Computer Aided Hull Design

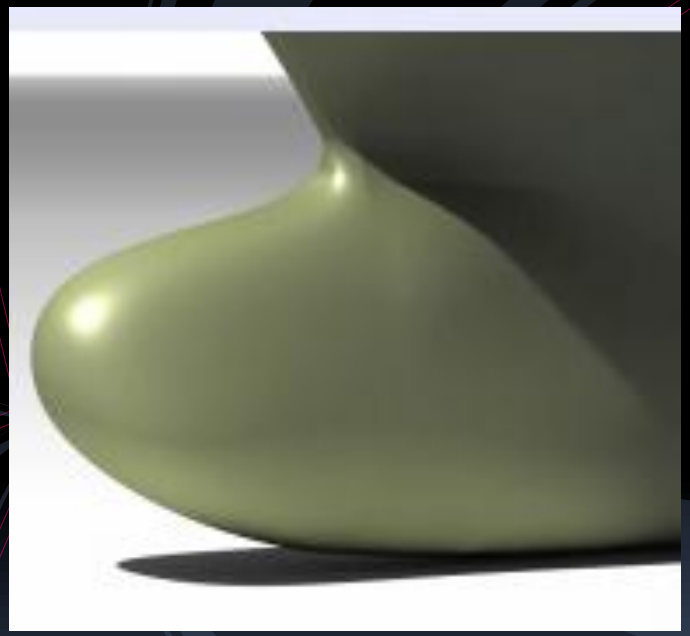
COURSE TEACHER

Dr. Md. Mashud Karim

Professor

Dept. of Naval Architecture and Marine Engineering

Bangladesh University of Engineering & Technology



NAME 6303: COMPUTER AIDED HULL DESIGN
3 CREDIT, 3 HRS./WK.

DETAILED SYLLABUS

- Analytic representation of a curve
- Advanced interpolation and control polygon techniques,
- Bezier and B-spline approximations,
- B-spline curve fitting.
- Form parameter of curves,
- Development of lines plan,

- Parametric surface representation,
- Blend generation,
- Partial differential equation (PDE) method for surface generation,
- Free form surface generation,
- Bezier surfaces, B-spline surfaces, Non-uniform rational B-spline (NURBS) surfaces,
- Surface design with volume constraints,
- Gaussian curvature and surface fairness.

- Generation and optimization of ship hull and propeller blade geometry.

Suggested Books

- Computational Geometry for Ships by H. Nowacki, M. I. G. Bloor and B. Oleksiewicz (Apr 1995)

- Mathematical Elements for Computer Graphics by David F. Rogers and J.Alan Adams (Nov 1, 1989)

- Curves and Surfaces for CAGD, Fifth Edition: A Practical Guide (The Morgan Kaufmann Series in Computer Graphics… by Gerald Farin (Nov 5, 2001)

Computer Aided Geometric Design

- Computer-aided geometric design deals with the mathematical description of shape for use in computer graphics, numerical analysis, approximation theory, data structures, and computer algebra.

- Computer-Aided Geometric Design(CAGD) is the basis for modern design in most branches of industry, from naval architecture and aeronautic to textile industry.

- In principle, students should just get familiar with some specific design applications, such as Rhino3D, Maxsurf, Foran,... But a thorough knowledge of their capabilities comes from learning at least the algorithms that lie behind the application, even if the students are not to become developers themselves.

- In recent decades, mathematical methods of computational geometry have made much progress and are successfully applied in many engineering disciplines

Computer Aided Hull Design

- **Computer Aided Hull Design (CAHD) is a specialized branch of CAGD.**

- The objective of this course is to review the mathematical background needed to understand the methods used by modern computer systems for ship design and manufacture

1.4 Parametric, Implicit, and Explicit Equations

There are basically three types of equations that can be used to define a planar curve: parametric, implicit, and explicit. The parametric equation of a plane curve takes the form

$$x = \frac{x(t)}{w(t)} \qquad y = \frac{y(t)}{w(t)}.$$

The implicit equation of a curve is of the form

$$f(x, y) = 0.$$

An explicit equation of a curve is a special case of both the parametric and implicit forms:

$$y = f(x).$$

Parametric Curve representation

A familiar way of representing a plane curve by using explicit non-parametric equations of the form ***y=f(x)***.

However, CAGD, it is more common to define curves parametrically in terms of a single scalar parameter.

Consider a curve $\mathbf{r}(t)$ defined by a vector-valued function of the parameter t by

$$\mathbf{r}(t) = (r_1(t), r_2(t), r_3(t)) \tag{2}$$

where the scalar functions $r_1(t), r_2(t), r_3(t)$ are called the coordinate functions of $\mathbf{r}$ and give the x, y and z coordinates of points along the line.

Thus we define a parametrized differentiable curve as a differentiable function $\mathbf{r}$ and can regard it as a mapping from the real-line R to Euclidean 3-space E^3 (Fig. 4). In CAGD, it is usual to consider a limited region of the real line, e.g. $0 \leq t \leq 1$.

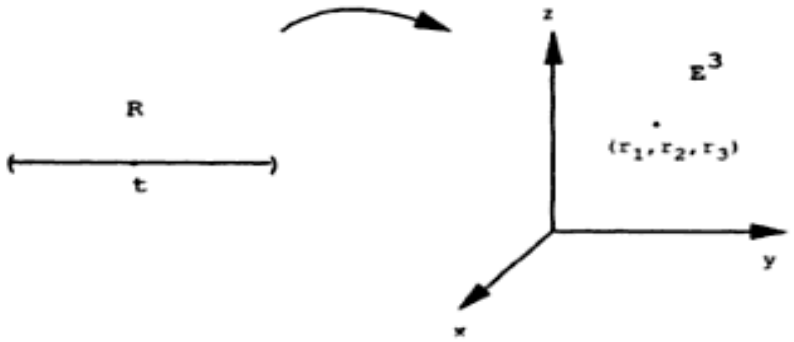


Fig. 4. The mapping from R into E^3

We further say that a curve $\mathbf{r}(t)$ of a real variable t is smooth (or differentiable) if its derivatives $\frac{dr_1}{dt}$, $\frac{dr_2}{dt}$ and $\frac{dr_3}{dt}$ all exist [13].

Example

- The parametrized differentiable curve given by $\mathbf{r}(t) = (a\cos t, a\sin t, bt)$ where t is real represents a helix of fixed pitch $2\pi b$ on the cylinder $x^2+y^2 = a^2$ in E^3. The parameter t measures the angle which the x-axis makes with the line joining the origin O to the projection of the point $\mathbf{r}(t)$ over the (x, y) plane.

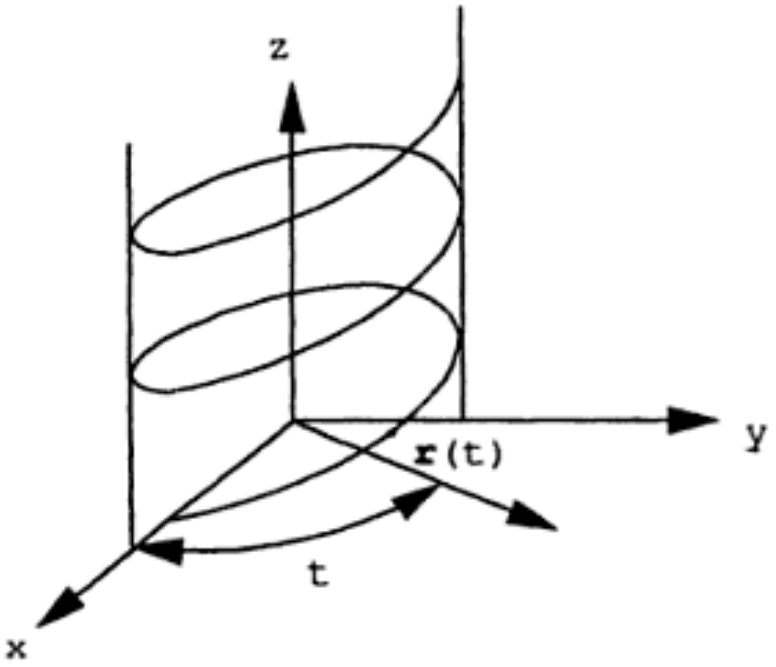


Fig. 5. The helix $\mathbf{r}(t) = (a\cos t, a\sin t, bt)$

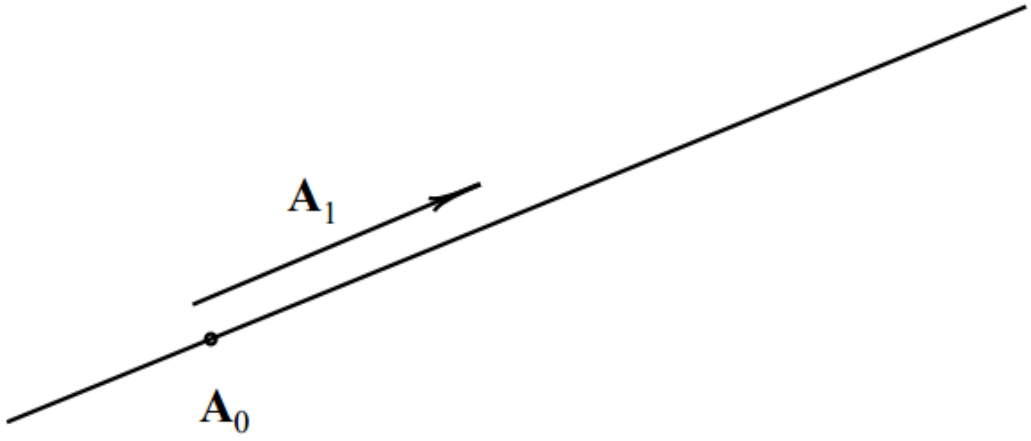
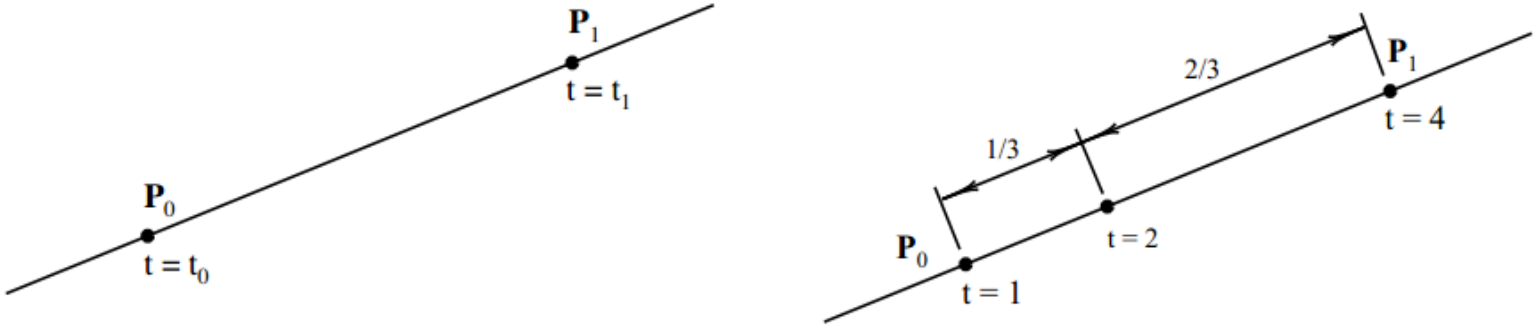


Figure 1.7: Line given by $\mathbf{A}_0 + \mathbf{A}_1 t$.



