An introduction to NURBS

Philippe Lavoie

January 20, 1999

A three dimensional (3D) object is composed of curves and surfaces. One must find a way to represent these to be able to model accurately an object. The two most common methods to represent a curve or a surface are the implicit and the parametric method.

The implicit method is a function which depends on the axis variables and is usually equal to 0. It describes a relationship between the axis variables. For example, the function $f(x,y) = x^2 + y^2 - 1 = 0$ represents a circle of radius 1.

In the parametric method each of the axis variables is a function of an independent parameter. In this form, a curve would be defined with the independent variable u as

$$C(u) = [x(u), y(u)] \qquad a \leq u \leq b$$

To represent the first quadrant of a circle in a parametric form, one can write

$$C(u) = [cos(u), sin(u)] \qquad 0 \leq u \leq \frac{\pi}{2}$$

or he can also write

$$C(u) = \left[\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right] \qquad 0 \leq t \leq 1$$

This shows that the parametric representation of a curve is not unique.

To visualize how a parametric curve is drawn, imagine that as time increases a new point on the curve is plotted. In the function above, the time is represented with the variable t, it goes from 0 to 1 and it generates a curve like the one in figure 1.

A class of parametric curves and surface is the **N**on-**U**niform **R**ational **B**-**S**pline (NURBS) curve or surface. NURBS are being used for computational reasons such as being easily processed by a computer, being stable to floating points errors and having little memory requirements and for the ability to represent any kind of curves or surface. They are the generalization of non-rational B-splines which are based on rational Bézier curves. Finally, the rational Bézier curve is a generalization of the Bézier curve which is studied first.

A Bézier curve of degree n is defined by

$$C(u) = \sum_{i=0}^{n} B_{i,n}(u) P_i \qquad 0 \leq u \leq 1 \tag{1}$$

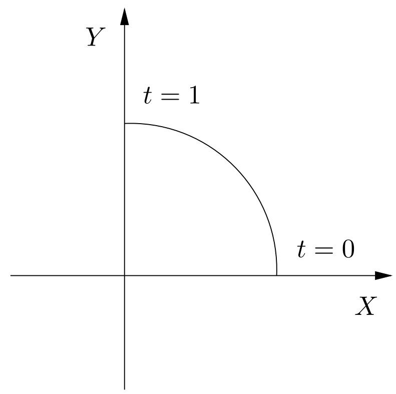

Figure 1: A quadrant of a circle generated with $C(u) = \left[\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right]$ for $0 \leq t \leq 1$

The geometric coefficients P_i are called *control points*. The basis functions $B_{i,n}$ are the classical nth-degree Bernstein polynomials [1] given by

$$B_{i,n}(u) = \frac{n!}{i!(n-i)!} u^i (1-u)^{n-i} \tag{2}$$

For interactive shape design, the control points of the Bézier curve convey a lot of geometric information as it can be seen with figure 2.

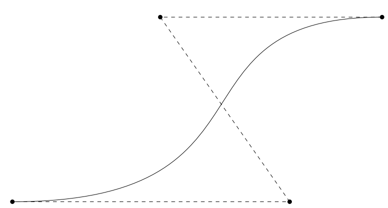

Figure 2: A Bézier curve of degree 3.

The Bézier curves can not represent conic curves. A conic curve (such as a circle) can be represented using a *rational function* which is defined as the ratio of two polynomials such as

$$x(u) = \frac{X(u)}{W(u)} \qquad y(u) = \frac{Y(u)}{W(u)} \qquad z(u) = \frac{Z(u)}{W(u)} \tag{3}$$

From this observation, the rational Bézier curve is defined as

$$C(u) = \frac{\sum_{i=0}^{n} B_{i,n}(u) w_i P_i}{\sum_{i=0}^{n} B_{i,n}(u) w_i} \qquad 0 \leq u \leq 1 \tag{4}$$

The P_i and $B_{i,n}$ are defined as before, the w_i are scalars called the *weights*. When the weights are varied, a *control point* will "attract" or "repulse" the curve more. This is best explained with an example. Four Bézier curves are drawn in figure 3. The only difference between them, is the weight of the control point P_2. The weight of 0.5 makes the curve go outside the boundary drawn by its control points. With the weight equal to 1, the curve is equivalent to the one depicted in figure 2. The weight of 2 tends to push the curve away from the second point and the weight of 10 pushes it even farther.

