

Original Lecture #10,11: 3,5 November 1992
Topics: Rational Curves; Tensor Product Surfaces
Scribe: Dave Wilkins and Nina Billawala

1 Knot insertion

Suppose that you want more control over a section of a curve; one possibility is to raise the degree. Unfortunately, you must raise the degree of the entire spline. If it is cubic, you must degree-raise all of the segments to be quartic. Note also that the knot multiplicities will change. A C^2 joint in a cubic spline corresponds to a single knot. If we degree-raise that spline to a quartic, however, the joint is still C^2 and hence corresponds to a double knot.

Instead of degree-raising, it is often easier to add a new knot, which adds a new de Boor point and produces a new control polygon. For example, if we start with the knot sequence

$$(\dots, 0, 2, 4, 6, 8, 10, \dots), \quad (1)$$

we could add a new knot at point 5 giving the knot sequence

$$(\dots, 0, 2, 4, 5, 6, 8, 10, \dots). \quad (2)$$

Figure 1 shows the spline prior to the knot insertion and figure 2 shows the spline with the additional knot and new control polygon. This additional control point gives the user one more degree of freedom to perturb the spline.

2 Rational Curves

Consider the rectangular hyperbola $XY = 1$. It can be parametrized by

$$\begin{aligned} X(T) &= T \\ Y(T) &= \frac{1}{T} \end{aligned} \quad (3)$$

Consider a hyperbola parameterized by the above equations. What is the degree? The naive way is to find the highest degree of any polynomial involved. But that leads to the conclusion that the hyperbola has degree 1, which is clearly not the case. A single line intersects the curve twice, which implies that the curve must be at least quadratic. Furthermore, it goes to infinity with two different slopes.

Rule 1. *Put all coordinates of any rational curve or surface over a common denominator.*

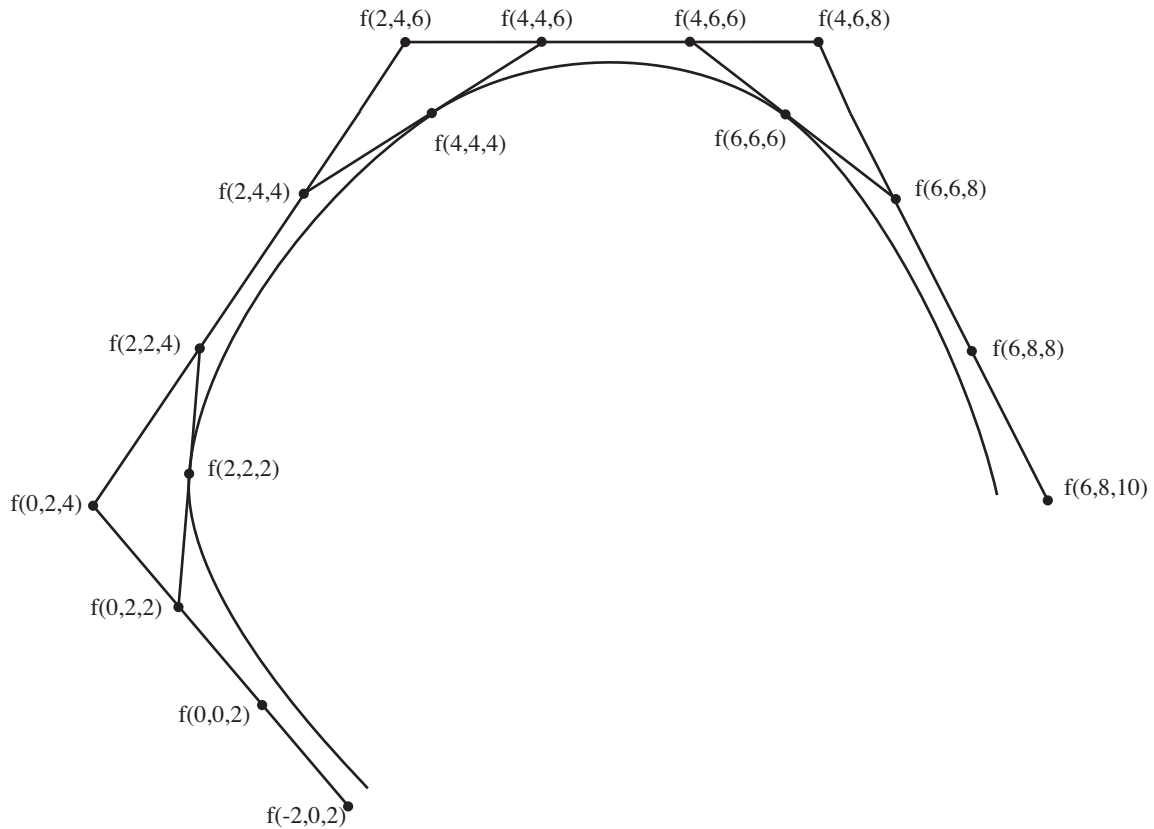


Figure 1: Before knot insertion

In our hyperbola example, we rewrite (3) as

$$\begin{aligned} X(T) &= \frac{T^2}{T} \\ Y(T) &= \frac{1}{T} \end{aligned}$$

Rewriting the equation shows that the curve is quadratic. In general, we have

$$X(T) = \frac{x(T)}{w(T)}, \quad Y(T) = \frac{y(T)}{w(T)}$$

where the two denominators are the same. For example:

$$X(T) = \frac{3T^3+7T^2+2T+5}{7T^3-2T+4}, \quad Y(T) = \frac{6T^3+4T^2+6}{7T^3-2T+4}$$

The denominator acts as a weight coordinate. We can save writing the common denominator twice if we express the rational curve in terms of homogeneous coordinates: $[w(T); x(T), y(T)]$ where $x(T)$, $y(T)$, and $w(T)$ are polynomials. Rational curves, therefore, are best thought of as curves whose homogeneous coordinates are given by polynomials.