(a) Line given by $\mathbf{P}(t) = \frac{(t_1 - t)\mathbf{P}_0 + (t - t_0)\mathbf{P}_1}{t_1 - t_0}$.

(b) Affine Example.

Figure 1.8: Affine parametric equation of a line.

Affine parametric equation of a line

A line can also be expressed

$$\mathbf{P}(t) = \frac{(t_1 - t)\mathbf{P}_0 + (t - t_0)\mathbf{P}_1}{t_1 - t_0}$$

where $\mathbf{P}_0$ and $\mathbf{P}_1$ are two points on the line and t_0 and t_1 are any parameter values. Note that $\mathbf{P}(t_0) = \mathbf{P}_0$ and $\mathbf{P}(t_1) = \mathbf{P}_1$. Note in Figure 1.8.a that the line *segment* $\mathbf{P}_0$–$\mathbf{P}_1$ is defined by restricting the parameter:

$$t_0 \leq t \leq t_1.$$

Sometimes this is expressed by saying that the line segment is the portion of the line in the *parameter interval* or *domain* $[t_0, t_1]$.

We will see that the line in Figure 1.8.a is actually a degree one Bézier curve. Most commonly, we have $t_0 = 0$ and $t_1 = 1$ in which case

$$\mathbf{P}(t) = (1 - t)\mathbf{P}_0 + t\mathbf{P}_1.$$

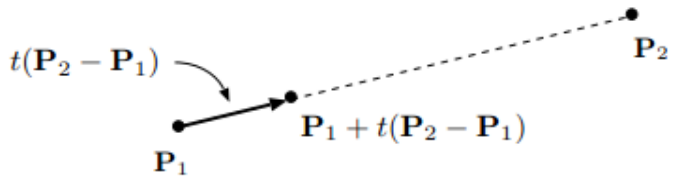
Equation 1.36 is called an *affine* equation, whereas equation 1.34 is called a *linear* equation. An affine equation is coordinate system independent, and is mainly concerned with ratios and proportions. An affine equation can be thought of as answering the question: "If a line is defined through two points $\mathbf{P}_0$ and $\mathbf{P}_1$, and if point $\mathbf{P}_0$ corresponds to parameter value t_0 and point $\mathbf{P}_1$ corresponds to parameter value t_1, what point corresponds to an arbitrary parameter value t?" Figure 1.8.b shows a line on which $\mathbf{P}_0$ corresponds to parameter $t = t_0 = 1$ and $\mathbf{P}_1$ is assigned

Affine Combinations of Two Points

Let $\mathbf{P}_1$ and $\mathbf{P}_2$ be points, and consider the expression

$$\mathbf{P} = \mathbf{P}_1 + t(\mathbf{P}_2 - \mathbf{P}_1)$$

This equation is meaningful, as $\mathbf{P}_2 - \mathbf{P}_1$ is a vector, and thus so is $t(\mathbf{P}_2 - \mathbf{P}_1)$. Therefore $\mathbf{P}$ is the sum of a point and a vector which is again a point. For each t, the point $\mathbf{P}$ represents a point on the line that passes through $\mathbf{P}_1$ and $\mathbf{P}_2$.



We note that if $0 \leq t \leq 1$ then $\mathbf{P}$ is somewhere on the line segment joining $\mathbf{P}_1$ and $\mathbf{P}_2$. If $t > 1$ then the point $\mathbf{P}$ is still on the line, but to the right of $\mathbf{P}_2$ in our illustration. Similarly if $t < 0$, then $\mathbf{P}$ is to the left of $\mathbf{P}_1$.

We can now define an **affine combination** of two points $\mathbf{P}_1$ and $\mathbf{P}_2$ to be

$$\mathbf{P} = \alpha_1\mathbf{P}_1 + \alpha_2\mathbf{P}_2$$

where $\alpha_1 + \alpha_2 = 1$. [$\mathbf{P} = (1-t)\mathbf{P}_1 + t\mathbf{P}_2$ is shown to be an affine transformation by setting $\alpha_2 = t$.]

1.5 Lines

The simplest case of a curve is a line. Even so, there are several different equations that can be used to represent lines.

1.5.1 Parametric equations of lines

Linear parametric equation

A line can be written in parametric form as follows:

$$x = a_0 + a_1 t; \qquad y = b_0 + b_1 t$$

In vector form,

$$\mathbf{P}(t) = \left\{ \begin{array}{c} x(t) \\ y(t) \end{array} \right\} = \left\{ \begin{array}{c} a_0 + a_1 t \\ b_0 + b_1 t \end{array} \right\} = \mathbf{A}_0 + \mathbf{A}_1 t. \tag{1.34}$$

In this equation, $\mathbf{A}_0$ is a point on the line and $\mathbf{A}_1$ is the direction of the line (see Figure 1.7)

Analytic Curves vs. Synthetic Curves

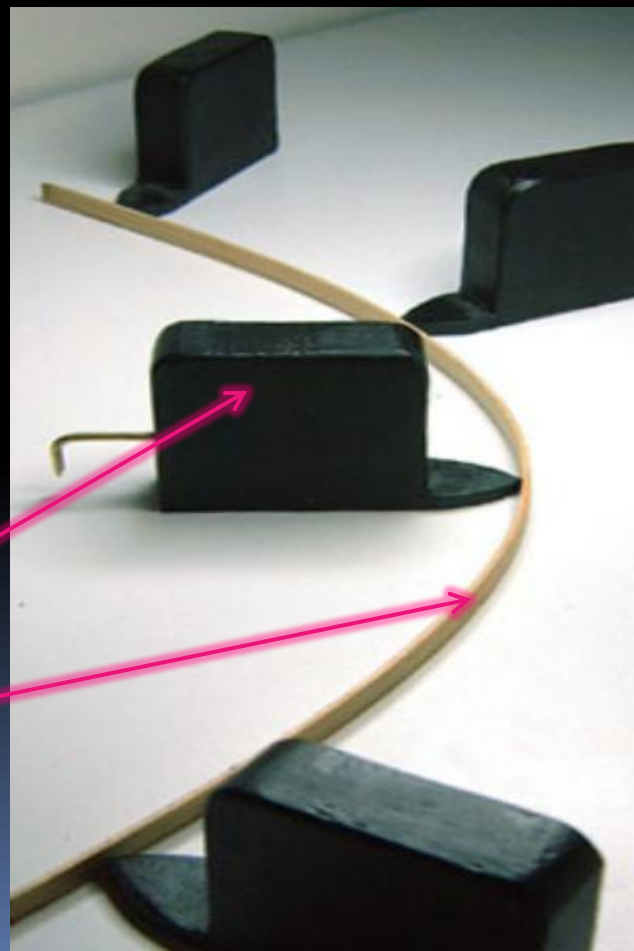
Analytic Curves

- Analytic Curves are points, lines, arcs and circles, fillets and chamfers, and conics (ellipses, parabolas, and hyperbolas)

Synthetic Curves

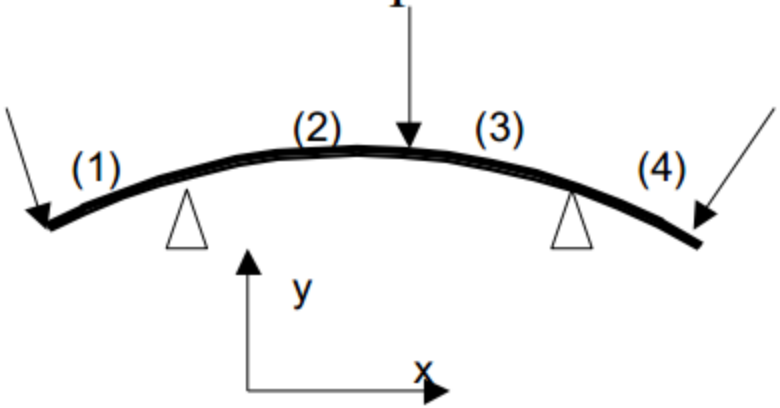
- Synthetic curves include various types of splines (cubic spline, B-spline, Beta-spline) and Bezier curves.

Fig.: Physical Spline and ducks



Spline

Splines — a mechanical beam with bending deflections, or a smooth curve under multiple constraints.



$$y''(x) = R(x) = \frac{M(x)}{EI} = \frac{a_i x + b_i}{EI}$$

$$y(x) = \frac{1}{EI}\left[\frac{a_i}{6}x^3 + \frac{b_i}{2}x^2 + c_i x + d_i\right]$$

Cubic Spline

spline between ducks is mathematically described by cubic polynomials. Thus we consider a mathematical spline modeled using cubic polynomials.

In general, the mathematical spline is a piecewise polynomial of degree K with continuity of derivatives of order $K-1$ at the common joints between segments. Thus, the cubic spline has second-order or C^2 continuity at the joints. Piecewise splines of low-degree polynomials are most useful for curve fitting because low-degree polynomials both reduce the computational requirements and also reduce numerical instabilities that arise with higher degree curves. These instabilities cause undesirable oscillations when several points are joined in a common curve. However, since low-degree polynomials cannot span an arbitrary series of points, adjacent polynomial segments are used. Based on these considerations and the analogy with the physical spline, a common technique is to use a series of cubic spline segments with each segment spanning only two points. Further, the cubic spline is advantageous since it is the lowest degree curve which allows a point of inflection and which has the ability to twist through space.

The equation for a single parametric cubic spline segment is given by

$$P(t) = \sum_{i=1}^{4} B_i \, t^{i-1} \qquad t_1 \leq t \leq t_2 \qquad (5-1)$$

where t_1 and t_2 are the parameter values at the beginning and end of the segment. $P(t)$ is the position vector of any point on the cubic spline segment. $P(t) = [\, x(t) \quad y(t) \quad z(t) \,]$ is a vector valued function. The three components of $P(t)$ are the Cartesian coordinates of the position vector.† Each component has a similar formulation to $P(t)$, i.e.,

$$x(t) = \sum_{i=1}^{4} B_{i_x}\, t^{i-1} \qquad t_1 \le t \le t_2$$

$$y(t) = \sum_{i=1}^{4} B_{i_y}\, t^{i-1} \qquad t_1 \le t \le t_2$$

$$z(t) = \sum_{i=1}^{4} B_{i_z}\, t^{i-1} \qquad t_1 \le t \le t_2$$

The constant coefficients B_i are determined by specifying four boundary conditions for the spline segment. Writing out Eq. (5–1) yields

$$P(t) = B_1 + B_2 t + B_3 t^2 + B_4 t^3 \qquad t_1 \le t \le t_2 \qquad (5-2)$$

Let P_1 and P_2 be the position vectors at the ends of the spline segment (see

---

†$P(t) = [\, r(t) \quad \theta(t) \quad z(t) \,]$, where $r(t), \theta(t), z(t)$ are considered components of a cylindrical coordinate system, is also perfectly acceptable, as are representations in other coordinate systems. (See Prob. 5–8.)

