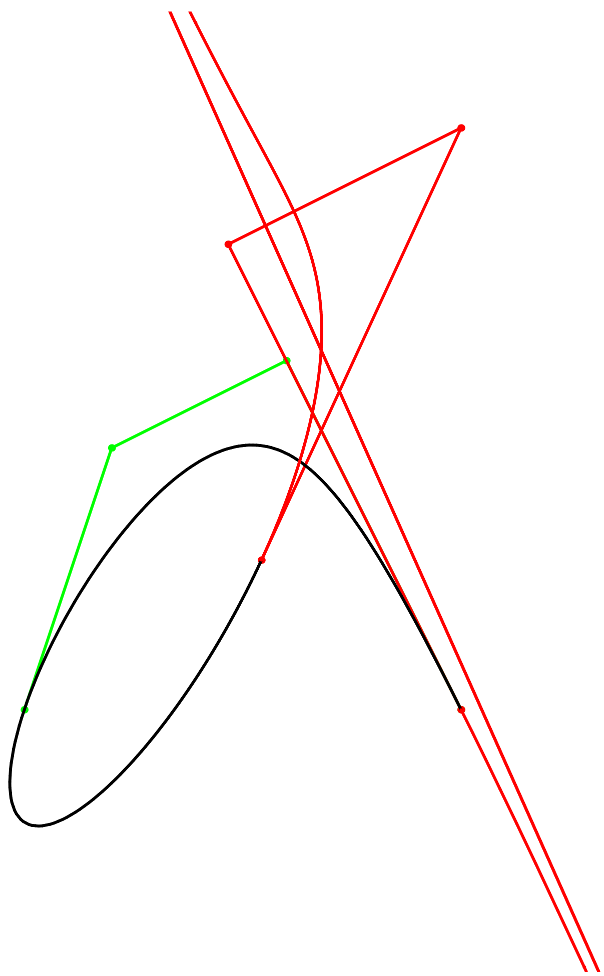


Figure 22.12: A Cubic with Asymptote

22.5 A "Gallery" of Rational Curves

The purpose of this section is to illustrate the power and flexibility of the method of the previous section, as well as take a stroll in a small gallery of the immense museum of rational curves. We wish you a happy stroll.

It will become clear that polarizing polynomials and computing control polygons by hand is very painful. Thus, we wrote *Mathematica* programs doing this, and we urge the readers to do the same.

We begin with a classic, known as the "Lemniscate of Bernoulli",

Example 1. The "Lemniscate of Bernoulli" is a quartic defined by the polynomials

$$\begin{aligned} G_1(X) &= X + X^3 \\ G_2(X) &= X - X^3 \\ G_3(X) &= 1 + X^4. \end{aligned}$$

We get the polar forms

$$\begin{aligned} g_1(X_1, X_2, X_3, X_4) &= \frac{X_1 + X_2 + X_3 + X_4}{4} + \frac{X_1X_2X_3 + X_1X_2X_4 + X_1X_3X_4 + X_2X_3X_4}{4} \\ g_2(X_1, X_2, X_3, X_4) &= \frac{X_1 + X_2 + X_3 + X_4}{4} - \frac{X_1X_2X_3 + X_1X_2X_4 + X_1X_3X_4 + X_2X_3X_4}{4} \\ g_3(X_1, X_2, X_3, X_4) &= 1 + X_1X_2X_3X_4. \end{aligned}$$

With respect to the affine frame $(0, 1)$ in $\mathbb{A}$, we get the following five control points θ_0, θ_1, θ_2, θ_3, θ_4:

$$\begin{aligned} \theta_0 &= (0, 0, 1) \\ \theta_1 &= \left(\frac{1}{4}, \frac{1}{4}, 1\right) \\ \theta_2 &= \left(\frac{1}{2}, \frac{1}{2}, 1\right) \\ \theta_3 &= \left(1, \frac{1}{2}, 1\right) \\ \theta_4 &= (2, 0, 2). \end{aligned}$$

Letting $a_3 = \left(\frac{1}{4}, \frac{1}{4}\right)$, $a_4 = \left(\frac{1}{2}, \frac{1}{2}\right)$, $a_5 = \left(1, \frac{1}{2}\right)$, and $a = (1, 0)$, we have the control points $\theta_0 = \langle \Omega_1, 1\rangle$, $\theta_1 = \langle a_3, 1\rangle$, $\theta_2 = \langle a_4, 1\rangle$, $\theta_3 = \langle a_5, 1\rangle$, and $\theta_4 = \langle a, 2\rangle$. When $t \in [0, 1]$, the point $F(t) = \Pi(G(t))$ travels on the curve segment between the origin and a (a fourth of the curve).