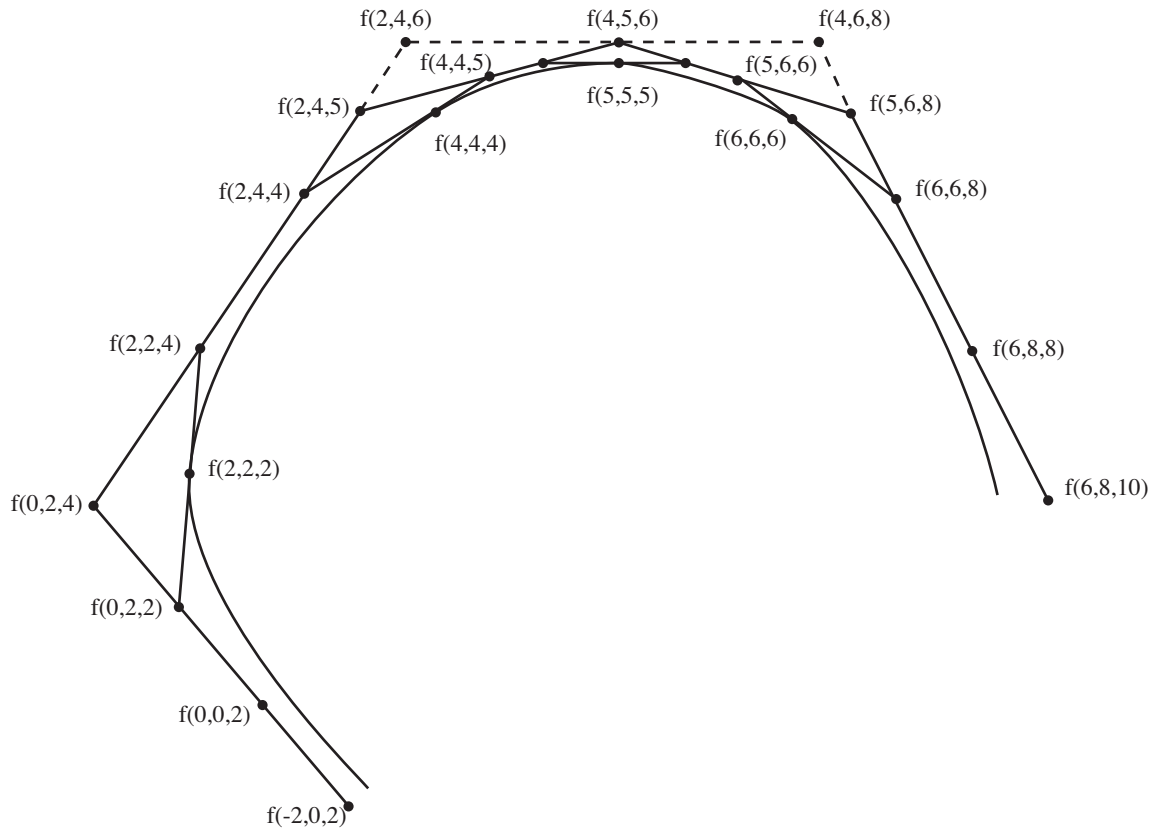


Figure 2: After knot insertion

$$X(T) = \frac{x(T)}{w(T)} = \frac{T^2}{T}, \quad Y(T) = \frac{y(T)}{w(T)} = \frac{1}{T}$$

What happens when $T \rightarrow 0$ or $T \rightarrow \infty$?

$$T \rightarrow 0, \quad \frac{T^2}{T} \rightarrow 0, \quad \frac{1}{T} \rightarrow \infty, \quad (T; T^2, 1) \rightarrow (0; 0, 1)$$

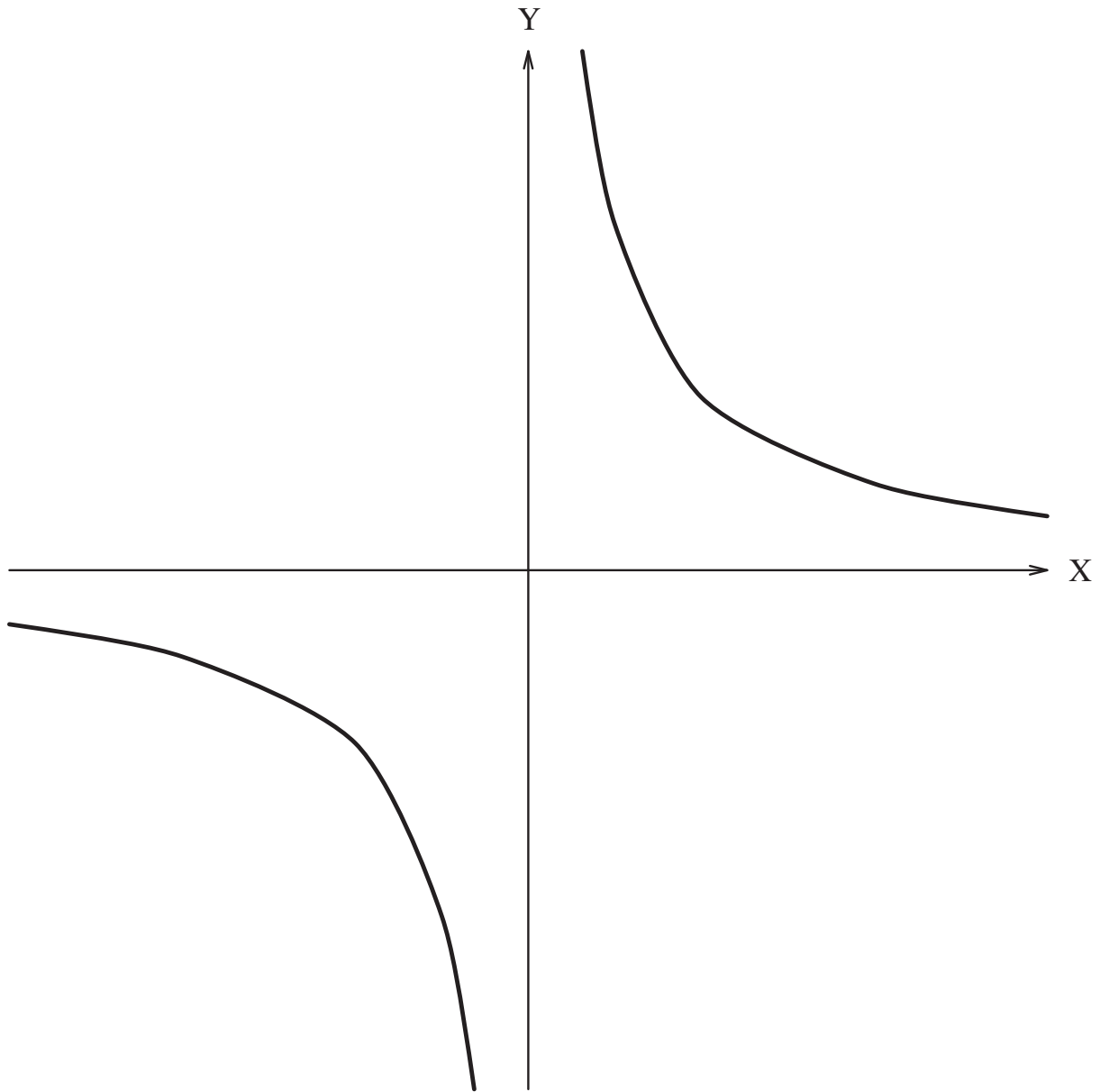
$$T \rightarrow \infty, \quad \frac{T^2}{T} \rightarrow \infty, \quad \frac{1}{T} \rightarrow 0, \quad (T; T^2, 1) \rightarrow (0; 1, 0)$$