Fig. 5–5). Also let P_1' and P_2', the derivatives with respect to t, be the tangent vectors at the ends of the spline segment. Differentiating Eq. (5–1) yields

$$P'(t) = [\, x'(t) \quad y'(t) \quad z'(t) \,] = \sum_{i=1}^{4} B_i \, (i-1) t^{i-2} \qquad t_1 \le t \le t_2 \qquad (5-3)$$

Writing this result out gives

$$P'(t) = B_2 + 2B_3 t + 3B_4 t^2 \qquad t_1 \le t \le t_2 \qquad (5-4)$$

Assuming, without loss of generality, that $t_1 = 0$, and applying the four boundary conditions,

$$P(0) = P_1 \qquad (5-5a)$$

$$P(t_2) = P_2 \qquad (5-5b)$$

$$P'(0) = P_1' \qquad (5-5c)$$

$$P'(t_2) = P_2' \qquad (5-5d)$$

yields four equations for the unknown B_i's. Specifically,

$$P(0) = B_1 = P_1 \qquad (5-6a)$$

$$P'(0) = \sum_{i=1}^{4} (i-1)\, t^{i-2} B_i \Big|_{t=0} = B_2 = P_1' \qquad (5-6b)$$

yields four equations for the unknown B_i's. Specifically,

$$P(0) = B_1 = P_1 \qquad (5-6a)$$

$$P'(0) = \sum_{i=1}^{4} (i-1)\, t^{i-2} B_i \Big|_{t=0} = B_2 = P_1' \qquad (5-6b)$$

$$P(t_2) = \sum_{i=1}^{4} B_i\, t^{i-1} \Big|_{t=t_2} = B_1 + B_2 t_2 + B_3 t_2^2 + B_4 t_2^3 \qquad (5-6c)$$

$$P'(t_2) = \sum_{i=1}^{4} (i-1) t^{i-2} B_i \Big|_{t=t_2} = B_2 + 2B_3 t_2 + 3B_4 t_2^2 \qquad (5-6d)$$

Figure 5–5 Single cubic spline segment.

Solving for B_3 and B_4 yields

$$B_3 = \frac{3(P_2 - P_1)}{t_2^2} - \frac{2P_1'}{t_2} - \frac{P_2'}{t_2} \qquad (5-7a)$$

and

$$B_4 = \frac{2(P_1 - P_2)}{t_2^3} + \frac{P_1'}{t_2^2} + \frac{P_2'}{t_2^2} \qquad (5-7b)$$

These values of B_1, B_2, B_3 and B_4 determine the cubic spline segment. Note that the shape of the segment depends on the position and tangent vectors at the ends of the segment. Further, notice that the value of the parameter $t = t_2$ at the end of the segment occurs in the results. Since each of the end position and tangent vectors has three components, the parametric equation for a cubic space curve depends on twelve vector components and the parameter value t_2 at the end of the segment.

Substituting Eqs. (5–6) and (5–7) into Eq. (5–1) yields the equation for a single cubic spline segment:

$$P(t) = P_1 + P_1't + \left[\frac{3(P_2 - P_1)}{t_2^2} - \frac{2P_1'}{t_2} - \frac{P_2'}{t_2}\right]t^2 + \left[\frac{2(P_1 - P_2)}{t_2^3} + \frac{P_1'}{t_2^2} + \frac{P_2'}{t_2^2}\right]t^3 \qquad (5-8)$$

Equation (5–8) is for a single cubic spline segment. However, to represent a complete curve, multiple segments are joined together. Two adjacent segments

Equation (5–8) is for a single cubic spline segment. However, to represent a complete curve, multiple segments are joined together. Two adjacent segments are shown in Fig. 5–6. Provided that the position vectors P_1, P_2, P_3, the tangent vectors P_1', P_2', P_3' and the parameter values t_2, t_3 are known, then Eq. (5–8), applied to each of the two segments, yields their shapes. However, it is unlikely that the tangent vector P_2' at the internal joint between the two segments is known. Fortunately, the internal tangent vector P_2' can be determined by imposing a continuity condition at the internal joint.

Recall that a piecewise spline of degree K has continuity of order $K-1$ at the internal joints. Thus, a cubic spline has second-order continuity at the internal joints. This means that the second derivative $P_2''(t)$ is continuous across the joint; i.e., the curvature is continuous across the joint. Differentiating Eq. (5–1) twice yields

$$P''(t) = \sum_{i=1}^{4} (i-1)(i-2)B_i\, t^{i-3} \qquad t_1 \le t \le t_2 \qquad (5-9)$$

Noting that for the first cubic spline segment the parameter range is $0 \le t \le t_2$, evaluating Eq. (5–9) at the end of the segment where $t = t_2$ gives

$$P'' = 6B_4 t_2 + 2B_3$$

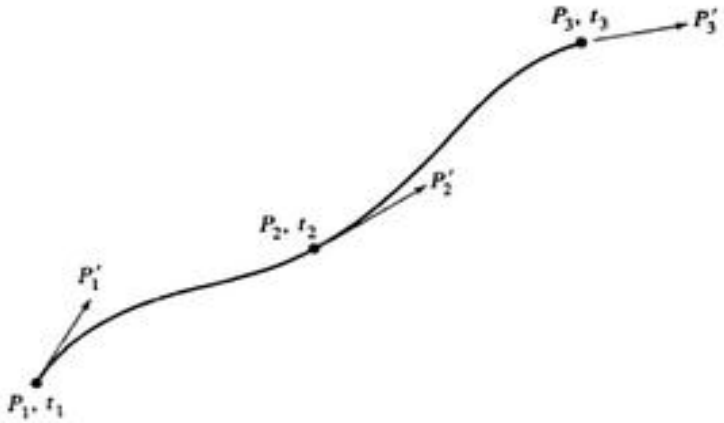


Figure 5–6 Two piecewise cubic spline segments.

For the second cubic spline segment the parameter range is $0 \le t \le t_3$. Evaluating Eq. (5–9)at the beginning of this second segment where $t = 0$ yields

$$P'' = 2B_3$$

Equating these two results and using Eqs. (5–6a, b) and (5–7a) yields

$$6t_2\left[\frac{2(P_1 - P_2)}{t_2^3} + \frac{P_1'}{t_2^2} + \frac{P_2'}{t_2^2}\right] + 2\left[\frac{3(P_2 - P_1)}{t_2^2} - \frac{2P_1'}{t_2} - \frac{P_2'}{t_2}\right]$$

$$= 2\left[\frac{3(P_3 - P_2)}{t_3^2} - \frac{2P_2'}{t_3} - \frac{P_3'}{t_3}\right]$$

Here the left-hand side of the equation represents the curvature at the end of the first segment and the right-hand side the curvature at the beginning of the second segment. Multiplying by t_2t_3 and collecting terms gives

$$t_3P_1' + 2(t_3 + t_2)P_2' + t_2P_3' = \frac{3}{t_2t_3}\left[t_2^2(P_3 - P_2) + t_3^2(P_2 - P_1)\right] \qquad (5-10)$$

which can be solved for P_2', the unknown tangent vector at the internal joint. Again notice that the end values of the parameter t, i.e., t_2 and t_3, occur in the resulting equation.

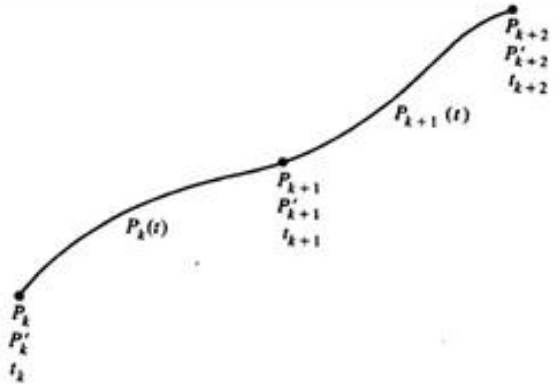


Figure 5–7 Notation for multiple piecewise cubic spline segments.

$$P_k(t) = P_k + P'_k t + \left[\frac{3(P_{k+1} - P_k)}{t_{k+1}^2} - \frac{2P'_k}{t_{k+1}} - \frac{P'_{k+1}}{t_{k+1}}\right]t^2 + \left[\frac{2(P_k - P_{k+1})}{t_{k+1}^3} + \frac{P'_k}{t_{k+1}^2} + \frac{P'_{k+1}}{t_{k+1}^2}\right]t^3 \qquad (5-11)$$

for the first segment, and

$$P_{k+1}(t) = P_{k+1} + P'_{k+1} t + \left[\frac{3(P_{k+2} - P_{k+1})}{t_{k+2}^2} - \frac{2P'_{k+1}}{t_{k+2}} - \frac{P'_{k+2}}{t_{k+2}}\right]t^2 + \left[\frac{2(P_{k+1} - P_{k+2})}{t_{k+2}^3} + \frac{P'_{k+1}}{t_{k+2}^2} + \frac{P'_{k+2}}{t_{k+2}^2}\right]t^3 \qquad (5-12)$$

for the second segment. Recalling that the parameter range begins at zero for each segment, for the first segment $0 \le t \le t_{k+1}$ and for the second $0 \le t \le t_{k+2}$.

For any two adjacent spline segments, equating the second derivatives at the common internal joint, i.e., letting $P''_k(t_k) = P''_{k+1}(0)$, yields the generalized result, equivalent to Eq. (5–10), i.e.,

$$t_{k+2}P'_k + 2(t_{k+1} + t_{k+2})P'_{k+1} + t_{k+1}P'_{k+2} = \frac{3}{t_{k+1}t_{k+2}}\left[t_{k+1}^2(P_{k+2} - P_{k+1}) + t_{k+2}^2(P_{k+1} - P_k)\right] \qquad 1 \le k \le n-2 \qquad (5-13)$$

for determining the tangent vector at the internal joint between any two spline segments P_k and P_{k+1}.