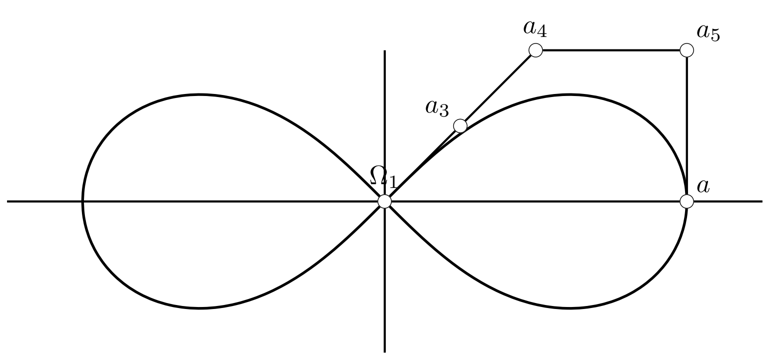


Figure 22.13: Lemniscate of Bernoulli

The Lemniscate of Bernoulli is also defined by the implicit equation

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

Example 2. Consider the following rational quartic

$$\begin{aligned}
G_1(t) &= t(3 - t^2),\\
G_2(t) &= t^2(3 - t^2),\\
G_3(t) &= (1 + t^2)^2.
\end{aligned}$$

The implicit equation of the above quartic is

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

If we parameterize the curve in terms of $u = 1/t$, we get

$$\begin{aligned}
G_1(u) &= u(3u^2 - 1),\\
G_2(u) &= 3u^2 - 1,\\
G_3(u) &= (1 + u^2)^2.
\end{aligned}$$

The curve displayed in Figure 22.14 looks like a three-leafed rose.

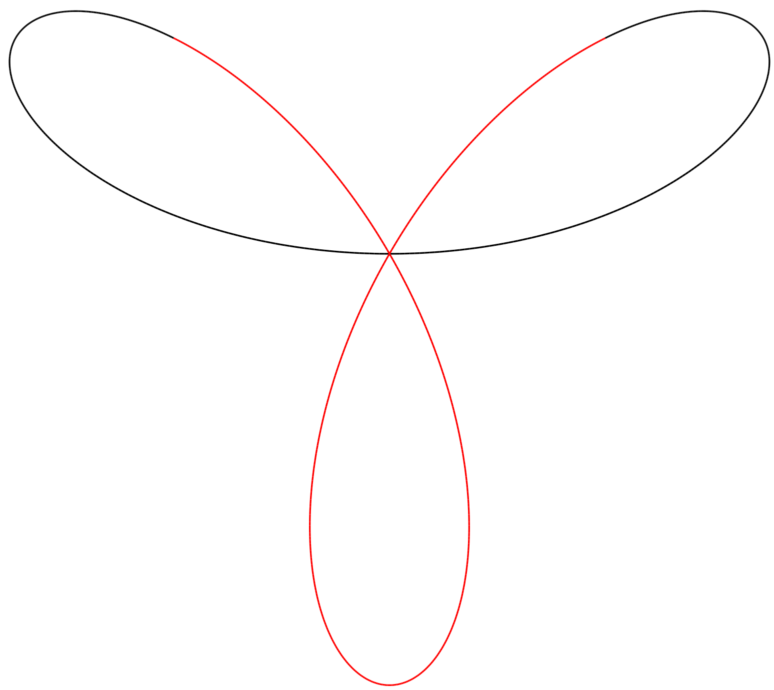


Figure 22.14: A three-leafed rose (quartic)

The curve is invariant under a rotation by $2\pi/3$. We leave as exercise to show that the control polygon w.r.t. $(0, 1)$ is

```
cpoly =  {{0, 0, 1}, {3/4, 0, 1}, {9/8, 3/8, 4/3},
          {1, 3/4, 2}, {1/2, 1/2, 4}}
```

The next four examples show some pretty and perhaps surprising curves. The details of the construction of these curves are left as instructive exercises.

Example 3. Consider the following rational curve of degree 6:

$$
\begin{aligned}
G_1(t) &= 4t(1-t^2)^2,\\
G_2(t) &= 8t^2(1-t^2),\\
G_3(t) &= (1+t^2)^3.
\end{aligned}
$$

The curve shown in Figure 22.15 looks like a four-leafed rose.

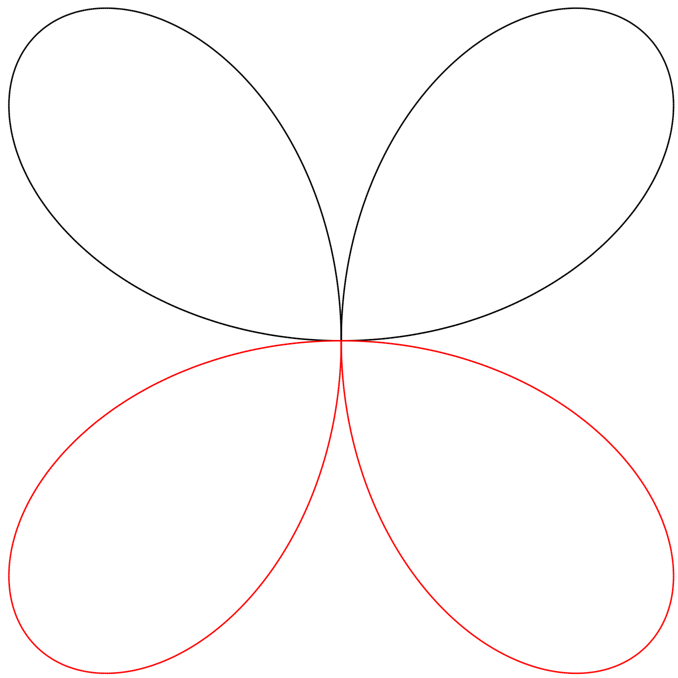


Figure 22.15: A four-leafed rose (sextic)