These points are the vertical point at infinity and the horizontal point at infinity. As T gets small, X gets small and as T increases, X also increases; but the converse is true about Y .

2.1 Homogenizing the time coordinate

Formally, we express the time coordinate T in terms of homogeneous coordinates. Let $T = \frac{t}{s}$ giving us $(s; t)$. Homogenizing the hyperbola of Equation (3) gives us:

$$(T; T^2, 1) \rightarrow \left(\frac{t}{s}; \frac{t^2}{s^2}, 1\right) \rightarrow (st; t^2, s^2) \tag{4}$$

Figure 3: The curve $XY = 1$

2.2 Rational parameterization of a line

Figure 4 shows the midpoint of the projected line between $(1;0)$ and $(3;3)$. In this figure, the w value is plotted along the normal y -axis. This means that the point $(1;0)$ shows up as $(0,1)$. This was done, so that we could keep the weight coordinate on the left.

As before, the affine points lie along the line $w = 1$. If we look at $(w,x) = (1;0)$ or $(2;0)$ or $(3;0)$ all these points will project onto the point $(1;0)$. As time moves from $(1;0)$ to $(3;3)$ at a constant rate, where is the midpoint? We can interpolate and find the point $(2;\frac{3}{2})$. Projecting

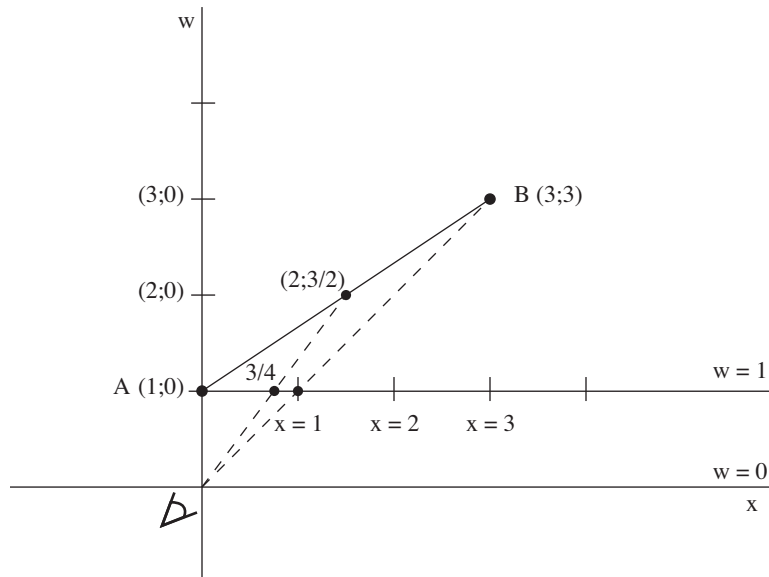


Figure 4: Projected midpoint of a line

this point onto the line $w = 1$, the point becomes $(1; \frac{3}{4})$. As shown in Figure 4, the projection $(1; \frac{3}{4})$ of the midpoint of the sites A and B is different from the midpoint of the projections of A and B . Since A and B project to $(1;0)$ and $(1;1)$, the midpoint of the projections is simply $(1; \frac{1}{2})$.

Until now, we have assumed that the weights of all control points are equal. Now, by varying the weights, we can move the halfway points around.

Assigning a weight of 1 to point A and 3 to point B , the halfway point becomes $(1; \frac{3}{4})$. Geometrically, the halfway point is pulled towards the control point with the larger weight. To see why this is the case, compute the value T using

$$H(T) = \frac{(1-T)w_a A + T w_b B}{(1-T)w_a + T w_b}.$$

Here, w_a is the weight associated with the control site A and w_b is the weight associated with B .

2.3 Interpolating in site space

To parameterize interpolation between points, first compute the the site coordinates and then interpolate the old way.