Applying Eq. (5–13) recursively over all the spline segments yields $n-2$ equations for the tangent vectors P'_k, $2 \le k \le n-1$. In matrix form the result is

$$\begin{bmatrix} t_3 & 2(t_2+t_3) & t_2 & 0 & \cdot & \cdot & \cdot & \cdot \\ 0 & t_4 & 2(t_3+t_4) & t_3 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & t_5 & 2(t_4+t_5) & t_4 & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 & t_n & 2(t_n+t_{n-1}) & t_{n-1} \end{bmatrix} \times$$

$$\begin{bmatrix} P'_1 \\ P'_2 \\ P'_3 \\ \cdot \\ \cdot \\ P'_n \end{bmatrix} = \begin{bmatrix} \frac{3}{t_2 t_3}\{t_2^2(P_3-P_2)+t_3^2(P_2-P_1)\} \\ \frac{3}{t_3 t_4}\{t_3^2(P_4-P_3)+t_4^2(P_3-P_2)\} \\ \cdot \\ \cdot \\ \cdot \\ \frac{3}{t_{n-1}t_n}\{t_{n-1}^2(P_n-P_{n-1})+t_n^2(P_{n-1}-P_{n-2})\} \end{bmatrix} \qquad (5-14)$$

or

$$[M^*][P'] = [R]$$

Since there are only $n-2$ equations for the n tangent vectors, $[M^*]$ is not square and thus cannot be inverted to obtain the solution for $[P']$; i.e., the problem is indeterminate. By assuming that the end tangent vectors P'_1 and P'_n are known, the problem becomes determinant. The matrix formulation is now

$$\begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ t_3 & 2(t_2+t_3) & t_2 & 0 & \cdot & \cdot & \cdot & \cdot \\ 0 & t_4 & 2(t_3+t_4) & t_3 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & t_5 & 2(t_4+t_5) & t_4 & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 & t_n & 2(t_n+t_{n-1}) & t_{n-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 1 \end{bmatrix} \times$$

$$\begin{bmatrix} P'_1 \\ P'_2 \\ P'_3 \\ \cdot \\ \cdot \\ \cdot \\ P'_n \end{bmatrix} = \begin{bmatrix} P'_1 \\ \frac{3}{t_2 t_3}\{t_2^2(P_3 - P_2) + t_3^2(P_2 - P_1)\} \\ \frac{3}{t_3 t_4}\{t_3^2(P_4 - P_3) + t_4^2(P_3 - P_2)\} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \frac{3}{t_{n-1} t_n}\{t_{n-1}^2(P_n - P_{n-1}) + t_n^2(P_{n-1} - P_{n-2})\} \\ P'_n \end{bmatrix} \qquad (5-15)$$

or $$[M][P'] = [R]$$

where $[M]$ is now square and invertible. Notice also that $[M]$ is tridiagonal,† which reduces the computational work required to invert it. Further, $[M]$ is diagonally dominant.‡ Hence it is nonsingular, and inversion yields a unique solution. The solution for $[P']$ is thus

$$[P'] = [M]^{-1}[R] \qquad (5-16)$$

Once the P'_k's are known, the B_i coefficients for each spline segment can be determined. Generalizing Eqs. (5–6) – (5–11) yields

$$B_{1k} = P_k$$

$$B_{2k} = P'_k$$

$$B_{3k} = \frac{3(P_{k+1} - P_k)}{t_{k+1}^2} - \frac{2P'_k}{t_{k+1}} - \frac{P'_{k+1}}{t_{k+1}}$$

$$B_{4k} = \frac{2(P_k - P_{k+1})}{t_{k+1}^3} + \frac{P'_k}{t_{k+1}^2} + \frac{P'_{k+1}}{t_{k+1}^2}$$

Recalling that the P_k's and P'_k's are vector valued confirms that the B_i's are also vector valued; i.e., if the P_k's and P'_k's have x, y, z components then the B_i's also have x, y, z components.

---

†A tridiagonal matrix is one in which coefficients appear only on the main, first upper and first lower diagonals.

‡In a diagonally dominant matrix the magnitude of the terms on the main diagonal exceed that of the off-diagonal terms on the same row.

In matrix form these equations for any spline segment k are

$$[\,B\,] = \begin{bmatrix} B_{1k} \\ B_{2k} \\ B_{3k} \\ B_{4k} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -\dfrac{3}{t_{k+1}^2} & -\dfrac{2}{t_{k+1}} & \dfrac{3}{t_{k+1}^2} & -\dfrac{1}{t_{k+1}} \\ \dfrac{2}{t_{k+1}^3} & \dfrac{1}{t_{k+1}^2} & -\dfrac{2}{t_{k+1}^3} & \dfrac{1}{t_{k+1}^2} \end{bmatrix} \begin{bmatrix} P_k \\ P'_k \\ P_{k+1} \\ P'_{k+1} \end{bmatrix} \qquad (5-17)$$

To generate a piecewise cubic spline through n given position vectors P_k, $1 \le k \le n$, with end tangent vectors P'_1 and P'_n, Eq. (5–16) is used to determine the internal tangent vectors P'_k, $2 \le k \le n-1$. Then for each piecewise cubic spline segment the end position and tangent vectors for that segment are used to determine the B_{ik}'s, $1 \le i \le 4$ for that segment using Eq. (5–17). Finally the generalization of Eq. (5–1)

$$P_k(t) = \sum_{i=1}^{4} B_{ik} t^{i-1} \qquad 0 \le t \le t_{k+1}, \quad 1 \le k \le n \leftarrow 1 \qquad (5-18)$$

is used to determine points on the spline segment.

In matrix form Eq. (5–18) becomes

$$P_k(t) = [\,1 \quad t \quad t^2 \quad t^3\,] \begin{bmatrix} B_{1k} \\ B_{2k} \\ B_{3k} \\ B_{4k} \end{bmatrix} \qquad 0 \le t \le t_{k+1} \qquad (5-19)$$

Substituting Eq. (5–17) and rearranging yields

$$P_k(\tau) = [\,F_1(\tau) \quad F_2(\tau) \quad F_3(\tau) \quad F_4(\tau)\,] \begin{bmatrix} P_k \\ P_{k+1} \\ P'_k \\ P'_{k+1} \end{bmatrix} \qquad \begin{array}{l} 0 \le \tau \le 1 \\ 1 \le k \le n-1 \end{array} \qquad (5-20)$$

where

$$\tau = (t/t_{k+1})$$

$$F_{1k}(\tau) = 2\tau^3 - 3\tau^2 + 1 \qquad (5-21a)$$

$$F_{2k}(\tau) = -2\tau^3 + 3\tau^2 \qquad (5-21b)$$

$$F_{3k}(\tau) = \tau(\tau^2 - 2\tau + 1)t_{k+1} \qquad (5-21c)$$

$$F_{4k}(\tau) = \tau(\tau^2 - \tau)t_{k+1} \qquad (5-21d)$$

are called blending or weighting functions.

Using the definitions of the blending functions, Eq. (5–20) is written in matrix form as

$$P_k(\tau) = [\,F\,][\,G\,] \qquad (5-22)$$

where $[\,F\,]$ is a blending function matrix given by

$$[\,F\,] = [\,F_1(\tau) \quad F_2(\tau) \quad F_3(\tau) \quad F_4(\tau)\,] \qquad (5-23)$$

and

$$[\,G\,]^T = [\,P_k \quad P_{k+1} \quad P'_k \quad P'_{k+1}\,] \qquad (5-24)$$

contains the geometric information. As we shall see, equations of the form of Eqs. (5–22), i.e., a matrix of blending functions times a matrix of geometric conditions, frequently appear in curve and surface descriptions.

Notice from Eqs. (5–21) that each of the blending functions is a cubic. Any point on a cubic spline segment is a weighted sum of the end position and tangent vectors. The F_{ik}'s act as the blending or weighting functions. Figure 5–8 shows

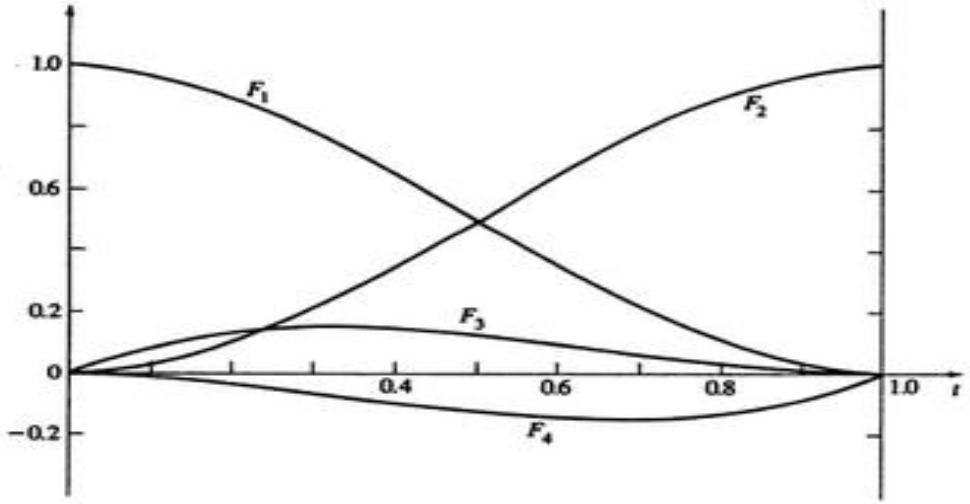


Figure 5–8 Cubic spline blending functions for $t_{k+1} = 1.0$.

the F_i's for $t_{k+1} = 1.0$. Notice from Fig. 5–8 that $F_1(0) = 1$ and $F_2(0) = F_3(0) = F_4(0) = 0$. Thus, the curve passes through the end position vector P_1. Similarly $F_2(1) = 1$ and $F_1(1) = F_3(1) = F_4(1) = 0$. Thus, the curve also passes through the end position vector P_2. Further, note the symmetry of F_1 and F_2 and of F_3 and F_4. In fact, $F_2(\tau) = 1 - F_1(\tau)$. Finally, note the relative magnitudes of F_1 and F_2 and of F_3 and F_4. This significant difference in magnitude shows that in general the end position vectors have relatively more influence than the end tangent vectors.

Recall that a piecewise cubic spline curve is determined by the position vectors, tangent vectors *and* the parameter values, i.e., the t_k's at the end of each segment. The choice of the t_k's affects the curve smoothness.

Continuity of the second derivatives at the internal joints does not in itself produce a fair or smooth cubic spline curve in the sense of minimum curvature along the curve. To obtain minimum curvature and hence maximum smoothness, the coefficients B_3 and B_4 must be minimized for each segment by choosing appropriate values for the t_k's for each segment. This additional computational effort is normally not required. Simpler methods, such as those described here, are used to generate curves smooth and fair enough for most practical purposes.

One approach used to determine the t_k's is to set the parameter values equal to the chord lengths between successive data points. Acceptably smooth curves for most practical purposes are obtained using this technique. A second approach is to normalize the variation by choosing $t_k = 1.0$ for each cubic segment. This choice simplifies the computational requirements (see Sec. 5–4). As can be seen from the previous equations, each choice of t_k produces different coefficient values and, hence, different curves through the given data points.

An example more fully illustrates the procedure.†

Example 5–2 Cubic Spline Curve

Considering the four two-dimensional position vectors $P_1[0 \quad 0]$, $P_2[1 \quad 1]$, $P_3[2 \quad -1]$ and $P_4[3 \quad 0]$ (see Fig. 5–9), determine the piecewise cubic spline curve through them using the chord approximation for the t_k's. The tangent vectors at the ends are $P_1'[1 \quad 1]$ and $P_4'[1 \quad 1]$. Calculate intermediate points at $\tau = 1/3,\ 2/3$ for each segment.

