

A Simple Method For Drawing a Rational Curve as Two Bézier Segments

Jean Gallier

Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104, USA
`jean@saoul.cis.upenn.edu`

April 7, 2024

Abstract. In this paper, we give a simple method for drawing a closed rational curve specified in terms of control points as two Bézier segments. The main result is the following:

For every affine frame (r, s) (where $r < s$), for every rational curve $F(t)$ specified over $[r, s]$ by some control polygon $(\beta_0, \dots, \beta_m)$ (where the β_i are weighted control points or control vectors), the control points $(\theta_0, \dots, \theta_m)$ (w.r.t. $[r, s]$) of the rational curve $G(t) = F(\varphi(t))$ are given by

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

where $\varphi: \mathbb{RP}(1) \rightarrow \mathbb{RP}(1)$ is the projectivity mapping $[r, s]$ onto $\mathbb{RP}(1) -]r, s[$. Thus, in order to draw the entire trace of the curve F over $[-\infty, +\infty]$, we simply draw the curve segments $F([r, s])$ and $G([r, s])$.

The correctness of the method is established using a simple geometric argument about ways of partitioning the real projective line into two disjoint segments. Other known methods for drawing rational curves can be justified using similar geometric arguments.

1 Introduction

In this paper, we give a simple method for drawing a closed rational curve specified in terms of control points as two Bézier segments. In order to explain the basic intuition behind our method as clearly as possible, we first consider rational curves specified in terms of rational functions. For example, consider the following rational curve of degree 8 (a rose, see Section 4) 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.

Nevertheless, one might rightly ask why this is not a trivial problem! There are at least two reasons. Firstly, it is not obvious how to do draw a closed rational curve very efficiently, and secondly, in most CAGD applications, the curve is given in terms of control points rather than parametrically. When 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 the curve G defined by the fractions

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

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 compute the control points of G . This is obviously a lot of work, and it may be computationally rather expensive.

The main contribution of this paper is that there is a very simple (and cheap) method for getting the control points of G from the control points of F . Indeed, if $(\beta_0, \dots, \beta_m)$ are the control points (really, weighted control points or control vectors, see section 2) of the curve F w.r.t. the affine frame $(-1, 1)$, the control points $(\theta_0, \dots, \theta_m)$ of the 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. An implementation of our method in *Mathematica* using a subdivision version of the de Casteljau algorithm can be found in Gallier [9].

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 [14] and Patterson [15]. 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 [1], and by DeRose [4], who credits Patterson [15] for the original idea behind the method. We will compare these methods with ours in section 3.

For the sake of brevity, we do not review how polynomial curves are defined in terms of control points. The deep reason why polynomial curves can be effectively handled in terms of control points is that polynomials arise from multiaffine symmetric maps (see de Casteljau [3], Ramshaw [16], Farin [6, 5], Hoschek and Lasser [13], or Gallier [9]). However, we briefly review how to handle rational curves in terms of control points.

2 Rational Curves and Control Points

A rational curve $F: \mathbb{R} \rightarrow \mathbb{R}^n$ of degree m is specified by n fractions

$$\begin{aligned} x_1(t) &= \frac{F_1(t)}{F_{n+1}(t)}, \\ &\dots = \dots \\ x_n(t) &= \frac{F_n(t)}{F_{n+1}(t)}, \end{aligned}$$