Let $A = (a_x, a_y)$ and $B = (b_x, b_y)$ be Cartesian coordinates for a rational curve. Homogenizing this gives

$$\begin{aligned} A &= (1; a_x, a_y) \\ B &= (1; b_x, b_y) \end{aligned} \tag{5}$$

For rational curves, weights — say u and v — are associated with (5) which gives us:

$$\begin{aligned} A &= (1; a_x, a_y) \times u \rightarrow (u; ua_x, ua_y) \\ B &= (1; b_x, b_y) \times v \rightarrow (v; vb_x, vb_y) \end{aligned}$$

For time T in $[0..1]$, interpolation using weights gives

$$\begin{aligned} &= (1 - T)(u; ua_x, ua_y) + (T)(v; vb_x, vb_y) \\ &= ((1 - T)u + Tv; (1 - T)ua_x + (T)vb_x, (1 - T)ua_y + (T)vb_y) \end{aligned} \quad (6)$$

Dividing through to get the Cartesian coordinates of the result, we have

$$\begin{aligned} X(T) &= \frac{x(T)}{w(T)} = \frac{(1-T)ua_x + Tvb_x}{(1-T)u + Tv} \\ Y(T) &= \frac{y(T)}{w(T)} = \frac{(1-T)ua_y + Tvb_y}{(1-T)u + Tv} \end{aligned} \quad (7)$$

For example, to find $T = \frac{1}{2}$ between A and B with the associated weights $u = 2$ and $v = 3$ using (7), we obtain

$$\frac{(1-T)2A + T3B}{(1-T)2 + 3T} = \frac{A + \frac{3}{2}B}{\frac{5}{2}} = \frac{2A + 3B}{5} \quad (8)$$

The midpoint is $\frac{3}{5}$ the way from A to B at time $T = \frac{1}{2}$ and the weight coordinate of the midpoint is $\frac{5}{2}$, precisely the value of the denominator of (8)

If the weights were all 1, notice that the denominator drops out and we get the unweighted parametric equation of

$$(1 - T)P + TQ = \frac{1}{2}P + \frac{1}{2}Q \quad (9)$$

2.4 Rational Quadratic Curves

Consider the first quadrant of a circle with the 3 Bézier control sites:

$$P = (1; 1, 0), \quad Q = (1; 1, 1), \quad R = (1; 0, 1), \quad (10)$$

If all the weights are 1, the 3 control sites determine a parabolic arc, and $T = \frac{1}{2}$ corresponds to

$$\begin{aligned} F(T) &= (1 - T)^2P + 2T(1 - T)Q + T^2R \\ F\left(\frac{1}{2}\right) &= \frac{1}{4}P + \frac{1}{2}Q + \frac{1}{4}R = \frac{1}{4}(1; 1, 0) + \frac{1}{2}(1; 1, 1) + \frac{1}{4}(1; 0, 1) = \left(1; \frac{3}{4}, \frac{3}{4}\right) \end{aligned} \quad (11)$$

As shown in Figure 5, we know that at time $T = \frac{1}{2}$, the quarter-circle is at $(1; \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. Suppose we want to put weights on the control points of the parabolic arc in order to get a quarter circle. In this example, we will leave the weights of the end control sites at 1 and vary the weight coordinate of Q .

We must determine the weight coordinate of Q , so that at time $T = \frac{1}{2}$, the resulting curve is at $(1; \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. Using a modified version of the Bernstein polynomials using weights gives us:

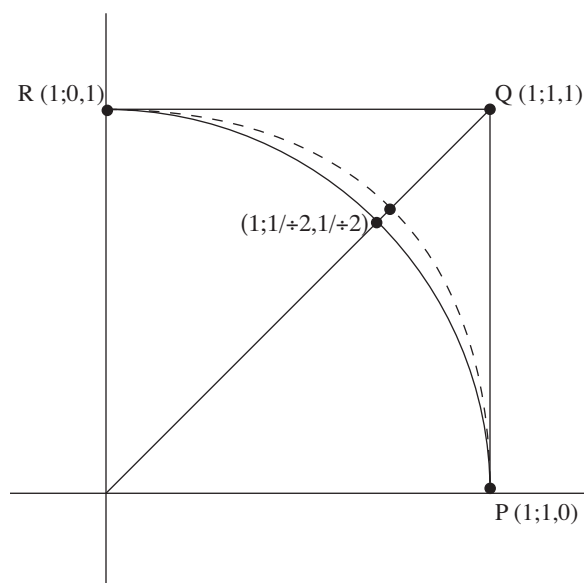


Figure 5: Unweighted (dotted) quadratic curve and (solid) quarter circle

$$F(T) = \frac{(1-T)^2 * 1 * P + 2T(1-T) * w * Q + T^2 * 1 * R}{(1-T)^2 * 1 + 2T(1-T) * w + T^2 * 1} \quad (12)$$

Substituting $\frac{1}{2}$ for T and P, Q, and R values into (12):

$$\frac{\frac{1}{4}(1;1,0) + \frac{w}{2}(1;1,1) + \frac{1}{4}(1;0,1)}{\frac{1}{4} + \frac{w}{2} + \frac{1}{4}}$$

Simplifying the equation and setting it to the desired point for $T = \frac{1}{2}$:

$$\left(1; \frac{2w+1}{2w+2}, \frac{2w+1}{2w+2}\right) = \left(1; \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

Which results in $w = \frac{1}{\sqrt{2}}$. Notice, since $\frac{1}{\sqrt{2}} < \frac{3}{4}$, we are relaxing the pull of the control site Q .

We can also use the de Casteljau algorithm to find where the curve is at time $T = \frac{1}{2}$. Determine point A by interpolating between points P and Q, with weights of 1 and $\frac{1}{\sqrt{2}}$, respectively. Similarly, determine point B by interpolating between Q and R with respective weights of $\frac{1}{\sqrt{2}}$ and 1. The weight of A is $\frac{1+\frac{1}{\sqrt{2}}}{2}$. Similarly, for B.