The implicit equation of the above sextic is

$$(x^2 + y^2)^3 = 4x^2y^2,$$

and its control polygon w.r.t. $(0, 1)$ is

```
cpoly = {{0, 0, 1}, {2/3, 0, 1}, {10/9, 4/9, 6/5},
         {1, 1, 8/5}, {4/9, 10/9, 12/5}, {0, 2/3, 4}, {0, 0, 8}}
```

Example 4. Consider the following rational curve of degree 6:

$$\begin{aligned}
G_1(t) &= t(5 - 10t^2 + t^4),\\
G_2(t) &= t^2(5 - 10t^2 + t^4),\\
G_3(t) &= (t^2 + 1)^3.
\end{aligned}$$

The implicit equation of the above sextic is

$$(x^2 + y^2)^3 - 5x^4y + 10x^2y^3 - y^5 = 0,$$

and its control polygon w.r.t. $(0, 1)$ is

```
cpoly = {{0, 0, 1}, {2/3, 0, 1}, {10/9, 4/9, 6/5},
         {1, 1, 8/5}, {4/9, 10/9, 12/5}, {0, 2/3, 4}, {0, 0, 8}}
```

The curve looks like a five-leafed rose.

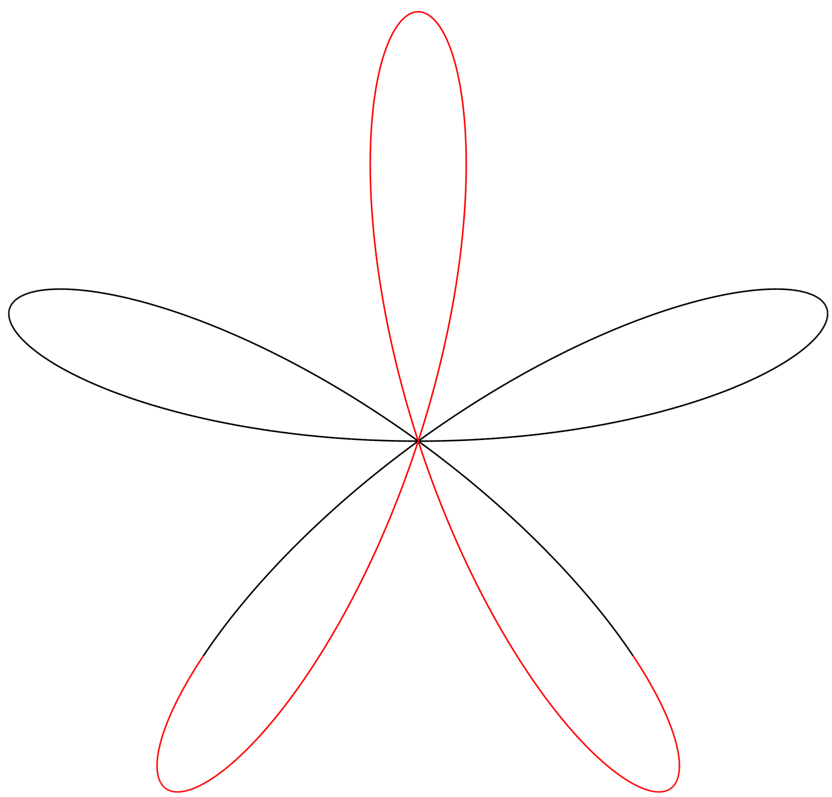


Figure 22.16: A five-leafed rose (sextic)

Example 5. Consider the following rational curve of degree 8:

$$\begin{aligned} G_1(t) &= t(7 - 35t^2 + 21t^4 - t^6), \\ G_2(t) &= t^2(7 - 35t^2 + 21t^4 - t^6), \\ G_3(t) &= (t^2 + 1)^4. \end{aligned}$$

The implicit equation of the above curve is

$$(x^2 + y^2)^4 - 7x^6y + 35x^4y^3 - 21x^2y^5 + y^7 = 0,$$

and its control polygon w.r.t. $(0, 1)$ is

```
cpoly =  {{0, 0, 1}, {7/8, 0, 1}, {49/32, 7/32, 8/7}, {7/5, 21/40, 10/7},
         {35/68, 35/68, 68/35}, {-21/40, 0, 20/7}, {-35/32, -21/32, 32/7},
         {-1, -7/8, 8}, {-1/2, -1/2, 16}}
```

The curve looks like a seven-leafed rose.