To begin, determine the t_k's:

$$t_2 = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} = \sqrt{(1)^2 + (1)^2} = \sqrt{2}$$

$$t_3 = \sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2} = \sqrt{(1)^2 + (-2)^2} = \sqrt{5}$$

$$t_4 = \sqrt{(x_4 - x_3)^2 + (y_4 - y_3)^2} = \sqrt{(1)^2 + (1)^2} = \sqrt{2}$$

---

†Two-dimensional examples are used throughout this chapter to simplify the calculations and the presentation of results. Three dimensions is a simple extension (see the problems for Chapter 5).

$$\begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ t_3 & 2(t_2+t_3) & t_2 & 0 & \cdot & \cdot & \cdot & \cdot \\ 0 & t_4 & 2(t_3+t_4) & t_3 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & t_5 & 2(t_4+t_5) & t_4 & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 & t_n & 2(t_n+t_{n-1}) & t_{n-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 1 \end{bmatrix} \times$$

$$\begin{bmatrix} P_1' \\ P_2' \\ P_3' \\ \cdot \\ \cdot \\ \cdot \\ P_n' \end{bmatrix} = \begin{bmatrix} P_1' \\ \frac{3}{t_2 t_3}\{t_2^2(P_3-P_2)+t_3^2(P_2-P_1)\} \\ \frac{3}{t_3 t_4}\{t_3^2(P_4-P_3)+t_4^2(P_3-P_2)\} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \frac{3}{t_{n-1} t_n}\{t_{n-1}^2(P_n-P_{n-1})+t_n^2(P_{n-1}-P_{n-2})\} \\ P_n' \end{bmatrix} \qquad (5-15)$$

$$F_{1k}(\tau) = 2\tau^3 - 3\tau^2 + 1 \qquad (5-21a)$$

$$F_{2k}(\tau) = -2\tau^3 + 3\tau^2 \qquad (5-21b)$$

$$F_{3k}(\tau) = \tau(\tau^2 - 2\tau + 1)t_{k+1} \qquad (5-21c)$$

$$F_{4k}(\tau) = \tau(\tau^2 - \tau)t_{k+1} \qquad (5-21d)$$

The internal tangent vectors P_2' and P_3' are found using Eq. (5–15). Specifically,

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ t_3 & 2(t_2+t_3) & t_2 & 0 \\ 0 & t_4 & 2(t_3+t_4) & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} P_1' \\ P_2' \\ P_3' \\ P_4' \end{bmatrix} =$$

$$\begin{bmatrix} P_1' \\ \dfrac{3}{t_2 t_3}\{t_2^2(P_3 - P_2) + t_3^2(P_2 - P_1)\} \\ \dfrac{3}{t_3 t_4}\{t_3^2(P_4 - P_3) + t_4^2(P_3 - P_2)\} \\ P_4' \end{bmatrix}$$

Substituting yields

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ \sqrt{5} & 2(\sqrt{2}+\sqrt{5}) & \sqrt{2} & 0 \\ 0 & \sqrt{2} & 2(\sqrt{5}+\sqrt{2}) & \sqrt{5} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} P_1' \\ P_2' \\ P_3' \\ P_4' \end{bmatrix}$$

$$= \begin{bmatrix} [1 \quad 1] \\ \frac{3}{\sqrt{2}\sqrt{5}}\{2([2 \quad -1]-[1 \quad 1])+5([1 \quad 1]-[0 \quad 0])\} \\ \frac{3}{\sqrt{5}\sqrt{2}}\{5([3 \quad 0]-[2 \quad -1])+2([2 \quad -1]-[1 \quad 1])\} \\ [1 \quad 1] \end{bmatrix}$$

$$= \begin{bmatrix} [1 \quad 1] \\ \frac{3}{\sqrt{2}\sqrt{5}}\{[2 \quad -4]+[5 \quad 5]\} \\ \frac{3}{\sqrt{5}\sqrt{2}}\{[5 \quad 5]+[2 \quad -4]\} \\ [1 \quad 1] \end{bmatrix} = \begin{bmatrix} [1 \quad 1] \\ 0.949\,[7 \quad 1] \\ 0.949\,[7 \quad 1] \\ [1 \quad 1] \end{bmatrix}$$

or

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2.236 & 7.300 & 1.414 & 0 \\ 0 & 1.414 & 7.300 & 2.236 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} P_1' \\ P_2' \\ P_3' \\ P_4' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 6.641 & 0.949 \\ 6.641 & 0.949 \\ 1 & 1 \end{bmatrix}$$

Inverting and premultiplying yields the tangent vectors

$$\begin{bmatrix} P_1' \\ P_2' \\ P_3' \\ P_4' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -0.318 & 0.142 & -0.028 & 0.062 \\ 0.062 & -0.028 & 0.142 & -0.318 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 6.641 & 0.949 \\ 6.641 & 0.949 \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ 0.505 & -0.148 \\ 0.505 & -0.148 \\ 1 & 1 \end{bmatrix}$$

Recalling Eqs. (5–21)the blending functions for the first segment are

$$F_1(1/3) = 2(1/3)^3 - 3(1/3)^2 + 1 = \frac{20}{27} = 0.741$$

$$F_2(1/3) = -2(1/3)^3 + 3(1/3)^2 = \frac{7}{27} = 0.259$$

$$F_3(1/3) = (1/3)[(1/3)^2 - 2(1/3) + 1]\sqrt{2} = \frac{4\sqrt{2}}{27} = 0.210$$

$$F_4(1/3) = (1/3)[(1/3)^2 - 1/3]\sqrt{2} = -\frac{2\sqrt{2}}{27} = -0.105$$

and at $\tau = 2/3$

$$F_1(2/3) = 2(2/3)^3 - 3(2/3)^2 + 1 = \frac{7}{27} = 0.259$$

$$F_2(2/3) = -2(2/3)^3 + 3(2/3)^2 = \frac{20}{27} = 0.741$$

$$F_3(2/3) = (2/3)[(2/3)^2 - 2(2/3) + 1]\sqrt{2} = \frac{2\sqrt{2}}{27} = 0.105$$

$$F_4(2/3) = (2/3)[(2/3)^2 - 2/3]\sqrt{2} = \frac{-4\sqrt{2}}{27} = -0.210$$

The point on the first spline segment at $\tau = 1/3$ is given by Eq. (5–22), $P(\tau) = [\,F\,][\,G\,]$, where here

$$P(1/3) = \begin{bmatrix} 0.741 & 0.259 & 0.210 & -0.105 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 1 \\ 0.505 & -0.148 \end{bmatrix}$$

$$= \begin{bmatrix} 0.416 & 0.484 \end{bmatrix}$$

and at $\tau = 2/3$ by

$$P(2/3) = \begin{bmatrix} 0.259 & 0.741 & 0.105 & -0.210 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 1 \\ 0.505 & -0.148 \end{bmatrix}$$

$$= \begin{bmatrix} 0.740 & 0.876 \end{bmatrix}$$

Table 5–1 Results for Cubic Spline Curve

Segment	τ	$P_x(\tau)$	$P_y(\tau)$
1	1/3	0.416	0.484
	2/3	0.740	0.876
2	1/3	1.343	0.457
	2/3	1.657	−0.457
3	1/3	2.260	−0.876
	2/3	2.584	−0.484

Complete results are shown in Table 5–1. The cubic spline curve is shown in Fig. 5–9.

Although the blending functions (see Eq. 5–21) indicate that the end tangent vectors have less influence than the end position vectors on the shape of the spline segment, the effect can still be significant. Figure 5–10 shows a single plane symmetric spline segment with constant tangent vector directions and varying magnitudes. The tangent vector directions are indicated by the angle α, and the relative magnitudes by the length of the vectors. When the magnitude is a small fraction of the chord length l, the curve is convex at the ends and lies

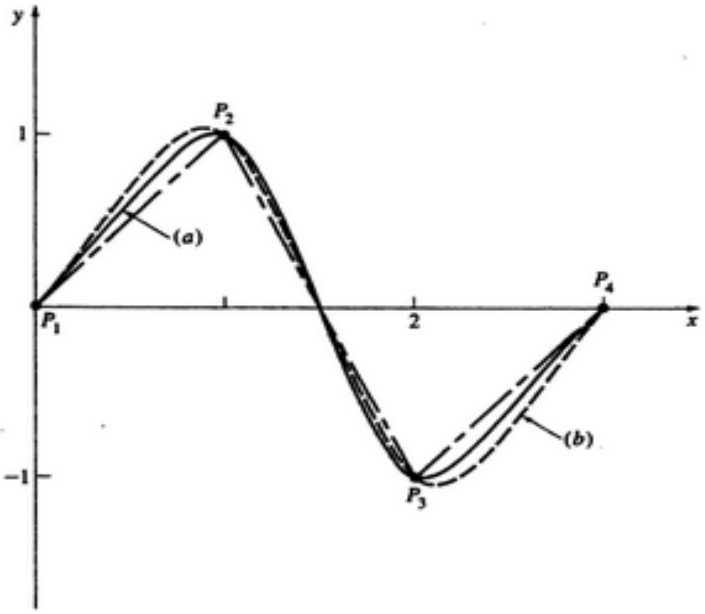


Figure 5–9 Piecewise cubic spline. (a) t_k determined using the chord approximation; (b) t_k normalized to 1.

sometimes used to improve the shape of the curve, is to restrict the tangent vector magnitude to less than or equal to the chord length.

5–4 NORMALIZED CUBIC SPLINES

An alternate approximation for the t_k spline segment parameter values to that previously suggested is to normalize them to unity. Thus, $0 \leq t \leq 1$ for all segments.

The blending functions (see Eq. 5–21) now become†

$$F_1(t) = 2t^3 - 3t^2 + 1 \qquad (5-25a)$$

$$F_2(t) = -2t^3 + 3t^2 \qquad (5-25b)$$

$$F_3(t) = t^3 - 2t^2 + t \qquad (5-25c)$$

$$F_4(t) = t^3 - t^2 \qquad (5-25d)$$

For the normalized cubic spline the blending function matrix is now written as

$$[F] = [T][N] = [t^3 \quad t^2 \quad t \quad 1] \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad (5-26)$$

The matrix equation for a cubic spline segment (see Eq. 5–22) can now be written as

$$P(t) = [F][G] = [T][N][G] \qquad (5-27)$$

Notice that $[T]$ and $[N]$ are constant for all cubic spline segments. Only the geometry matrix $[G]$ changes from segment to segment.

Equation (5–15), used to determine the internal tangent vectors required in $[G]$, now becomes

$$\begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & \cdot \\ 1 & 4 & 1 & 0 & \cdot & \cdot \\ 0 & 1 & 4 & 1 & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0 & 1 & 4 & 1 \\ \cdot & \cdot & \cdot & \cdot & 0 & 1 \end{bmatrix} \begin{bmatrix} P_1' \\ P_2' \\ \cdot \\ \cdot \\ \cdot \\ P_n' \end{bmatrix} = \begin{bmatrix} 3\{(P_3 - P_2) + (P_2 - P_1)\} \\ 3\{(P_4 - P_3) + (P_3 - P_2)\} \\ \cdot \\ \cdot \\ \cdot \\ 3\{(P_n - P_{n-1}) + (P_{n-1} - P_{n-2})\} \end{bmatrix} \qquad (5-28)$$

†These blending functions are the cubic Hermite polynomial blending functions on the interval $0 \leq t \leq 1$.

Example 5–3 Normalized Cubic Spline Curve

Again consider the four two-dimensional position vectors of Ex. 5–2, i.e., $P_1[0\ \ 0], P_2[1\ \ 1], P_3[2\ \ -1]$ and $P_4[3\ \ 0]$ with tangent vectors $P_1'[1\ \ 1]$ and $P_4'[1\ \ 1]$. Determine the normalized piecewise cubic spline curve through them.

Here the t_k parameter values are $t_2 = t_3 = t_4 = 1.0$.