2.5 Degrees of freedom

How many degrees of freedom are there? In the previous example, the weights of P and R remained fixed, while the weight of the inner Bézier control point varied. It is also possible to

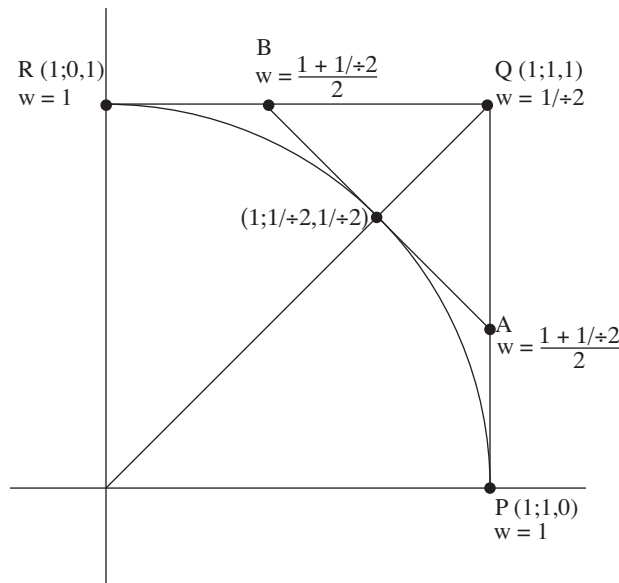


Figure 6: Weighting the middle Bézier point

vary the weight of one endpoint and achieve the same resulting curve. Though the curve will be the same, a different parameterization will give a different point for $T = \frac{1}{2}$.

For example, suppose P and Q each have a weight of 1 and R has a weight of 2. Using the values of $P = (1; 1, 0)$, $Q = (1; 1, 1)$, $R = (1; 0, 1)$, compute A, B, and C.

$$A = \frac{\frac{1}{2}P + \frac{1}{2}Q}{\frac{1}{2} + \frac{1}{2}} = \frac{P+Q}{2} = (1; 1, \frac{1}{2}), w_A = \frac{w_P + w_Q}{2} = 1$$

$$B = \frac{\frac{1}{2}Q + \frac{1}{2} * 2R}{\frac{1}{2} + \frac{1}{2} * 2} = \frac{Q+2R}{3} = (1; \frac{1}{3}, 1), w_B = \frac{w_R + w_Q}{2} = \frac{3}{2}$$

$$C = \frac{\frac{1}{2} * 1 * A + \frac{1}{2} * \frac{3}{2} B}{\frac{1}{2} * 1 + \frac{1}{2} * \frac{3}{2}} = \frac{2}{5}A + \frac{3}{5}B = (1; \frac{3}{5}, \frac{4}{5})$$

So, point C is certainly on the unit circle in the first quadrant. The point $F(\frac{1}{2})$ is referred to as the *shoulder point*. Note that it is the *ratio* of weights, rather than the values of the weights individually that determines the nature of the curve.

Given the three Bézier sites $f(t_1, t_1)$, $f(t_1, t_2)$, and $f(t_2, t_2)$ of a rational, quadratic segment $F([t_1 .. t_2])$, we can get different rational curves through $F(t_1)$ and $F(t_2)$ by juggling the weights of the three Bézier Sites. Two rules determine the behavior of a rational curve when we modify the weights of its Bézier sites:

1. If you change the weights of all Bézier sites by the same factor, you will get the same curve.

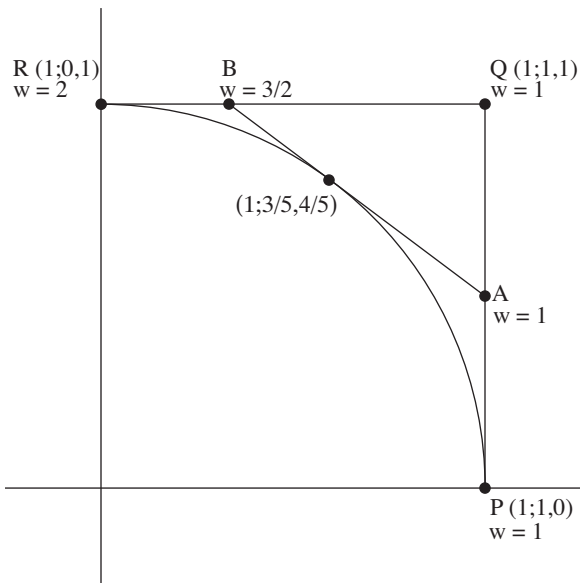


Figure 7: Weighting end Bézier point

Curve	Weights of Bézier Sites		
$F(t)$	w_0	w_1	w_2
$F(t)$	cw_0	cw_1	cw_2
Reparametrization of $F(t)$	w_0	cw_1	c^2w_2

Table 1: Effect of Weight on a Rational Curve

2. If you multiply the weights of consecutive Bézier sites by consecutive powers of a constant, then you will get a reparametrization of the same curve. The new curve will pass through the same points, but at different times.

Given that we have three degrees of freedom in choosing the weights of three Bézier sites of a quadratic, the above two rules leave us with only one degree of freedom that matters. The parameter

$$\frac{w_1^2}{w_0w_2}$$

controls which conic arc is determined by the three Bézier sites. If we change the weights w_0 , w_1 , and w_2 , but keep the ratio $w_1^2 : w_0w_2$ constant, we will get either the same curve or a reparametrization of it.