where $F_1(t), \dots, F_{n+1}(t)$ 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 $F: \mathbb{R} \rightarrow \mathbb{R}^n$ as the projection of the polynomial curve $G: \mathbb{R} \rightarrow \mathbb{R}^{n+1}$ defined by the polynomials $F_1(X), \dots, F_{n+1}(X)$ onto the affine hyperplane of equation $x_{n+1} = 1$. Furthermore, in order to handle rigorously the situation where $t = \pm\infty$, we can view the curve $G: \mathbb{R} \rightarrow \mathbb{R}^{n+1}$ as a rational map $\tilde{F}: \mathbb{RP}(1) \rightarrow \mathbb{RP}(n)$, where $\mathbb{RP}(n)$ is the real projective space associated with the vector space \mathbb{R}^{n+1} (and thus, $\mathbb{RP}(1)$ is the real projective line). For this, we homogenize the polynomials $F_1(X), \dots, F_{n+1}(X)$ as polynomials of the same total degree m (replacing X by X/Z). Now, the projective space $\mathbb{RP}(n)$ has a well known model (for instance, see Berger [2], Samuel [17], or Stolfi [18]) in which points at infinity correspond to the lines through the origin contained in the hyperplane of equation $x_{n+1} = 0$, and other points correspond to points in the affine hyperplane of equation $x_{n+1} = 1$ (since these points correspond to lines through the origin and not contained in the hyperplane of equation $x_{n+1} = 0$, and such lines intersect the affine hyperplane of equation $x_{n+1} = 1$ in a single point). Thus, the original rational curve $F: \mathbb{R} \rightarrow \mathbb{R}^n$ is the projection of the curve $\tilde{F}: \mathbb{RP}(1) \rightarrow \mathbb{RP}(n)$ onto the affine hyperplane $x_{n+1} = 1$, under the projection map $\Pi: \mathbb{RP}(n) \rightarrow \mathbb{R}^n$ defined such that

$$\Pi(x_1, \dots, x_n, x_{n+1}) = \left(\frac{x_1}{x_{n+1}}, \dots, \frac{x_n}{x_{n+1}} \right),$$

provided that $x_{n+1} \neq 0$. If $x_{n+1} = 0$ and $x_i \neq 0$ for some i , $1 \leq i \leq n$, $\Pi(x_1, \dots, x_n, x_{n+1})$ is a point at infinity, while $\Pi(0, \dots, 0)$ is undefined.

For example, a circle C can be defined by the following rational functions:

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

After homogenizing, we get the following curve $\tilde{C}: \mathbb{RP}(1) \rightarrow \mathbb{RP}(2)$:

$$\begin{aligned}x(X, Z) &= Z^2 - X^2 \\y(X, Z) &= 2XZ \\w(X, Z) &= Z^2 + X^2.\end{aligned}$$

Now, it is well known that the control points of a polynomial curve $F: \mathbb{R} \rightarrow \mathbb{R}^n$ are obtained from the *polar form (also called blossom)* of the curve w.r.t. to some affine frame (see de Casteljau [3], Ramshaw [16], Farin [6, 5], Hoschek and Lasser [13], or Gallier [9]). The only small difference between the classical situation and the present situation is that points are represented by homogeneous coordinates, and that the polar form of homogeneous polynomials are multilinear maps (as opposed to the classical case in which polar forms are only multiaffine). Thus, given some affine frame $((r_1, r_2), (s_1, s_2))$, if $f: (\mathbb{R}^2)^m \rightarrow \mathbb{R}^{n+1}$ is the polar form of the curve $\tilde{F}: \mathbb{RP}(1) \rightarrow \mathbb{RP}(n)$, the $m + 1$ control points $(\theta_0, \dots, \theta_m)$ specifying \tilde{F} are given by the formulae

$$\theta_i = f(\underbrace{(r_1, r_2), \dots, (r_1, r_2)}_{m-i}, \underbrace{(s_1, s_2), \dots, (s_1, s_2)}_i).$$

Going back to the example of the circle, we get the following polar forms:

$$\begin{aligned}f_1((X_1, Z_1), (X_2, Z_2)) &= Z_1 Z_2 - X_1 X_2, \\f_2((X_1, Z_1), (X_2, Z_2)) &= X_1 Z_2 + X_2 Z_1, \\f_3((X_1, Z_1), (X_2, Z_2)) &= Z_1 Z_2 + X_1 X_2.\end{aligned}$$

If $(r_1, r_2) = (0, 1)$ and $(s_1, s_2) = (1, 1)$, we get the following control points (in \mathbb{R}^3):

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

The third control point θ_3 corresponds to the weighted point $((0, 1), 2)$, where $(0, 1)$ is a point in \mathbb{R}^2 .