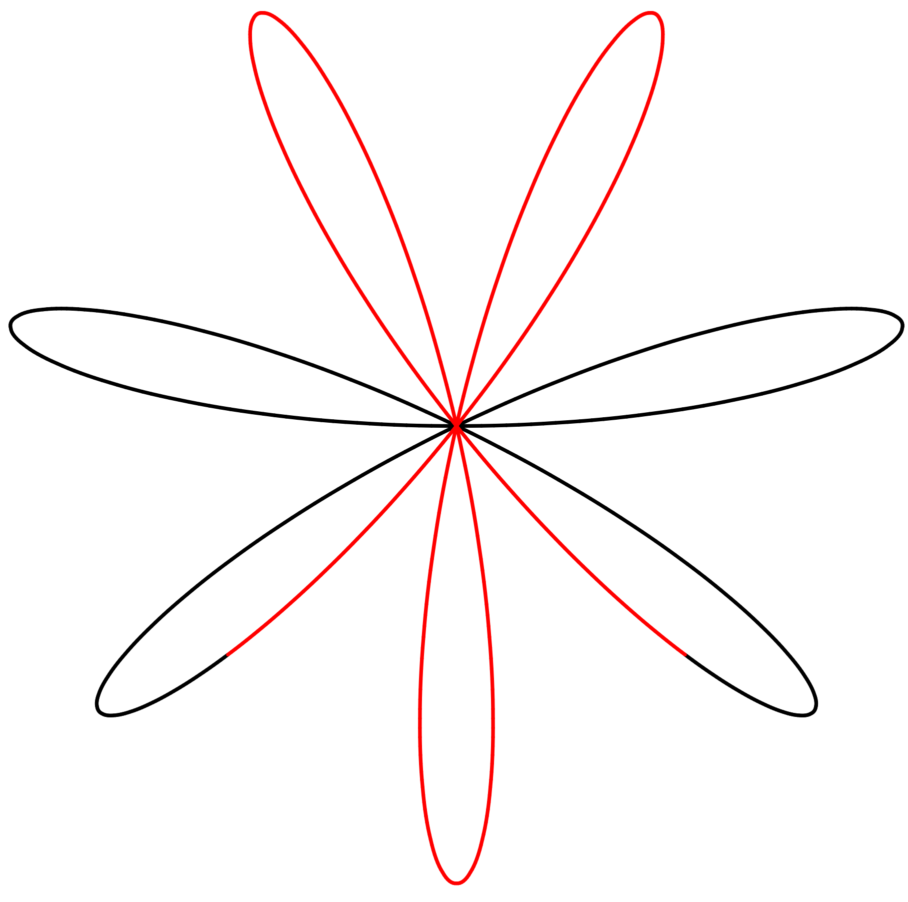


Figure 22.17: A seven-leafed rose

Observant readers may have noticed a common pattern in the equations defining the previous roses. Indeed, we leave as an exercise (nontrivial) to show that these roses are examples of the curves defined in polar coordinates by

$$\rho = \sin n\theta.$$

Example 6. Consider the following rational curve of degree 6:

$$\begin{aligned} G_1(t) &= (1-t^2)(1-14t^2+t^4), \\ G_2(t) &= 4t(1-t^2)(1+t^2), \\ G_3(t) &= (1+t^2)^3. \end{aligned}$$

We leave as exercise to show that the control polygon w.r.t. $(0, 1)$ is

```
cpoly =   {{1, 0, 1}, {1, 2/3, 1}, {0, 10/9, 6/5}, {-5/4, 5/4, 8/5},
           {-5/3, 10/9, 12/5}, {-1, 2/3, 4}, {0, 0, 8}}
```

The curve looks as follows:

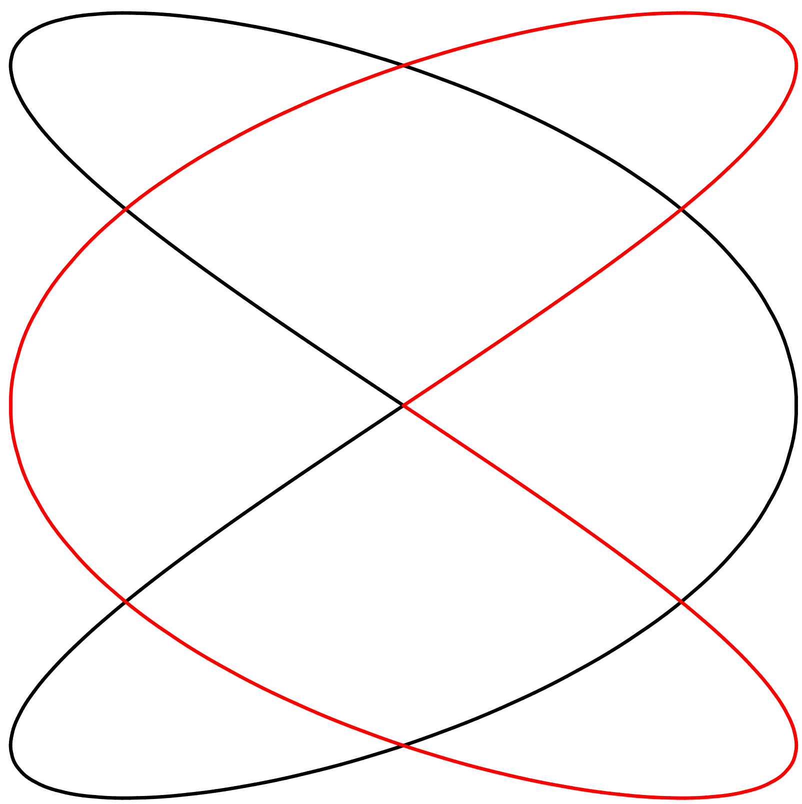


Figure 22.18: A Lissajous curve

We leave as exercise to show that the curve is also defined parametrically by

$$x(u) = \cos 3u,$$
$$y(u) = \sin 2u.$$

It is a “Lissajous” curve. Such curves occur in dynamics when considering the small oscillations of a particle with two degrees of freedom. Specifically, these curves are integral curves of the system of second-order differential equations

$$\ddot{x}_1 = -x_1,$$
$$\ddot{x}_2 = -\omega^2 x_2.$$

We now give a three-dimensional example.