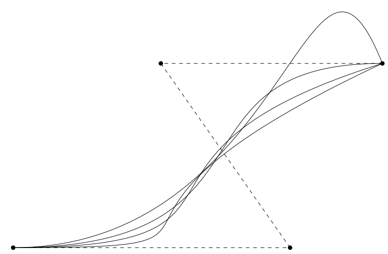

Figure 3: The effects of changing the weight of the second control point with $w = 0.5, 1, 2$ and 10.

A curve consisting of only one rational Bézier curve segment is often inadequate. The problems with a single segment range from the need of a high degree curve to accurately fit a complex shape, which is inefficient to process and is numerically unstable, to the need of interactive design for which a single segment has limitations as far as local control of the shape is concerned. To overcome these problems, a piecewise rational curve is used.

A piecewise Bézier curve or a B-spline curve is constructed from several Bézier curves joined together at some *breakpoints* with some level of continuity between them. Such a curve is depicted in figure 4.

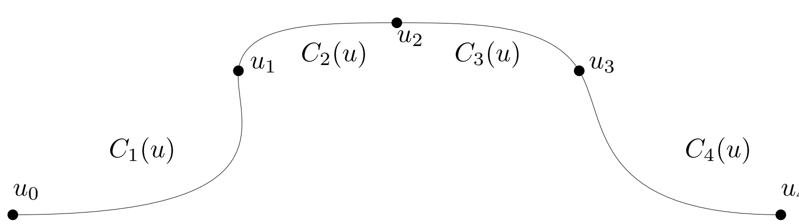

Figure 4: A piecewise rational Bézier curve.

The curve $C(u)$ is defined on $u \in [0, 1]$ and it is composed of the segments $C_i(u), 1 \leq i \leq m$. The segments are joined together at the breakpoints $u_0 = 0 < u_1 < u_2 < u_3 < u_4 = 1$ with some level of continuity. The curve is said to be C^k continuous at a breakpoint u_i if $C_i^{(j)}(u_i) = C_{i+1}^{(j)}(u_i)$ for all $0 \leq j \leq k$ where $C_i^{(j)}$ represents the jth derivative of C_i.

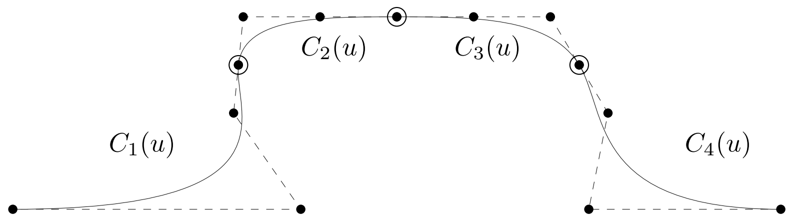

Figure 5: The Bézier segments composing the curve.

The different control points composing the B-spline curve are shown in figure 5. The control points circled with another circle are control points which are used by more than one Bézier segment. It should be clear that storing these points more than once is not memory efficient. If the curve is said to be C^1 continuous, then some of the points inside a Bézier segment are dependent on the position of the previous points to satisfy the continuity constraint. Therefore storing these points in memory is not necessary.

The equation of the B-spline should therefore be memory efficient and should also allows for the local control of the curve; *i.e.* the basis functions should not be defined over $[u_0, u_m]$, instead, they should be constrained to a limited number of subintervals. It is defined as

$$C(u) = \sum_{i=0}^{n} N_{i,p} P_i \qquad a \leq u \leq b \tag{5}$$

where P_i are the control points and $N_{i,p}$ are the pth degree B-spline basis functions.

There are different methods to define the B-spline basis functions: divided differences of truncated power functions [3], blossoming [6] and recurrence formula [2, 4, 5]. The recurrence definition is used since it is well suited to a computer implementation.

The B-spline has breakpoints which are named *knots* as seen in figure 4. A sequence of these knots is named the *knot vector* and it is defined as $U = u_0, \ldots, u_m$ which is a nondecreasing sequences of real numbers, *i.e.*, $u_i \leq u_{i+1}$ for $i = 0, \ldots, m$. The B-spline basis function of p-degree is defined using the recurrence formula as

$$\begin{aligned} N_{i,0}(u) &= \begin{cases} 1 & \text{if } u_i \leq u < u_{i+1} \\ 0 & \text{otherwise} \end{cases} \\ N_{i,p}(u) &= \frac{u - u_i}{u_{i+p} - u_i} N_{i,p-1}(u) + \frac{u_{i+p+1} - u}{u_{i+p+1} - u_{i+1}} N_{i+1,p-1}(u) \end{aligned} \tag{6}$$

The above equation can result in a $\frac{0}{/}0$ quotient; that quotient is defined to be zero. The knot vector of a B-spline curve is a non-periodic and non-uniform knot vector of the form

$$U = \{\underbrace{a, \ldots, a}_{p+1}, u_{p+1}, \ldots, u_{m-p+1}, \underbrace{b, \ldots, b}_{p+1}\} \tag{7}$$

The B-spline curve of figure 4 is presented with its *control polygon* in figure 6. The control polygon is the polygon formed by joining the control points P_i.