Remark: To be perfectly rigorous, the curve $\tilde{F}: \mathbb{RP}(1) \rightarrow \mathbb{RP}(n)$ is defined by the multi-projective map $\mathbf{P}(f): (\mathbb{RP}(1))^m \rightarrow \mathbb{RP}(n)$ induced by the multilinear map $f: (\mathbb{R}^2)^m \rightarrow \mathbb{R}^{n+1}$. Furthermore, the control points θ_i specified in homogeneous coordinates are really control points for the polynomial curve $G: \mathbb{R} \rightarrow \mathbb{R}^{n+1}$. Thus, the θ_i really live in a certain vector space $\widehat{\mathbb{R}^n}$ (isomorphic to \mathbb{R}^{n+1}). They are weighted points when $x_{n+1} \neq 0$, and control vectors when $x_{n+1} = 0$. There are some subtleties that we are glossing over in this paper. A rigorous account of this approach is given in Fiorot and Jeannin [7, 8] and Gallier [9]. However, going into such matters would distract us from the main goal of this paper.

For the sake of notational simplicity, from now on, we will identify F and \tilde{F} , thus writing $F: \mathbb{RP}(1) \rightarrow \mathbb{RP}(n)$ instead of $\tilde{F}: \mathbb{RP}(1) \rightarrow \mathbb{RP}(n)$.

3 Proof of the Main Result

Since a rational curve $F: \mathbb{RP}(1) \rightarrow \mathbb{RP}(n)$ is a map with domain the real projective line $\mathbb{RP}(1)$, to draw the entire trace of the curve F over $\mathbb{RP}(1)$, we can partition $\mathbb{RP}(1)$ into closed intervals I_0, \dots, I_p (intersecting only at boundary points) and find some simple projectivities $\varphi_1, \dots, \varphi_p$ 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 in Section 1, we described a way of partitioning $\mathbb{RP}(1)$ where $I_0 = [-1, 1]$, $I_1 = \mathbb{RP}(1) -]-1, 1[$, and the projectivity $\varphi_1: \mathbb{RP}(1) \rightarrow \mathbb{RP}(1)$

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

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

Besides the *hyperplane model* of the projective space $\mathbb{RP}(n)$ reviewed in Section 2, there is another well known model in terms of the n -sphere, the *sphere model*. (see Berger [2], Samuel [17], or Stolfi [18]). Indeed, $\mathbb{RP}(n)$ is the quotient space obtained by identifying antipodal points on S^n (recall that $S^n = \{(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1^2 + \dots + x_{n+1}^2 = 1\}$, and the antipodal point of (x_1, \dots, x_{n+1}) is $(-x_1, \dots, -x_{n+1})$). In particular, 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 indispensable, 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 1.

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

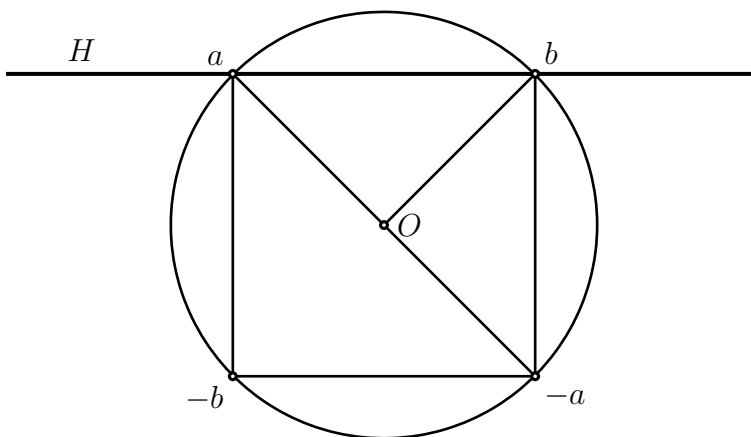


Figure 1: Model of the projective line, Case 1: $I_0 = [-1, 1]$

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 $\varphi: \mathbb{RP}(1) \rightarrow \mathbb{RP}(1)$, 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 2 illustrates the case where a is inside the circle.