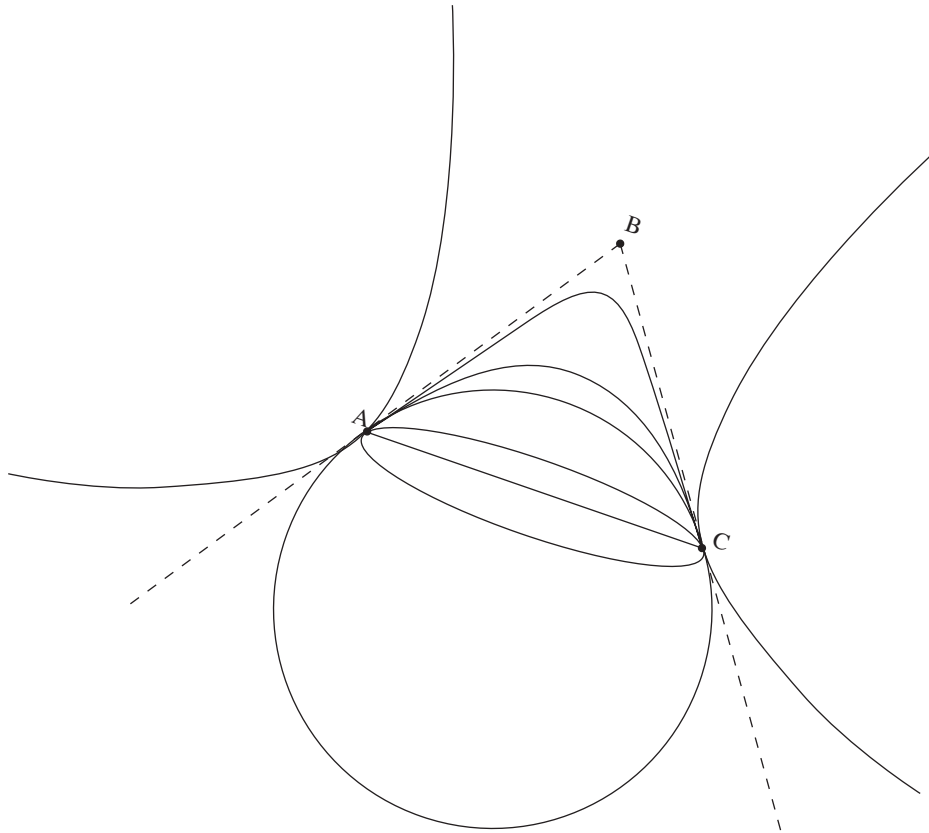


Figure 8: Quadratic rational curves

2.6 Visual character of a quadratic rational curve

What is the visual nature of a rational quadratic curve? What shape does it trace out? Consider the 3 Bézier points A, B and C, each with respective weights of w_a , w_b , and w_c . Suppose $w_a = w_c = 1$. How does the weight of B, w_b , affect the curve? If all weights have the same sign, the curve remains within $\triangle ABC$.

1. As w_b increases with respect to w_a and w_c , the curve is pulled tighter and moves toward point B. If w_b is very large, you get part of one branch of a hyperbola.
2. If $w_b = 1$, you get a portion of a parabolic curve.
3. As w_b decreases and remains positive, the curve relaxes and moves away from point B. So, a small w_b gives a curve close to the line AB.
4. If $w_b = 0$, the curve lies along the line AB. However, this means that it goes to ∞ . Remember the relation: $\frac{w_1^2}{w_0 w_2}$. A weight of 0 corresponds to ∞ .

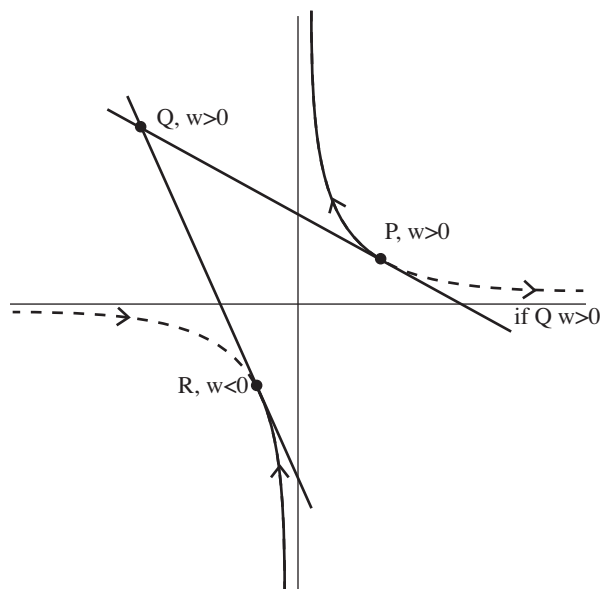


Figure 9: Section of quadratic when endpoints have different signs

- If $w_b < 0$, remembering that $w_a = w_c = 1$, you get the the other “half” of the curve, i.e., the portion that lies outside of $\triangle ABC$ but within $\angle ABC$.

When A and B have opposite signs, the resulting curve lies outside the $\angle ABC$. The segment of the curve followed depends on the sign of the weights. In Figure 9, P and Q have weights > 0 and R has a weight < 0 . The segment from P to R starts at P, moves toward Q, goes through the line at ∞ and continues to R.

Consider the following parametrization of the unit circle:

$$C(T) = [1 + T^2; 1 - T^2, 2T]$$

It is easy to see that the above curve parametrizes the unit circle, since for any point $[w; x, y] = [1 + T^2; 1 - T^2, 2T]$ on the circle, we have

$$\left(\frac{x}{w}\right)^2 + \left(\frac{y}{w}\right)^2 = \frac{(1 - T^2)^2 + 4T^2}{(1 + T^2)^2} = \frac{1 + 2T^2 + T^4}{(1 + T^2)^2} = 1.$$

The polarized form of the unit circle is:

$$c(T_1, T_2) = [1 + T_1 T_2; 1 - T_1 T_2, T_1 + T_2]$$

And the homogenized, polar form of the unit circle is:

$$c((s_1; t_1), ((s_2; t_2))) = [s_1 s_2 + t_1 t_2; s_1 s_2 - t_1 t_2, s_2 t_1 + s_1 t_2]$$

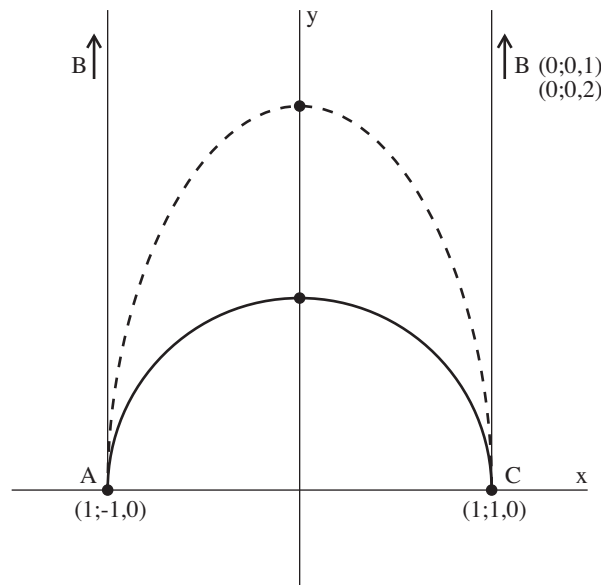


Figure 10: Middle Bézier point at infinity

Consider the curve over the interval $C[(0..1)]$. One way to specify the rational curve is by its 3 Bézier points, each of which is a value of the homogenized polar form.