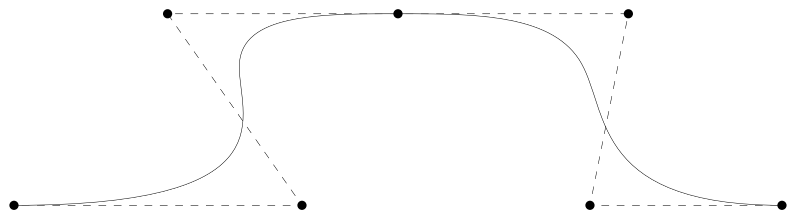

Figure 6: A B-spline curve with its control polygon.

As mentioned above, only rational functions can represent a conic curves therefore one could generalize the B-spline curve to obtain a rational representation. This generalization is named Non Uniform Rational B-Spline (NURBS) and it is defined as

$$C(u) = \frac{\sum_{i=0}^{n} N_{i,p}(u) w_i P_i}{\sum_{i=0}^{n} N_{i,p}(u) w_i} \qquad a \leq u \leq b \tag{8}$$

where P_i are the control points, w_i are the weights and $N_{i,p}$ are the B-spline basis functions defined on the non-periodic and non-uniform knot vector defined in equation 7.

Rational curves with coordinate functions in the form expressed in equation 3 have efficient processing and have an elegant geometric interpretation. The use of homogeneous coordinates can represent a rational curve in n dimensions as a polynomial curve of $n+1$ dimensions. The homogeneous control points are written as $P_i^w = w_i x_i, w_i y_i, w_i z_i, w_i$ in a four dimensional space where $w \neq 0$. To obtain P_i, we divide all the coordinates by the fourth coordinate w_i. This operation corresponds to a perspective map with the center at the origin. With these coordinates, the NURBS curve can be redefined as

$$C^w(u) = \sum_{i=0}^{n} N_{i,p}(u) P_i^w \tag{9}$$

NURBS are also used to represent surfaces. A NURBS surface in homogeneous coordinates is defined as

$$S^w(u,v) = \sum_{i=0}^{n} \sum_{j=0}^{m} N_{i,p}(u) N_{j,q}(v) P_{i,j}^w \tag{10}$$

where $P_{i,j}^w$ forms a bidirectional control net, and $N_{i,p}(u)$ and $N_{j,q}(v)$ are the non-rational B-spline basis functions defined on the knot vectors

$$\begin{aligned} U &= \{\underbrace{a, \ldots, a}_{p+1}, u_{p+1}, \ldots, u_{r-p+1}, \underbrace{b, \ldots, b}_{p+1}\} \\ V &= \{\underbrace{c, \ldots, c}_{q+1}, u_{q+1}, \ldots, u_{s-q+1}, \underbrace{d, \ldots, d}_{q+1}\} \end{aligned}$$

where $r = n+p+1$, $s = m+q+1$ and the limits $[a,b]$ and $[c,d]$ are usually always set to $[0,1]$.

References

[1] S.N. Bernstein. Démonstration du théorème de Weierstrass fondée sûr le calcul dès probabilités. *Commun. Soc. Math. Khrakow*, Vol. 12(No. 2):pages 1–2, 1912.

[2] M. G. Cox. The numerical evaluation of b-slines. *Jour. Inst. Math. Applic.*, Vol. 10:pages 134–149, 1972.

[3] H. B. Curry and I.J. Schoenberg. On spline distributions and their limits: the pòlya distribution functions. *Bull. Amer. Math. Soc.*, Vol. 53:page 109, 1947. Abstract 380t.

[4] C. De Boor. On calculating with b-splines. *Jour. Approx Theory*, Vol. 6:pages 50–62, 1972.

[5] C. De Boor. *A Practical Guide to Splines*. Springer-Verlag, New York, 1978.

[6] L. Ramshaw. Blossoming: A connect-the-dots approach to splines. Technical Report 19, Digital, System Research Center, Palo Alto, CA, 1987.