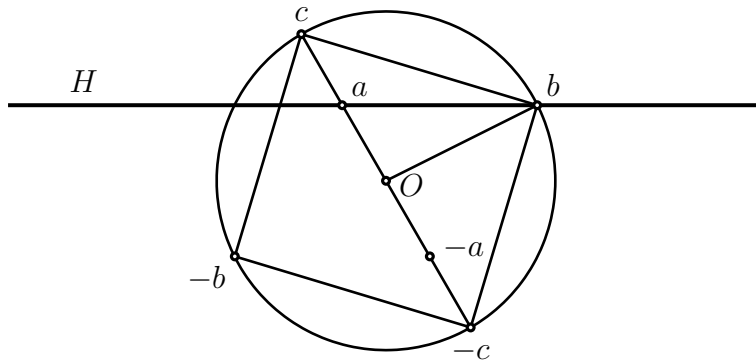


Figure 2: Model of the projective line, Case 2: $I_0 = [r, s]$

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 3.1 For every affine frame (r, s) ($r < s$) and every rational curve $F: \mathbb{RP}(1) \rightarrow \mathbb{RP}(n)$ of degree m specified by some symmetric multilinear map $f: (\mathbb{R}^2)^m \rightarrow \mathbb{R}^{n+1}$, the rational curve G is specified by the symmetric multilinear map $g: (\mathbb{R}^2)^m \rightarrow \mathbb{R}^{n+1}$ such that

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

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

$$f(\underbrace{(t, z), \dots, (t, z)}_m) = \vec{0}$$

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 $F_1(t), \dots, F_{n+1}(t)$ 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 $\vec{0} = \underbrace{(0, \dots, 0)}_{n+1}$

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 3.2 For every affine frame (r, s) ($r < s$) and every rational curve $F: \mathbb{RP}(1) \rightarrow \mathbb{RP}(n)$ of degree m specified by some symmetric multilinear map $f: (\mathbb{R}^2)^m \rightarrow \mathbb{R}^{n+1}$, if $g: (\mathbb{R}^2)^m \rightarrow \mathbb{R}^{n+1}$ is the symmetric multilinear map of definition 3.1, 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 $(\beta_0, \dots, \beta_m)$ are the control points (in \mathbb{R}^{n+1}) of F w.r.t. the affine frame (r, s) , the control points $(\theta_0, \dots, \theta_m)$ (in \mathbb{R}^{n+1}) of the curve G w.r.t. (r, s) are given by the equations

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

Proof. We have

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

and thus

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

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

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

and since

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

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

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

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 3.2 to this more computational proof, which, in our opinion, obscures what's really going on!

Lemma 3.2 shows that F and G have the same trace. It also shows that if the control points (in \mathbb{R}^{n+1}) of F w.r.t. (r, s) are $(\beta_0, \dots, \beta_m)$, if β_i is of the form (a, w) , where a is a point in \mathbb{R}^n and $w \neq 0$ is a weight, then

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

and if β_i is a control vector $\vec{u} \in \mathbb{R}^n$, then

$$\theta_i = (-1)^i \vec{u}.$$

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

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

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) [1]. 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 3.

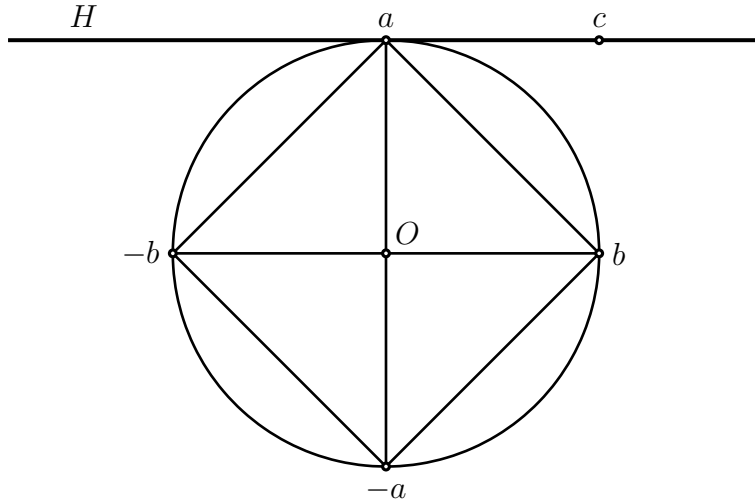


Figure 3: Model of the projective line, Case 3

However, Bajaj and Royappa 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 due to DeRose [4] can be explained as follows. If F is a rational curve of degree m defined by the multilinear map $f: (\mathbb{R}^2)^m \rightarrow \mathbb{R}^{n+1}$, 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 $w_i \neq 0$, it is easily shown that if $\sum_{i=0}^m w_i B_i^m(t, z) \neq 0$, then

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

for all homogeneous coordinates $[t, z] \in \mathbb{R}^2$. Then, $F(t, z)$ 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 $z b_i^r + u b_{i+1}^r$. Finally, 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 $(1, 0)$). 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. Further computer experimentation seems needed to compare the two methods.