Example 7. We consider the intersection of the cylinder of equation

$$X^2 + \left(Y - \frac{1}{2}\right)^2 = \frac{1}{4},$$

with the sphere of equation

$$X^2 + Y^2 + Z^2 = 1.$$

The resulting curve, called the "Viviani window", looks like a bent figure-eight in space, whose projection on the plane $Z = 0$ is a circle! It is defined by the polynomials

$$\begin{aligned}
G_1(X) &= 2X - 2X^3 \\
G_2(X) &= 4X^2 \\
G_3(X) &= 1 - X^4 \\
G_4(X) &= (1 + X^2)^2.
\end{aligned}$$

We get the polar forms

$$\begin{aligned}
g_1(X_1, X_2, X_3, X_4) &= \frac{X_1 + X_2 + X_3 + X_4}{2} - \frac{X_1X_2X_3 + X_1X_2X_4 + X_1X_3X_4 + X_2X_3X_4}{2} \\
g_2(X_1, X_2, X_3, X_4) &= \frac{2(X_1X_2 + X_1X_3 + X_1X_4 + X_2X_3 + X_2X_4 + X_3X_4)}{3} \\
g_3(X_1, X_2, X_3, X_4) &= 1 - X_1X_2X_3X_4 \\
g_4(X_1, X_2, X_3, X_4) &= 1 + \frac{X_1X_2 + X_1X_3 + X_1X_4 + X_2X_3 + X_2X_4 + X_3X_4}{3} + X_1X_2X_3X_4.
\end{aligned}$$

With respect to the affine frame $(0, 1)$ in $\mathbb{A}$, we leave as an exercise to show that the five control points $\theta_0, \theta_1, \theta_2, \theta_3, \theta_4$ are:

```
cpoly =   {{0, 0, 1, 1}, {1/2, 0, 1, 1}, {3/4, 1/2, 3/4, 4/3},
           {1/2, 1, 1/2, 2},  {0, 1, 0, 4}}
```

The curve is displayed in Figure 22.19.

Our method can also be used to render curves on surfaces. This way, we can render surfaces in terms of u-curves and v-curves. We will come back to this point when we deal with rational surfaces. A weakness of the method is that it only applies to rational curves. In particular, it does not apply to curves defined implicitly. Some methods to draw implicit curves are described in Hoffman [87]. Sometimes, it is desirable to convert parametric rational definitions into implicit form, and to apply methods to rasterize nonparametric curves. Some interesting algorithms tackling theses problems are investigated in Hobby [85, 86]. On the other hand, although restricted to rational curves, our method is very efficient.

22.6 Problems

Problem 1. Compute the control polygon for the following curve (known as a *tricuspoid*) with respect to $(0, 1)$:

$$\begin{aligned}
x &= \frac{a(3 - 6t^2 - t^4)}{1 + 2t^2 + t^4}, \\
y &= \frac{8at^3}{1 + 2t^2 + t^4},
\end{aligned}$$