In Figure 10, the middle Bézier point is the point at infinity.

$$C((1;0)(1;0)) = [1; 1, 0] = C(0)$$

$$C((1;0)(0;1)) = [0; 0, 1] = \text{point at infinity in the vertical direction}$$

$$C((0;1)(0;1)) = [1; -1, 0] = C(\infty)$$

Note that with rational curves, the control site $(0;0,1)$ is not the same as the site $(0;0,2)$ regarding its effect on the curve. When we scale the middle Bézier site up from $(0;0,1)$ to $(0;0,2)$, the semi-circle is pulled out into a half-ellipse (shown dashed). Though both sites $(0;0,1)$ and $(0;0,2)$ determine the same control point, the point at infinity in the vertical direction, the latter site pulls twice as hard as the former.

Note also that, when the control site is at infinity, it is no longer quite correct to talk about its *weight*. In the example above, for instance, both $(0;0,1)$ and $(0;0,2)$ have weight 0. More correctly, it is the *scale factor* that is really important. Increasing the scale factor increases the pull of that control site, and this happens even if the weight coordinate (before, and hence also after the change of scale factor) is 0.

2.7 Rational cubics

Using Bézier sites with weights, instead of Bézier points, works equally well for cubics, or rational curves of any degree. With cubics, we have four degrees of freedom in choosing

the weights, two of which aren't useful. Scaling all the weights by the same factor doesn't change the curve. Scaling the weights by factors in geometric progression just changes the parametrization. This leaves two degrees of freedom that we can use to get differently shaped curves.

2.8 Degenerate cases

Recall that we can treat a quadratic as a degenerate case of a cubic. For example, consider the parabola $F(T) = (1; T, T^2)$, which has the homogenized form $F(s, t) = (s^2; st, t^2)$. If G denotes the parabola F considered as a degenerate cubic, then we have $G(T) = (1; T, T^2)$ and $G(s, t) = (s^3; s^2t, st^2)$. Notice that, in the homogenized form, we have multiplied each coordinate by s .

In the rational case, we have more choice about how to raise the degree. Here is an example of a quadratic, rational curve:

$$F(T) = (T^2 - 3T + 4; 2T, 7T^2 - 6).$$

When we homogenize the quadratic F , we get

$$F(s, t) = (t^2 - 3st + 4s^2; 2st, 7t^2 - 6s^2).$$

One way to raise the degree of F is, as in the polynomial case, to multiply each homogeneous coordinate by s . This gives us the curve

$$G(s, t) = (st^2 - 3s^2t + 4s^3; 2s^2t, 7st^2 - 6s^3).$$

But we could equally well multiply each coordinate by t , or by some linear combination ($as + bt$). The resulting curve

$$H(s, t) = ((as + bt)(t^2 - 3st + 4s^2); 2st(as + bt), (as + bt)(7t^2 - 6s^2))$$

is also a degree-raised version of F .

3 Nonuniform Rational B-splines (NURBS)

Non-uniform means that the knot sequence is not necessarily uniformly spaced. Rational means that the curve is not necessarily a polynomial one. B-spline means that the curve is controlled by de Boor points.

Given the knot sequence $(0, 1, 2, 3, 4)$ with $n = 3$ (cubic), as shown in Figure 11, we can locate the point 223 by weighted interpolation between the de Boor points 123 and 234. If the weights of the de Boor points 123 and 234 were equal, then 223 would be located one-third of the way from 123 to 234. But, if the weights are not equal, we must take the weights into account. In the example in Figure 11, the weight of 123 is 1 while the weight of 234 is 2.

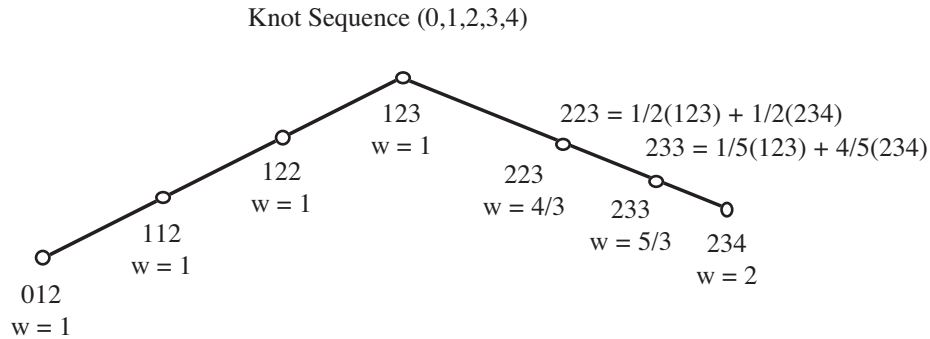


Figure 11: Interpolation on a Non-Uniform Rational B-spline

Using the point and weight method, we find that the location of P_{223} is

$$P_{223} = \frac{\frac{2}{3}(1)P_{123} + \frac{1}{3}(2)P_{234}}{\frac{2}{3}(1) + \frac{1}{3}(1)} = \frac{\frac{2}{3}P_{123} + \frac{2}{3}P_{234}}{\frac{4}{3}} = \frac{P_{123} + P_{234}}{2}$$

and the weight associated with P_{223} is $\frac{4}{3}$. In a similar manner, we can calculate the location of P_{233} to be

$$P_{233} = \frac{P_{123} + 4P_{234}}{5},$$

with a weight of $\frac{5}{3}$.