The internal tangent vectors are obtained using Eq. (5–28), i.e.,

$$\begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & \cdot \\ 1 & 4 & 1 & 0 & \cdot & \cdot \\ 0 & 1 & 4 & 1 & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0 & 1 & 4 & 1 \\ \cdot & \cdot & \cdot & \cdot & 0 & 1 \end{bmatrix} \begin{bmatrix} P_1' \\ P_2' \\ \cdot \\ \cdot \\ \cdot \\ P_n' \end{bmatrix} = \begin{bmatrix} 3\{(P_3 - P_2) + (P_2 - P_1)\} \\ 3\{(P_4 - P_3) + (P_3 - P_2)\} \\ \cdot \\ \cdot \\ \cdot \\ 3\{(P_n - P_{n-1}) + (P_{n-1} - P_{n-2})\} \end{bmatrix} \qquad (5-28)$$

$$[F] = [T][N] = [t^3 \quad t^2 \quad t \quad 1] \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad (5-26)$$

The matrix equation for a cubic spline segment (see Eq. 5–22) can now be written as

$$P(t) = [F][G] = [T][N][G] \qquad (5-27)$$

The solution is again given by Eq. (5–16). However, here $[M]$ is constant and need only be inverted once. When the number of position vectors is large, this represents a considerable savings in computational expense.

An example illustrates these results.

Example 5–3 Normalized Cubic Spline Curve

Again consider the four two-dimensional position vectors of Ex. 5–2, i.e., $P_1[0\ \ 0]$, $P_2[1\ \ 1]$, $P_3[2\ \ -1]$ and $P_4[3\ \ 0]$ with tangent vectors $P_1'[1\ \ 1]$ and $P_4'[1\ \ 1]$. Determine the normalized piecewise cubic spline curve through them.

Here the t_k parameter values are $t_2 = t_3 = t_4 = 1.0$.

The internal tangent vectors are obtained using Eq. (5–28), i.e.,

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 4 & 1 & 0 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} P_1' \\ P_2' \\ P_3' \\ P_4' \end{bmatrix} = \begin{bmatrix} [1\ \ 1] \\ 3\{[2\ \ -1]-[1\ \ 1]+[1\ \ 1]-[0\ \ 0]\} \\ 3\{[3\ \ 0]-[2\ \ -1]+[2\ \ -1]-[1\ \ 1]\} \\ [1\ \ 1] \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ 6 & -3 \\ 6 & -3 \\ 1 & 1 \end{bmatrix}$$

Inverting and premultiplying yields

$$\begin{bmatrix} P_1' \\ P_2' \\ P_3' \\ P_4' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -0.267 & 0.267 & -0.067 & 0.067 \\ 0.067 & -0.067 & 0.267 & -0.267 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 6 & -3 \\ 6 & -3 \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ 1 & -0.8 \\ 1 & -0.8 \\ 1 & 1 \end{bmatrix}$$

Notice that here the internal tangent vectors P_2', P_3' are considerably different from those determined with the chord approximation used in Ex. 5–2.

Using Eq. (5–26) the blending function matrix for the first segment at $t = 1/3$ is

$$[F] = [T][N] = [1/27 \quad 1/9 \quad 1/3 \quad 1] \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$= [20/27 \quad 7/27 \quad 4/27 \quad -2/27]$$

and at $t = 2/3$,

$$[F] = [\,7/27 \quad 20/27 \quad 2/27 \quad -4/27\,]$$

The point on the first spline segment at $t = 1/3$ is

$$P(t) = [F][G] = [\,20/27 \quad 7/27 \quad 4/27 \quad -2/27\,] \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & -0.8 \end{bmatrix}$$

$$= [\,1/3 \quad 63/135\,] = [\,0.333 \quad 0.467\,]$$

and at $t = 2/3$,

$$P(t) = [F][G] = [\,7/27 \quad 20/27 \quad 2/27 \quad -4/27\,] \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & -0.8 \end{bmatrix}$$

$$= [\,2/3 \quad 126/135\,]$$

Complete results are shown in Table 5–2.

Table 5–2 Results for Normalized Cubic Spline Curve

Segment	t	$P_x(t)$	$P_y(t)$
1	1/3	0.333	0.467
	2/3	0.667	0.933
2	1/3	1.333	0.422
	2/3	1.667	−0.422
3	1/3	2.333	−0.933
	2/3	2.667	−0.467

A Bezier curve is determined by a defining polygon as shown in the figure.

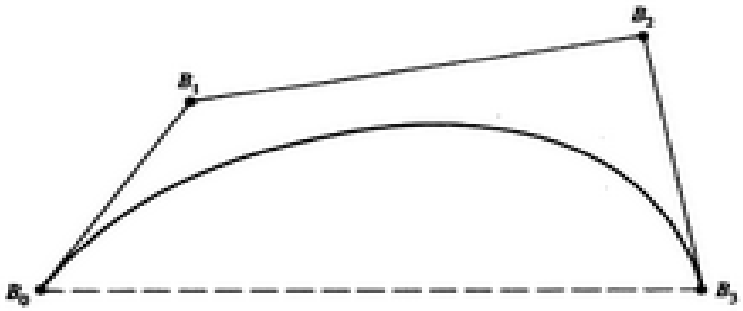


Figure 5–25 A Bézier curve and its defining polygon.

Properties of Bezier Cureve

The basis functions are real.

The degree of the polynomial defining the curve segment is one less than the number of defining polygon points.

The curve generally follows the shape of the defining polygon.

The first and last points on the curve are coincident with the first and last points of the defining polygon.

The tangent vectors at the ends of the curve have the same direction as the first and last polygon spans, respectively.

The curve is contained within the convex hull of the defining polygon, i.e., within the largest convex polygon defined by the polygon vertices. In Fig. 5–25, the convex hull is shown by the polygon and the dashed line.

The curve exhibits the variation diminishing property. Basically this means that the curve does not oscillate about any straight line more often than the defining polygon.

The curve is invariant under an affine transformation.

Several four-point Bézier polygons and the resulting cubic curves are shown in Fig. 5–26. With just the information given above a user quickly learns to predict the shape of a curve generated by a Bézier polygon.

Mathematically a parametric Bézier curve is defined by

$$P(t) = \sum_{i=0}^{n} B_i \, J_{n,i}(t) \qquad 0 \le t \le 1 \qquad (5-62)$$

where the Bézier or Bernstein basis or blending function is

$$J_{n,i}(t) = \binom{n}{i} t^i (1-t)^{n-i} \qquad (5-63)$$

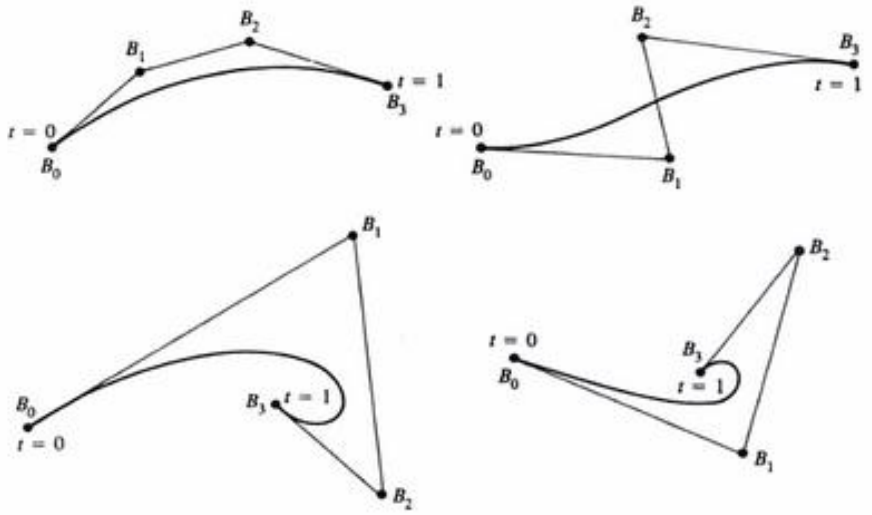


Figure 5–26 Bézier polygons for cubics.

$$\binom{n}{i} = \frac{n!}{i!\,(n-i)!} \tag{5-64}$$

$J_{n,i}(t)$ is the ith nth-order Bernstein basis function. Here n, the degree of the defining Bernstein basis function and thus of the polynomial curve segment, is one less than the number of points in the defining Bézier polygon. The vertices of the Bézier polygon are numbered from 0 to n as shown in Fig. 5–25.† Also $(0)^0 \equiv 1$ and $0! \equiv 1$.

Figure 5–27 shows the blending functions for several values of n. Notice the symmetry of the functions. Each of the blending functions is of degree n. For

†This notation and polygon numbering scheme is chosen to be consistent with the vast body of existing literature on Bézier curves and Bernstein basis functions. For programming purposes an alternate formulation is somewhat more convenient, i.e.,

$$P(t) = \sum_{I=1}^{N} B_I J_{N,I}(t)$$

where

$$J_{N,I} = \binom{N-1}{I-1} t^{I-1} (1-t)^{N-I}$$

and

$$\binom{N-1}{I-1} = \frac{(N-1)!}{(I-1)!\,(N-I)!}$$

The polygon points are numbered from 1 to N. The transformations $n = N-1$ and $i = I-1$ convert between the two notations.

example, each of the four blending functions shown in Fig. 5–27b for $n = 3$ is a cubic. The maximum value of each blending function occurs at $t = i/n$ and is given in Ref. 5–14 by

$$J_{n,i}\left(\frac{i}{n}\right) = \binom{n}{i}\frac{i^i(n-i)^{n-i}}{n^n} \qquad (5-65)$$

For example, for a cubic $n = 3$. The maximum values for $J_{3,1}$ and $J_{3,2}$ occur at 1/3 and 2/3, respectively, with values

$$J_{3,1}\left(\frac{1}{3}\right) = \frac{4}{9} \quad \text{and} \quad J_{3,1}\left(\frac{2}{3}\right) = \frac{4}{9}$$

Figure 5–27b illustrates this result.

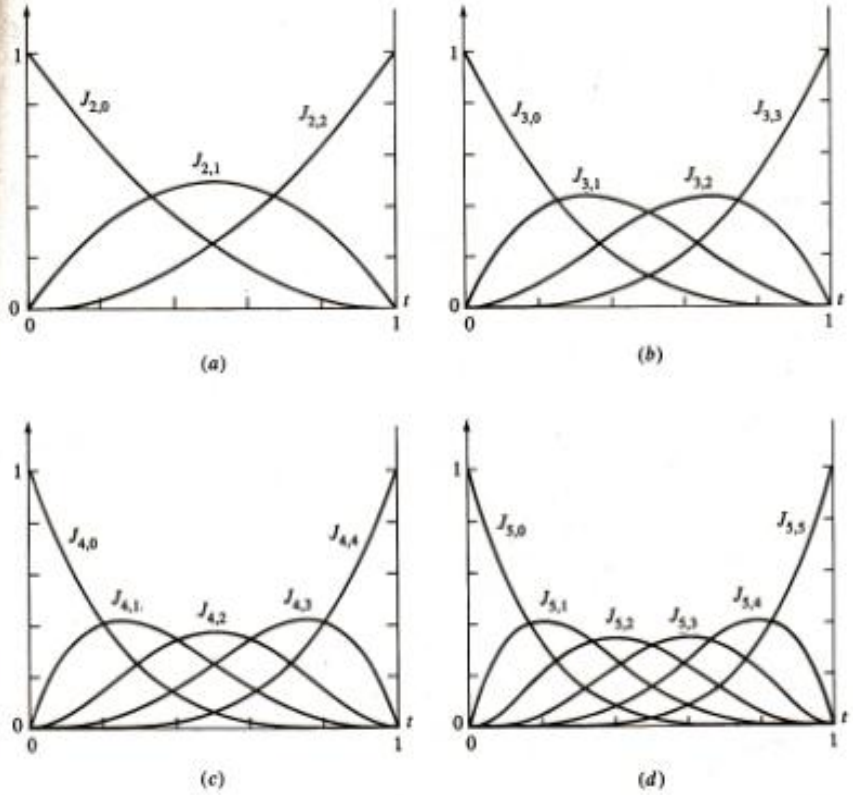


Figure 5–27 Bézier/Bernstein blending functions. (a) Three polygon points, $n = 2$; (b) four polygon points, $n = 3$; (c) five polygon points, $n = 4$; (d) six polygon points, $n = 5$.