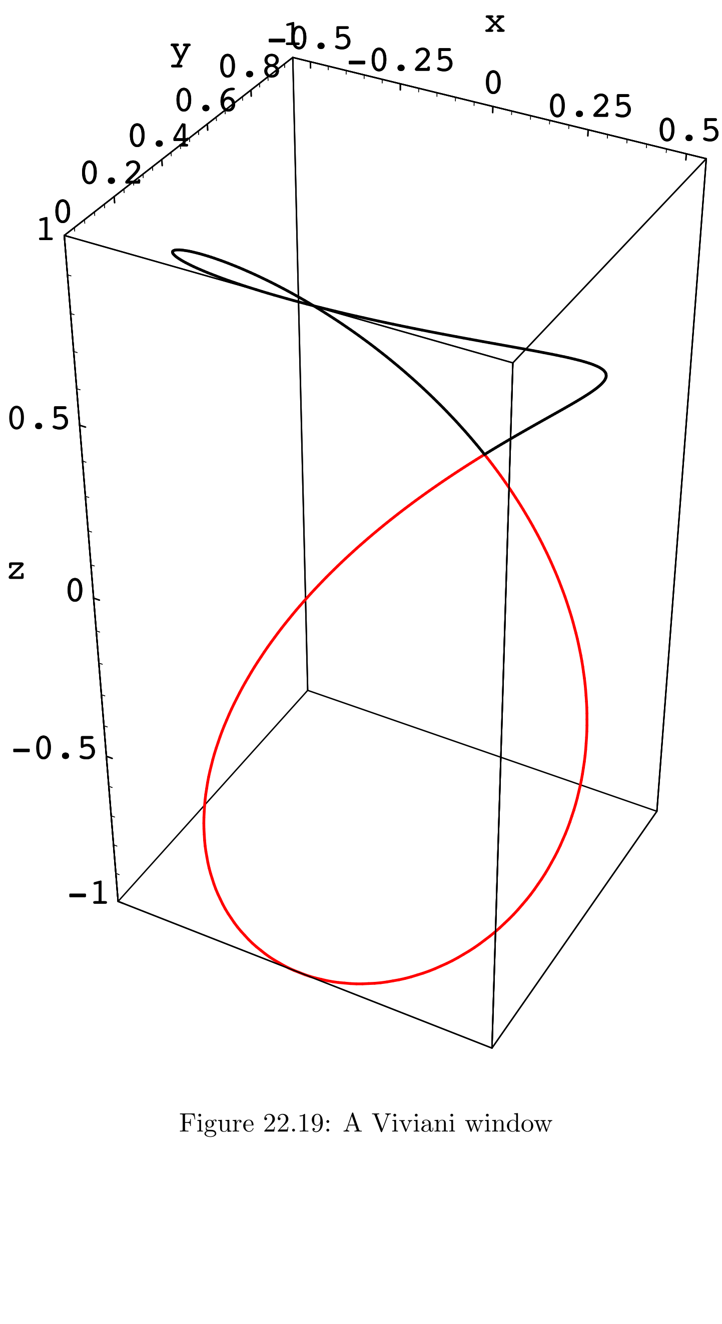


Figure 22.19: A Viviani window

Problem 2. Compute the control polygon for the following curve (known as *Freeth's nephroid*) with respect to $(0,1)$:

$$x = \frac{a(t^2+4t+1)(t^4-6t^2+1)}{t^6+3t^4+3t^2+1},$$
$$y = \frac{4at(t^2+4t+1)(1-t^2)}{t^6+3t^4+3t^2+1},$$

Problem 3. Compute the control polygon for the following curve (a type of rose) with respect to $(0,1)$:

$$x = \frac{4t(1-t^2)^2(1-14t^2+t^4)}{(1+t^2)^5},$$
$$y = \frac{8t^2(1-t^2)(3-10t^2+3t^4)}{(1+t^2)^5},$$

Problem 4. Compute the control polygon for the following curve (known as an *astroid*) with respect to $(0,1)$:

$$x = \frac{a(1-3t^2+3t^4-t^6)}{1+3t^2+3t^4+t^6},$$
$$y = \frac{8at^3}{1+3t^2+3t^4+t^6},$$

Problem 5. Compute the control polygon for the following curve (known as a *bicorne*) with respect to $(0,1)$:

$$x = \frac{a(-t^4+t^3-t+1)}{t^4-t^3+2t^2-t+1},$$
$$y = \frac{2at^2}{t^4-t^3+2t^2-t+1},$$

Problem 6. Compute the control polygon for the following curve (known as a *trisectrix of MacLaurin*) with respect to $(0,1)$:

$$x = \frac{a(3-t^2)}{(t^2+1)},$$
$$y = \frac{at(3-t^2)}{(t^2+1)},$$

Problem 7. A *limaçon of Pascal* is the curve defined as follows:

$$x = (a\cos\theta + l)\cos\theta,$$
$$y = (a\cos\theta + l)\sin\theta.$$

(i) Show that the following is a rational parameterization of the curve:

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

(ii) Compute a control net with respect to $(0, 1)$.

(iii) Study the different shapes of the curve depending on a and l, in particular when

(1) $l < a$, for example $a = 1, l = 1/2$;

(2) $a = l$ (a *cardioid*), for example $a = l = 1$;

(3) $a < l < 2a$, for example $a = 1, l = 3/2$;

(4) $2a \leq l$, for example $a = 1, l = 5/2$.

Problem 8. Write your own program for drawing a closed rational curve. Draw the curve specified by the following control polygon:

```
cpoly =  {{0, 0, 1}, {2/5, 0, 1}, {18/25, 12/25, 10/9}, {1/2, 6/5, 4/3},
   {-14/45, 79/45, 12/7}, {-45/37, 69/37, 148/63}, {-71/45, 14/9, 24/7},
   {-6/5, 11/10, 16/3}, {-12/25, 18/25, 80/9}, {0, 2/5, 16}, {0, 0, 32}}
```

Problem 9. (i) Consider a rational curve F specified by the control polygon $(\theta_0, \theta_1, \ldots, \theta_m)$ over $(0, 1)$, where each θ_i is a vector in $\widehat{\mathcal{E}}$. Show that for every constant $\rho \neq 0$, the control polygon $(\theta_0, \rho\theta_1, \ldots, \rho^m\theta_m)$ also specifies the curve F. Show that if $0 < \rho \leq 1$, the curve segment defined over $(0, 1)$ by $(\theta_0, \rho\theta_1, \ldots, \rho^m\theta_m)$ is equal to the curve segment defined over $(0, 1)$ by $(\theta_0, \theta_1, \ldots, \theta_m)$. What happens if $\rho < 0$ or $\rho > 1$?

Hint: Use the change of variable

$$t \mapsto \frac{\rho t}{1 + (\rho - 1)t}.$$

(ii) Assuming that every θ_i is a control point of the form $\langle a_i, w_i \rangle$, show that if w_0 and w_m have the same sign, then by choosing $\rho = \sqrt[m]{\frac{w_0}{w_m}}$, show that the curve F is specified by a control polygon in which both the first and the last weight are 1.

Problem 10. Consider a conic specified by three weighted control points (b_0, w_0), (b_1, w_1), and (b_2, w_2), over $(0, 1)$. Show that if the conic is not a hyperbola having an asymptote in the interval $[0, 1]$, then it can be specified by weighted control points such that w_0, w_1, w_2 all have the same sign. In this case, show that the conic is specified by control points where $w_0 = 1$, $w_1 > 0$, and $w_2 = 1$. Show that the conic is an ellipse if $w_1 < 1$, a parabola if $w_1 = 1$, and a hyperbola if $w_1 > 1$.