Our method has been implemented in *Mathematica*, and the code can be found in Gallier [9]). It can also be used to render curves on surfaces.

4 Examples

Our first example is a 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}}
```

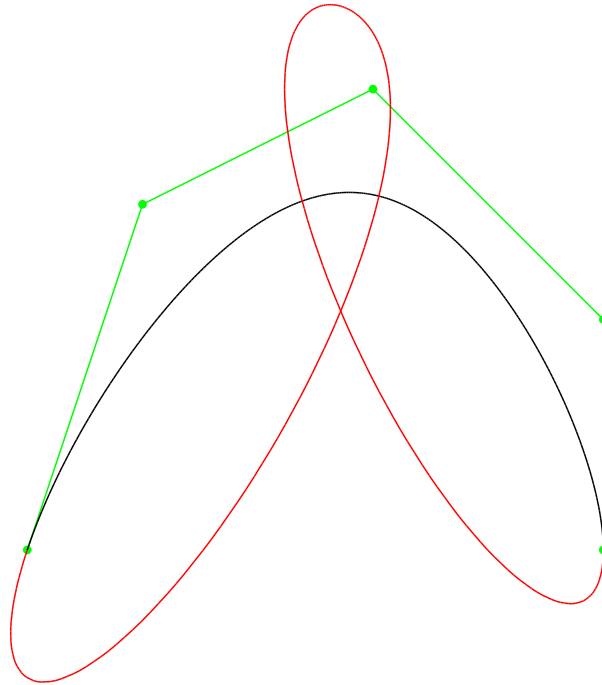


Figure 4: A Closed Quartic

Our second example is the following rational curve of degree 8:

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

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

The control polygon w.r.t. $(-1, 1)$ is