4 Tensor-product surfaces

Suppose that $F(u, v) = (X(u, v), Y(u, v), Z(u, v))$ is a parametric surface, and say that the function X is given by

$$X(u, v) = 2uv + 3uv^2 + 7u^2.$$

If we view F as a biquadratic, tensor-product surface, we treat u and v independently when computing the polar form $f = (x, y, z)$:

$$x(u_1, u_2; v_1, v_2) = 2 \left(\frac{u_1 + u_2}{2} \right) \left(\frac{v_1 + v_2}{2} \right) + 3 \left(\frac{u_1 + u_2}{2} \right) v_1 v_2 + 7u_1 u_2.$$

For a tensor-product surface, the natural domain region to take in the parameter plane is a rectangle aligned with the coordinate axes, since we are treating u and v separately. Such a rectangle maps to a rectangular surface patch. If we let u vary, say from r to s , and let v vary, say from p to q , the Bézier points of the rectangular surface patch $F([r..s] \times [p..q])$ can be found by substituting all combinations of r and s for the polar arguments u_i and all combinations of p and q for the polar arguments v_j . In the biquadratic case, the nine resulting Bézier points are:

$$\begin{matrix} f(s, s; p, p) & f(s, s; p, q) & f(s, s; q, q) \\ f(r, s; p, p) & f(r, s; p, q) & f(r, s; q, q) \\ f(r, r; p, p) & f(r, r; p, q) & f(r, r; q, q) \end{matrix}$$

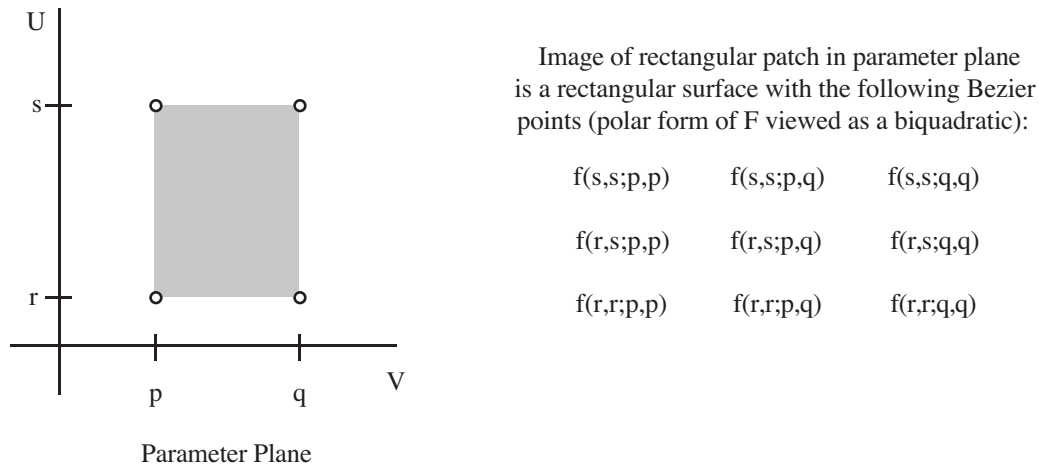


Figure 12: The nine Bézier points of a biquadratic patch

This is illustrated in Figure 12.

Of the nine Bézier points, the four corner Bézier points actually are the four corners of the patch $F([r..s] \times [p..q])$. The three Bézier points along one of the four edges are the Bézier points of one of the boundary curves of the patch $F([r..s] \times [p..q])$, each of which is a parabolic arc. For example, the boundary curve $F([r..s] \times \{p\})$, the image of the left edge of the domain rectangle, is simply the quadratic curve that has as its Bézier points the left column of Bezier points in Figure 12: $F(s,p) = f(s,s;p,p)$, $f(r,s;p,p)$, and $F(r,p) = f(r,r;p,p)$.

Given the Bézier points of a tensor-product surface, how do we evaluate a point on the surface, say $F(a,b)$? That is, how does the de Casteljau Algorithm generalize to the tensor-product case? Answer: We run the de Casteljau Algorithm in u and in v separately. For example, suppose that we want to evaluate the point $F(a,b)$ on the biquadratic surface of Figure 12. We can start by interpolating in u , controlled by the position of a in the interval $[r..s]$. This lets us compute $f(a,r;p,p)$ by interpolating between $f(r,r;p,p)$ and $f(r,s;p,p)$, and lets us compute $f(a,s;p,p)$ by interpolating between $f(r,s;p,p)$ and $f(s,s;p,p)$. Similarly, with four more interpolations, we can compute $f(a,r;p,q)$, $f(a,s;p,q)$, $f(a,r;q,q)$, and $f(a,s;q,q)$. The computation tree in Figure 13 shows one way to finish the process—the way in which we finish all of the u -interpolations before doing any of the v -interpolations.

If we liked, we could do the v -interpolations first. Or we could even intermix them, doing—say—one stage of u -interpolation, then one stage of v , then the second stage of u , then the second stage of v . Choosing different orders for the stages of interpolations in this way can lead to algorithms with different numbers of primitive interpolation steps, particularly in the case of tensor-product surface of degree $(m;n)$ with $m \neq n$. But we always get the same final answer.

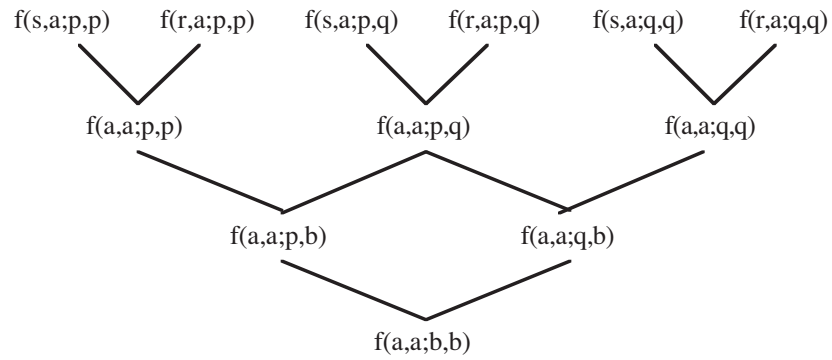


Figure 13: Part of the de Casteljau Algorithm for a biquadratic

References

- [1] Carl de Boor (1978), *A Practical Guide to Splines*, Springer-Verlag.
- [2] Gerald Farin (1988), *Curves and Surfaces for Computer Aided Geometric Design: A Practical Guide*, Academic Press.