Problem 11. Consider a conic defined by some control points $(b_0, 1)$, (b_1, w_1), and $(b_2, 1)$, over $(0, 1)$, with $w_1 > 0$. Show that the conic is a circle iff the triangle (b_0, b_1, b_2) is isoceles (with base (b_0, b_2)). Show that $w_1 = \cos\alpha$, where α is the angle between (b_0, b_2) and (b_0, b_1). Show that the full circle can be obtained by putting together three arcs corresponding to the choice $w_1 = 1/2$.

Problem 12. Let F be a rational curve of degree m specified by $m + 1$ control points (b_i, w_i), where $w_i \neq 0$ (w.r.t. $(0, 1)$). The *weight points* q_i $(0 \leq i \leq m - 1)$ are the points defined such that

$$q_i = \frac{w_i\, b_i + w_{i+1}\, b_{i+1}}{w_i + w_{i+1}}.$$

Prove that if $w_i > 0$ $(0 \leq i \leq m)$, then the curve segment F over $(0, 1)$ is contained in the convex hull of the polygon

$$(b_0, q_0, \ldots, q_{m-1}, b_m).$$

Problem 13. Let F be a rational curve of degree m specified by $m + 1$ control points (b_i, w_i) w.r.t. $(0, 1)$. Show that the $m + 2$ control points (b_i^1, w_i^1) of the curve obtained by raising the degree of F from m

to $m+1$ are given by the formulae:

$$\begin{aligned} b_i^1 &= \frac{iw_{i-1}\, b_{i-1} + (m+1-i)w_i\, b_i}{w_i^1}, \\ w_i^1 &= iw_{i-1} + (n+1-i)w_i, \end{aligned}$$

where $1 \leq i \leq m$, $(b_0^1, w_0^1) = (b_0, w_0)$, and $(b_{m+1}^1, w_{m+1}^1) = (b_m, w_m)$.

Problem 14. Show that the same rational curve can be specified by different control polygons (use degree raising and reparameterization).

Problem 15. (i) Given a polynomial cubic F specified by the control points (b_0, b_1, b_2, b_3) over $(0, 1)$, define the points l_1 and l_2 as follows:

$$\begin{aligned} l_1 &= \frac{3}{2}\, b_1 - \frac{1}{2}\, b_0, \\ l_2 &= \frac{3}{2}\, b_2 - \frac{1}{2}\, b_3. \end{aligned}$$

Prove that $F(t)$ is given by the formula

$$F(t) = (1-t)^2\, b_0 + 2(1-t)^2 t\, l_1 + 2t^2(1-t)\, l_2 + t^2\, b_3.$$

Show that if $l_1 = l_2$, then F is a curve of degree 2.

Remark: The points l_1 and l_2 are called the *Ball points*.

(ii) Assuming that $l_1 \neq l_2$, pick any point c on the line (l_1, l_2), let H be any plane in $\mathbb{A}^3$ such that $c \notin H$, and consider the image $F_{c,H}$ of F onto H under the central projection of center c. Show that $F_{c,H}$ is a curve of degree 2. Conclude that every (true) cubic lies on a cone of degree 2. Show that $F(\infty) = l_1 - l_2$ (a point at infinity in $\widetilde{\mathcal{E}}$). Thus, $c = l_1 - l_2$ is on the cubic (at infinity), and in this case, the cone is a parabolic cylinder.

(iii) Generalize (i) and (ii) to rational cubics.

Problem 16. (i) Given a polynomial cubic F specified by the control points (b_0, b_1, b_2, b_3) over $(0, 1)$, define the points a_1 and a_2 as follows:

$$\begin{aligned} a_1 &= \frac{3}{2}\, b_1 - \frac{1}{2}\, b_3, \\ a_2 &= \frac{3}{2}\, b_2 - \frac{1}{2}\, b_0. \end{aligned}$$

Consider the rectangular bilinear patch $X(u, v)$ defined such that

$$X(u, v) = (1-u)(1-v)\, b_0 + (1-u)v\, a_1 + uv\, a_2 + u(1-v)\, b_3.$$

Prove that the cubic F lies on the surface X.

Hint: Let $v = 2(1-u)u$ in $X(u, v)$.

(ii) Generalize (i) to rational cubics.

Conclude that every rational cubic lies on the intersection of a quadratic cone with a rational bilinear patch. Is the converse true?

Problem 17. Let F be a rational curve of degree m defined by the control points $((b_0, w_0), \ldots, (b_m, w_m))$ over $(0, 1)$. Prove that the curvature at b_0 is

$$\kappa(0) = \frac{(m-1)}{m} \frac{w_0 w_2}{w_1^2} \frac{\|\mathbf{b_0 b_1} \times \mathbf{b_1 b_2}\|}{\|\mathbf{b_0 b_1}\|^3}.$$

Show that the torsion at b_0 is given by

$$\tau(0) = -\frac{(m-2)}{m} \frac{w_0 w_3}{w_1 w_2} \frac{(\mathbf{b_0 b_1},\, \mathbf{b_0 b_2},\, \mathbf{b_0 b_3})}{\|\mathbf{b_0 b_1} \times \mathbf{b_1 b_2}\|^2}.$$

Use the above to solve the following problem: Given a sequence of four control points (b_0, b_1, b_2, b_3) and two numbers $\kappa_0 > 0$ and $\kappa_3 > 0$, find w_1, w_2 such that the rational cubic specified by $((b_0, 1), (b_1, w_1), (b_2, w_2), (b_3, 1))$ has curvature κ_0 at b_0 and curvature κ_3 at b_3.