Examining Eqs. (5–62) to (5–64) for the first point on the curve, i.e., at $t = 0$, shows that

$$J_{n,0}(0) = \frac{n!\,(1)(1-0)^{n-0}}{n!} = 1 \qquad i = 0$$

and

$$J_{n,i}(0) = \frac{n!\,(0)^i(1-0)^{n-i}}{i!\,(n-i)!} = 0 \qquad i \neq 0$$

Thus,

$$P(0) = B_0 J_{n,0}(0) = B_0$$

and the first point on the Bézier curve and on its defining polygon are coincident as previously claimed.

Similarly, for the last point on the curve, i.e., at $t = 1$,

$$J_{n,n}(1) = \frac{n!\,(1)^n(0)^{n-n}}{n!\,(1)} = 1 \qquad i = n$$

$$J_{n,i}(1) = \frac{n!}{i!\,(n-i)!}t^i(1-1)^{n-i} = 0 \qquad i \neq n$$

Thus,

$$P(1) = B_n J_{n,n}(1) = B_n$$

and the last point on the Bézier curve and the last point on its defining polygon are coincident. The blending functions shown in Fig. 5–27 illustrate these results.

Further, it can be shown that for any given value of the parameter t, the summation of the basis functions is precisely one; i.e.,

$$\sum_{i=0}^{n} J_{n,i}(t) = 1 \qquad (5-66)$$

An example illustrates the technique for determining a Bézier curve.

Example 5–7 Bézier Curve

Given $B_0\,[1 \quad 1]$, $B_1\,[2 \quad 3]$, $B_2\,[4 \quad 3]$ and $B_3\,[3 \quad 1]$ the vertices of a Bézier polygon (see Fig. 5–28), determine seven points on the Bézier curve.

Recall Eqs. (5–62) to (5–64):

$$P(t) = \sum_{i=0}^{n} B_i J_{n,i}(t)$$

where

$$J_{n,i}(t) = \binom{n}{i} t^i (1-t)^{n-i}$$

and

$$\binom{n}{i} = \frac{n!}{i!\,(n-i)!}$$

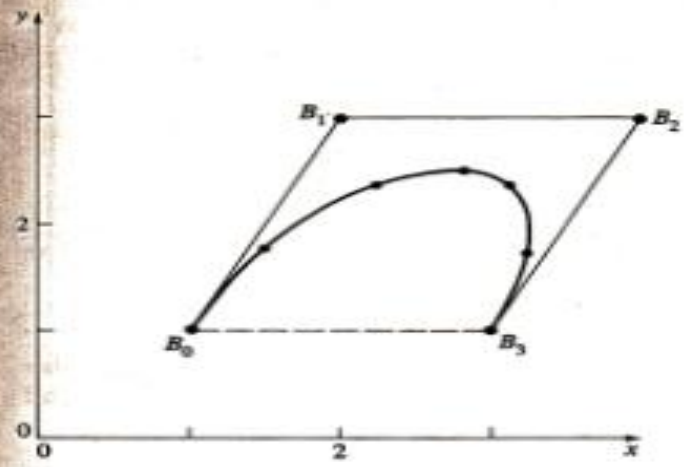


Figure 5–28 Results for Bézier curve segment for Ex. 5–7.

Here $n = 3$, since there are four vertices. Hence

$$\binom{n}{i} = \binom{3}{i} = \frac{6}{i!\,(3-i)!}$$

and

$$J_{3,0}(t) = (1)t^0(1-t)^3 = (1-t)^3$$
$$J_{3,1}(t) = 3t(1-t)^2$$
$$J_{3,2}(t) = 3t^2(1-t)$$
$$J_{3,3}(t) = t^3$$

Thus,

$$P(t) = B_0 J_{3,0} + B_1 J_{3,1} + B_2 J_{3,2} + B_3 J_{3,3}$$
$$= (1-t)^3 P_0 + 3t(1-t)^2 P_1 + 3t^2(1-t)P_2 + t^3 P_3$$

A table of $J_{n,i}$ for various values of t is given below.
The points on the curve are then

$$P(0) = B_0 = [\,1 \quad 1\,]$$
$$P(0.15) = 0.614B_0 + 0.325B_1 + 0.058B_2 + 0.003B_3 = [\,1.5 \quad 1.765\,]$$
$$P(0.35) = 0.275B_0 + 0.444B_1 + 0.239B_2 + 0.042B_3 = [\,2.248 \quad 2.367\,]$$
$$P(0.5) = 0.125B_0 + 0.375B_1 + 0.375B_2 + 0.125B_3 = [\,2.75 \quad 2.5\,]$$
$$P(0.65) = 0.042B_0 + 0.239B_1 + 0.444B_2 + 0.275B_3 = [\,3.122 \quad 2.367\,]$$
$$P(0.85) = 0.003B_0 + 0.058B_1 + 0.325B_2 + 0.614B_3 = [\,3.248 \quad 1.765\,]$$
$$P(1) = B_3 = [\,3 \quad 1\,]$$

Table 5-4 Coefficients for a Bézier curve

t	$J_{3,0}$	$J_{3,1}$	$J_{3,2}$	$J_{3,3}$
0	1	0	0	0
0.15	0.614	0.325	0.058	0.003
0.35	0.275	0.444	0.239	0.042
0.5	0.125	0.375	0.375	0.125
0.65	0.042	0.239	0.444	0.275
0.85	0.003	0.058	0.325	0.614
1	0	0	0	1

These points are shown along with the defining polygon in Fig. 5–28.

The equation for a Bézier curve can be expressed in a matrix form similar to those for cubic splines and parabolically blended curves (see Eqs. 5–27 and 5–44), i.e., as

$$P(t) = [T][N][G] = [F][G] \tag{5-67}$$

Here $[F] = [J_{n,0} \quad J_{n,1} \quad \cdots \quad J_{n,n}]$ and $[G]^T = [B_0 \quad B_1 \quad \cdots \quad B_n]$.

The specific matrix forms for low values of n are of interest. For four defining polygon points ($n = 3$), the cubic Bézier curve is given by

$$P(t) = [(1-t)^3 \quad 3t(1-t)^2 \quad 3t^2(1-t) \quad t^3] \begin{bmatrix} B_o \\ B_1 \\ B_2 \\ B_3 \end{bmatrix}$$

Collecting the coefficients of the parameter terms allows rewriting this as

$$P(t) = [T][N][G] = [t^3 \quad t^2 \quad t \quad 1] \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} B_o \\ B_1 \\ B_2 \\ B_3 \end{bmatrix} \tag{5-68}$$

Similarly, the quartic ($n = 4$) Bézier curve corresponding to five Bézier polygon points is

$$P(t) = [t^4 \quad t^3 \quad t^2 \quad t \quad 1] \begin{bmatrix} 1 & -4 & 6 & -4 & 1 \\ -4 & 12 & -12 & 4 & 0 \\ 6 & -12 & 6 & 0 & 0 \\ -4 & 4 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} B_o \\ B_1 \\ B_2 \\ B_3 \\ B_4 \end{bmatrix} \tag{5-69}$$

Cohen and Riesenfeld (Ref. 5–19) have generalized this representation to

$$P(t) = [T][N][G]$$

where here

$$[T] = [t^n \quad t^{n-1} \quad \cdots \quad t \quad 1]$$

$$[N] = \begin{bmatrix} \binom{n}{0}\binom{n}{n}(-1)^n & \binom{n}{1}\binom{n-1}{n-1}(-1)^{n-1} & \cdots & \binom{n}{n}\binom{n-n}{n-n}(-1)^0 \\ \binom{n}{0}\binom{n}{n-1}(-1)^{n-1} & \binom{n}{1}\binom{n-1}{n-2}(-1)^{n-2} & \cdots & 0 \\ \binom{n}{0}\binom{n}{1}(-1)^1 & \binom{n}{1}\binom{n-1}{0}(-1)^0 & \cdots & 0 \\ \binom{n}{0}\binom{n}{0}(-1)^0 & 0 & \cdots & 0 \end{bmatrix} \tag{5-70}$$

$[G]^T$ is again $[B_0 \quad B_1 \quad \cdots \quad B_n]$. The individual terms in $[N]$ are given by

$$(N_{i+1,j+1})_{i,j=0}^{n} = \begin{cases} \binom{n}{j}\binom{n-j}{n-i-j}(-1)^{n-i-j} & 0 \le i+j \le n \\ 0 & \text{otherwise} \end{cases}$$

Equation (5–70) can be decomposed into a sometimes more convenient form,

$$[N] = [C][D] \tag{5-71}$$

where

$$[C] = \begin{bmatrix} \binom{n}{n}(-1)^n & \binom{n}{1}\binom{n-1}{n-1}(-1)^{n-1} & \cdots & \binom{n}{n}\binom{n-n}{n-n}(-1)^0 \\ \binom{n}{n-1}(-1)^{n-1} & \binom{n}{1}\binom{n-1}{n-2}(-1)^{n-2} & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \binom{n}{1}(-1)^1 & \binom{n}{1}\binom{n-1}{0}(-1)^0 & & 0 \\ \binom{n}{0}(-1)^0 & 0 & \cdots & 0 \end{bmatrix}$$

$$[D] = \begin{bmatrix} \binom{n}{0} & & \cdots & 0 \\ \vdots & \binom{n}{1} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & \binom{n}{n} \end{bmatrix}$$

Equation (5–70) or (5–71) is more convenient to evaluate for arbitrary values of n. Notice that for each value of n the matrix $[\,N\,]$ is symmetrical about the main diagonal and that the lower right triangular corner is all zeros.

Although it is not necessary to numerically specify the tangent vectors at the ends of an individual Bézier curve, maintaining slope and curvature continuity when joining Bézier curves, determining surface normals for lighting or numerical control tool path calculation, or local curvature for smoothness or fairness calculations requires a knowledge of both the first and second derivatives of a Bézier curve.