```
inpol2 = {{1/2, -1/2, 16}, {8, -6, 0}, {-7/4, 7/2, 16/7},
          {-8, 2, 0}, {0, -35/6, 48/35}, {8, 2, 0}, {7/4, 7/2, 16/7},
          {-8, -6, 0}, {-1/2, -1/2, 16}}
```

There are four control vectors. The curve looks like a seven-leafed rose.

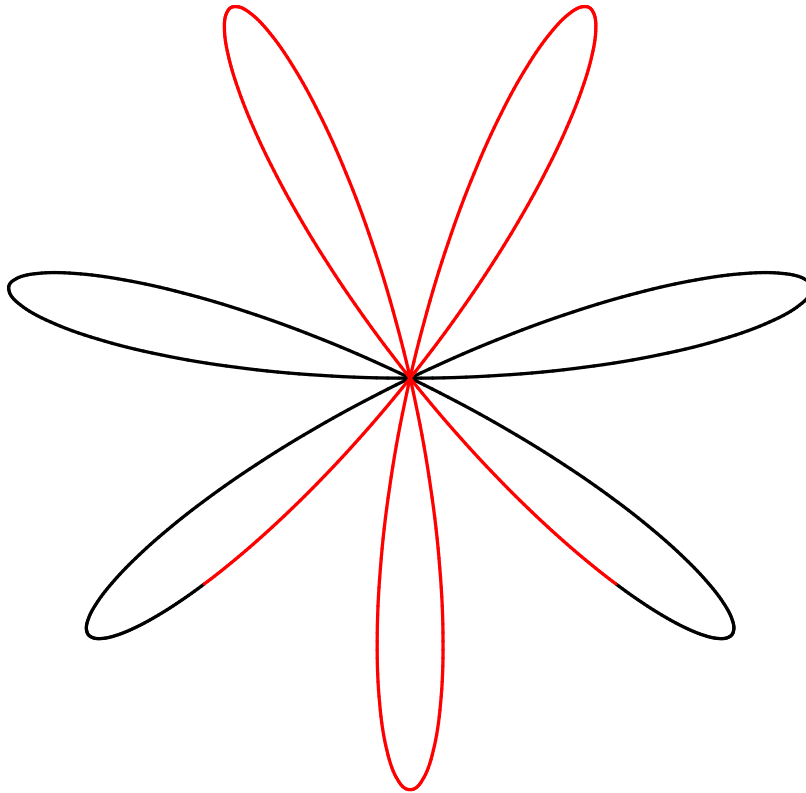


Figure 5: A seven-leafed rose, 6 iterations

Experimenting with our method, we observed that it is advantageous to compute a control polygon w.r.t. $(-1, 1)$. This is because the complementary curve segment specified by $F(1/t)$ is subdivided more evenly than the curve segment specified by $F(t/(2t - 1))$ over $(0, 1)$. In computing the new control polygon w.r.t. $(-1, 1)$, control vectors occasionally arise. Since we apply the de Casteljau algorithm in \mathbb{R}^{n+1} and apply the projection Π at the end (i.e., divide by the “weight factor” x_{n+1} at the end), there is absolutely no problem with control vectors. More examples, some rather pretty, can be found in Gallier [9].

5 Conclusion

We have presented an efficient method for drawing closed rational curves, defined parametrically or in terms of weighted control points. The method relies on a subdivision version of the standard de Casteljau algorithm. One of the advantages of our method is that it is incremental. Indeed, the algorithm produces an approximation of the curve as a sequence of control polygons. Thus, if we wish to get better accuracy, we can subdivide each control polygon in the list. We can also achieve a zooming effect by selectively subdividing some

subsequences of control polygons. We can also use our method for drawing rational curves on rational surfaces. This way, we can render surfaces in terms of u -curves and v -curves. 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 [12]. Sometimes, it is desirable to convert parametric rational definitions into implicit form, and to apply methods to rasterize nonparametric curves. Some interesting algorithms tackling these problems are investigated in Hobby [10, 11]. On the other hand, although restricted to rational curves, our method is very efficient.

References

- [1] C. Bajaj and A. Royappa. Finite representation of real parametric curves and surfaces. *Intl. J. of Computational Geometry and Applications*, pages 313–326, 1995.
- [2] Marcel Berger. *Géométrie 1*. Nathan, 1990. English edition: *Geometry 1*, Universitext, Springer Verlag.
- [3] Paul de Faget de Casteljaou. *Shape Mathematics and CAD*. Hermes, first edition, 1986.
- [4] Tony D. DeRose. Rational Bézier curves and surfaces on projective domains. In G. Farin, editor, *NURBS for Curve and Surface Design*, pages 35–45. SIAM, Philadelphia, Pa, 1991.
- [5] Gerald Farin. *NURB Curves and Surfaces, from Projective Geometry to practical use*. AK Peters, first edition, 1995.
- [6] Gerald Farin. *Curves and Surfaces for CAGD*. Academic Press, fourth edition, 1998.
- [7] J.-C. Fiorot and P. Jeannin. *Courbes et Surfaces Rationnelles*. RMA 12. Masson, first edition, 1989.
- [8] J.-C. Fiorot and P. Jeannin. *Courbes Splines Rationnelles*. RMA 24. Masson, first edition, 1992.
- [9] Jean H. Gallier. *Curves and Surfaces In Geometric Modeling: Theory And Algorithms*. Morgan Kaufmann, 1999.
- [10] John D. Hobby. Rasterization of nonparametric curves. *ACM Transactions on Graphics*, 9(3):262–277, 1990.
- [11] John D. Hobby. Numerically stable implicitization of cubic curves. *ACM Transactions on Graphics*, 10(3):255–296, 1991.
- [12] Christoph M. Hoffman. *Geometric and Solid Modeling*. Morgan Kaufmann, first edition, 1989.

- [13] J. Hoschek and D. Lasser. *Computer Aided Geometric Design*. AK Peters, first edition, 1993.
- [14] E. Lee. The rational Bézier representation for conics. In G. Farin, editor, *Geometric Modeling: Algorithms and New Trends*, pages 3–19. SIAM, Philadelphia, Pa, 1987.
- [15] R. Patterson. Projective transformations of the parameter of a rational Bernstein-Bézier curve. *ACM Transactions on Graphics*, 4:276–290, 1986.
- [16] Lyle Ramshaw. Blossoming: A connect-the-dots approach to splines. Technical report, Digital SRC, Palo Alto, CA 94301, 1987. Report No. 19.
- [17] Pierre Samuel. *Projective Geometry*. Undergraduate Texts in Mathematics. Springer Verlag, first edition, 1988.
- [18] J. Stolfi. *Oriented Projective Geometry*. Academic Press, first edition, 1991.