Problem 18. Let F be a rational curve of degree m defined by the control points $((b_0, w_0), \ldots, (b_m, w_m))$ over $(0, 1)$. Prove that the derivative at b_0 is given by

$$F'(0) = \frac{m\, w_1}{w_0}\,(b_1 - b_0).$$

Problem 19. Let F be a rational curve of degree m defined by the multilinear map $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$. Consider the change of variable

$$(t, z) \mapsto (at + bz,\, ct + dz),$$

where $ad - bc \neq 0$. Let g be the multilinear map defined such that

$$g((t_1, z_1), \ldots, (t_m, z_m)) = f((at_1 + bz_1, ct_1 + dz_1), \ldots, (at_m + bz_m, ct_m + dz_m)).$$

Show that the control point $g(\underbrace{(r, 1), \ldots, (r, 1)}_{m-i}, \underbrace{(s, 1), \ldots, (s, 1)}_{i})$ w.r.t. (r, s) are given by the expressions

$$(cr + d)^{m-i}(cs + d)^i f\Bigg(\underbrace{\left(\frac{ar + b}{cr + d}, 1\right), \ldots, \left(\frac{ar + b}{cr + d}, 1\right)}_{m-i}, \underbrace{\left(\frac{as + b}{cs + d}, 1\right), \ldots, \left(\frac{as + b}{cs + d}, 1\right)}_{i}\Bigg),$$

provided that $cr + d \neq 0$ and $cs + d \neq 0$.

Problem 20. Let F be a rational curve of degree m defined by the multilinear map $f\colon (\widehat{\mathbb{A}})^m \to \widehat{\mathcal{E}}$.

(i) Letting

$$B_i^m(u, v) = \binom{m}{i} u^i v^{m-i},$$

and assuming that the control points induced by f over $(0, 1)$ are of the form $\langle b_i, w_i \rangle$ where $w_i \neq 0$, prove that if $\sum_{i=0}^m w_i B_i^m(t, z) \neq 0$, then

$$F([t, z]) = \frac{\displaystyle\sum_{i=0}^m w_i B_i^m(t, z)\, b_i}{\displaystyle\sum_{i=0}^m w_i B_i^m(t, z)},$$

for all homogeneous coordinates $[t, z] \in \mathbb{R}^2$. Show that $F(t, z)$ can be computed by a simple modification of the de Casteljau algorithm (instead of computing the affine combination $(1 - u)\, b_i^r + u\, b_{i+1}^r$, compute the linear combination $z\, b_i^r + u\, b_{i+1}^r$).

(ii) Consider the C^0-continuous curve in $\mathbb{R}^2$ defined such that $t \mapsto (t,\, 1 - |t|)$ over $[-1, 1]$. This curve is a model of the projective line in $\mathbb{R}^2$ (identifying the points $(-1, 0)$ and $(1, 0)$). Explain how to draw a closed rational curve in a single piece using the results of (i) and (ii) (as the trace of $F([u, 1 - |u|])$). Apply this method to draw an ellipse. Experiment with other $2D$ models of the projective line.

Remark: This problem is an adaptation of results due to DeRose [45].

Problem 21. Let C be the curve $C\colon \mathbb{RP}^1 \to \mathbb{RP}^3$ (called the *twisted cubic*) defined such that

$$C(u, v) = (u^3,\, u^2 v,\, uv^2,\, v^3),$$

for all homogeneous coordinates $(u, v) \in \mathbb{R}^2$.

(i) Prove that C is the intersection of the three quadrics Q_1, Q_2, Q_3 of equations

$$\begin{aligned} &Q_1\colon xz - y^2 = 0,\\ &Q_2\colon xt - yz = 0,\\ &Q_3\colon yt - z^2 = 0, \end{aligned}$$

where (x, y, z, t) are homogeneous coordinates in $\mathbb{R}^4$.

Hint: Show that if $p = (x, y, z, t) \in Q_1 \cap Q_2 \cap Q_3$, then either $x \neq 0$ or $t \neq 0$. In the first case, p has homogeneous coordinates of the form $(x^3,\, x^2 y,\, xy^2,\, y^3)$, and in the second case, p has homogeneous coordinates of the form $(z^3,\, z^2 t,\, zt^2,\, t^3)$.

(ii) Prove that the intersection of any two of Q_i and Q_j (where $i \neq j$) is equal to the union of C with some line. Conclude that C is not equal to the intersection of only two of the Q_i's.

(iii) Consider the affine trace of C obtained by setting $u = 1$. Show that in the affine patch in $\mathbb{RP}^3$ corresponding to $x = 1$, this is the curve

$$t \mapsto (t,\, t^2,\, t^3).$$

Show that this curve is exactly the intersection of the affine traces of Q_1 and Q_2 (corresponding to $x = 1$). What are these two surfaces? How do you resolve the apparent conflict with the result of question (ii)?