Recalling Eq. (5–62), the first derivative of a Bézier curve is

$$P'(t) = \sum_{i=0}^{n} B_i J'_{n,i}(t) \qquad (5-72)$$

The second derivative is given by

$$P''(t) = \sum_{i=0}^{n} B_i J''_{n,i}(t) \qquad (5-73)$$

The derivatives of the basis function are obtained by formally differentiating Eq. (5–63). Specifically,

$$\begin{aligned} J'_{n,i}(t) &= \binom{n}{i}\{i\,t^{i-1}(1-t)^{n-i} - (n-i)t^i(1-t)^{n-i-1}\} \\ &= \binom{n}{i}t^i(1-t)^{n-i}\left\{\frac{i}{t} - \frac{(n-i)}{(1-t)}\right\} \\ &= \frac{(i-nt)}{t(1-t)}J_{n,i}(t) \end{aligned} \qquad (5-74)$$

Similarly the second derivative is

$$J''_{n,i}(t) = \left\{\frac{(i-nt)^2 - nt^2 - i(1-2t)}{t^2(1-t)^2}\right\}J_{n,i}(t) \qquad (5-75)$$

At the beginning and the ends of a Bézier curve, i.e., at $t = 0$ and $t = 1$, *numerical* evaluation of Eqs. (5–74) and (5–75) creates difficulties.†

An alternate evaluation for the rth derivative at $t = 0$ is given by

$$P^r(0) = \frac{n!}{(n-r)!}\sum_{i=0}^{r}(-1)^{r-i}\binom{r}{i}B_i \qquad (5-76)$$

---

†Algebraic evaluation of $J_{n,i}(t)$ and substitution *before* numerical evaluation yields correct results (see Ex. 5–8).

and at $t = 1$ by

$$P^r(1) = \frac{n!}{(n-r)!} \sum_{i=0}^{r} (-1)^i \binom{r}{i} B_{n-i} \qquad (5-77)$$

Thus, the first derivatives at the ends are

$$P'(0) = n(B_1 - B_0) \qquad (5-78)$$

and

$$P'(1) = n(B_n - B_{n-1}) \qquad (5-79)$$

This illustrates that the tangent vector for a Bézier curve at the initial and final points has the same direction as the initial and final polygon spans.

Similarly the second derivatives at the ends are

$$P''(0) = n(n-1)(B_0 - 2B_1 + B_2) \qquad (5-80a)$$

and

$$P''(1) = n(n-1)(B_n - 2B_{n-1} + B_{n-2}) \qquad (5-80b)$$

Thus, the second derivative of the Bézier curve at the initial and final points depends on the two nearest polygon spans, i.e., on the nearest three polygon vertices. In general, the rth derivative at an end point or starting point is determined by the end or starting point and its r neighboring polygon vertices. An example provides a more explicit illustration.

Example 5–8 Derivatives of Bézier Curves

Consider a four-point Bézier polygon as shown, for example, in Figs. 5–26 and 5–28. Recall that the Bézier curve is given by

$$P(t) = B_0 J_{3,0}(t) + B_1 J_{3,1}(t) + B_2 J_{3,2}(t) + B_3 J_{3,3}(t)$$

Hence the first derivative is

$$P(t) = B_0 J'_{3,0}(t) + B_1 J'_{3,1}(t) + B_2 J'_{3,2}(t) + B_3 J'_{3,3}(t)$$

Recalling Ex. 5–7 and differentiating the basis functions directly yields

$$\begin{aligned}
J_{3,0}(t) &= (1-t)^3 &\Rightarrow\quad J'_{3,0}(t) &= -3(1-t)^2 \\
J_{3,1}(t) &= 3t(1-t)^2 &\Rightarrow\quad J'_{3,1}(t) &= 3(1-t)^2 - 6t(1-t) \\
J_{3,2}(t) &= 3t^2(1-t) &\Rightarrow\quad J'_{3,2}(t) &= 6t(1-t) - 3t^2 \\
J_{3,3}(t) &= t^3 &\Rightarrow\quad J'_{3,3}(t) &= 3t^2
\end{aligned}$$

Evaluating these results at $t = 0$ yields

$$J'_{3,0}(0) = -3 \qquad J'_{3,1}(0) = 3 \qquad J'_{3,2}(0) = 0 \qquad J'_{3,3}(0) = 0$$

Substituting yields

$$P'(0) = -3P_0 + 3P_1 = 3(P_1 - P_0)$$

These results generalize to

$$C_i = \sum_{j=0}^{i} \binom{i}{j} \frac{B_j}{2^i} \qquad i = 0, 1, \cdots, n \tag{5-82a}$$

$$D_i = \sum_{j=i}^{n} \binom{n-i}{n-j} \frac{B_j}{2^{n-i}} \qquad i = 0, 1, \cdots, n \tag{5-82b}$$

Applied successively, the defining polygons converge to the Bézier curve itself.

5–9 B-SPLINE CURVES

From a mathematical point of view, a curve generated by using the vertices of a defining polygon is dependent on some interpolation or approximation scheme to establish the relationship between the curve and the polygon. This scheme is provided by the choice of basis function. As noted in Sec. 5–8, the Bernstein basis produces Bézier curves generated by Eq. (5–62). Two characteristics of the Bernstein basis, however, limit the flexibility of the resulting curves. First the number of specified polygon vertices fixes the order of the resulting polynomial which defines the curve. For example, a cubic curve must be defined by a polygon with four vertices and three spans. A polygon with six vertices always produces a fifth-degree curve. The only way to reduce the degree of the curve is to reduce the number of vertices, and conversely the only way to increase the degree of the curve is to increase the number of vertices.

The second limiting characteristic is due to the global nature of the Bernstein basis. This means that the value of the blending function $J_{n,i}(t)$ given by Eq. (5–63) is nonzero for all parameter values over the entire curve. Since any point on a Bézier curve is a result of blending the values of all defining vertices, a change in one vertex is felt throughout the entire curve. This eliminates the ability to produce a local change within a curve.

For example, since the end slopes of a Bézier curve are established by the directions of the first and last polygon spans, it is possible to change the middle vertex of a five-point polygon without changing the *direction* of the end slopes. However, the shape of the total curve is affected due to the global nature of the Bernstein basis. This lack of local control is detrimental in some applications.

There is another basis, called the B-spline basis (from Basis spline), which contains the Bernstein basis as a special case. This basis is generally nonglobal. The nonglobal behavior of B-spline curves is due to the fact that each vertex B_i is associated with a unique basis function. Thus, each vertex affects the shape of a curve only over a range of parameter values where its associated basis function is nonzero. The B-spline basis also allows the order of the basis function and hence the degree of the resulting curve to be changed without changing the number of defining polygon vertices. The theory for B-splines was first suggested by Schoenberg (Ref. 5–20). A recursive definition useful for numerical computation

was independently discovered by Cox (Ref. 5–21) and by de Boor (Ref. 5–22). Gordon and Riesenfeld (Refs. 5–15 and 5–23) applied the B-spline basis to curve definition.

Again letting $P(t)$ be the position vectors along the curve as a function of the parameter t, a B-spline curve is given by

$$P(t) = \sum_{i=1}^{n+1} B_i N_{i,k}(t) \qquad t_{\min} \le t < t_{\max}, \quad 2 \le k \le n+1 \tag{5-83}$$

where the B_i are the position vectors of the $n+1$ defining polygon vertices and the $N_{i,k}$ are the normalized B-spline basis functions.

For the ith normalized B-spline basis function of order k (degree $k-1$), the basis functions $N_{i,k}(t)$ are defined by the Cox-deBoor recursion formulas. Specifically,

$$N_{i,1}(t) = \begin{cases} 1 & \text{if } x_i \le t < x_{i+1} \\ 0 & \text{otherwise} \end{cases} \tag{5-84a}$$

and

$$N_{i,k}(t) = \frac{(t - x_i) N_{i,k-1}(t)}{x_{i+k-1} - x_i} + \frac{(x_{i+k} - t) N_{i+1,k-1}(t)}{x_{i+k} - x_{i+1}} \tag{5-84b}$$

The values of x_i are elements of a knot vector satisfying the relation $x_i \le x_{i+1}$. The parameter t varies from $t_{\min}$ to $t_{\max}$ along the curve $P(t)$.† The convention $0/0 = 0$ is adopted.

Formally a B-spline curve is defined as a polynomial spline function of order k (degree $k-1$) since it satisfies the following two conditions:

The function $P(t)$ is a polynomial of degree $k-1$ on each interval $x_i \le t < x_{i+1}$.

$P(t)$ and its derivatives of order $1, 2, \cdots, k-2$ are all continuous over the entire curve.

Thus, for example, a fourth-order B-spline curve is a piecewise cubic curve.

Because a B-spline basis is used to describe a B-spline curve, several properties in addition to those already mentioned above are immediately known:

The sum of the B-spline basis functions for any parameter value t can be shown (see Refs. 5–15 and 5–22) to be

$$\sum_{i=1}^{n+1} N_{i,k}(t) \equiv 1 \tag{5-85}$$

Each basis function is positive or zero for all parameter values, i.e., $N_{i,k} \ge 0$.

Except for $k = 1$ each basis function has precisely one maximum value.

The maximum order of the curve is equal to the number of defining polygon vertices.

The curve exhibits the variation diminishing property. Thus the curve does not oscillate about any straight line more often than its defining polygon.

The curve generally follows the shape of the defining polygon.

Any affine transformation can be applied to the curve by applying it to the defining polygon vertices; i.e., the curve is transformed by transforming the defining polygon vertices.

The curve lies within the convex hull of its defining polygon.

In fact, the convex hull property of B-spline curves is stronger than that for Bézier curves. For a B-spline curve of order k (degree $k-1$) a point on the curve lies within the convex hull of k neighboring points. Thus, all points on a B-spline curve must lie within the union of *all* such convex hulls formed by taking k successive defining polygon vertices. Figure 5–32, where the convex hulls are shown shaded, illustrates this effect for different values of k. Notice in particular that for $k = 2$ the convex hull is just the defining polygon itself. Hence, the B-spline curve is also just the defining polygon itself.

Using the convex hull property it is easy to see that if all the defining polygon vertices are colinear, then the resulting B-spline curve is a straight line for all orders k. Further, if l colinear polygon vertices occur in a noncolinear defining polygon, then the straight portions of the defining curve (if any) start and end at least $k-2$ spans from the beginning and end of the series of colinear polygon vertices. If the series of colinear polygon vertices is completely contained within a noncolinear defining polygon, then the number of colinear curve spans is at least $l-2k+3$. If the series of colinear polygon vertices occurs at the end of a noncolinear defining polygon, then the number of colinear curve spans is at least $l-k+1$. Figure 5–33 illustrates these results.

If at least $k-1$ coincident defining polygon vertices occur, i.e., $B_i = B_{i+1} \cdots = B_{i+k-2}$, then the convex hull of B_i to B_{i+k-2} is the vertex itself. Hence, the resulting B-spline curve must pass through the vertex B_i. Figure 5–34 illustrates this point for $k = 3$. Further, since a B-spline curve is everywhere C^{k-2} continuous, it is C^{k-2} continuous at B_i.

Finally, note that because of these continuity properties B-spline curves smoothly transition, with C^{k-2} continuity, into embedded straight segments as shown in Fig. 5–35.

Equations (5–84) clearly show that the choice of knot vector has a significant influence on the B-spline basis functions $N_{i,k}(t)$ and hence on the resulting B-spline curve. The only requirement for a knot vector is that it satisfy the relation $x_i \leq x_{i+1}$; i.e., it is a monotonically increasing series of real numbers. Fundamentally three types of knot vector are used: uniform, open uniform (or open) and nonuniform.

In a uniform knot vector, individual knot values are evenly spaced. Examples are

$$[0 \quad 1 \quad 2 \quad 3 \quad 4]$$

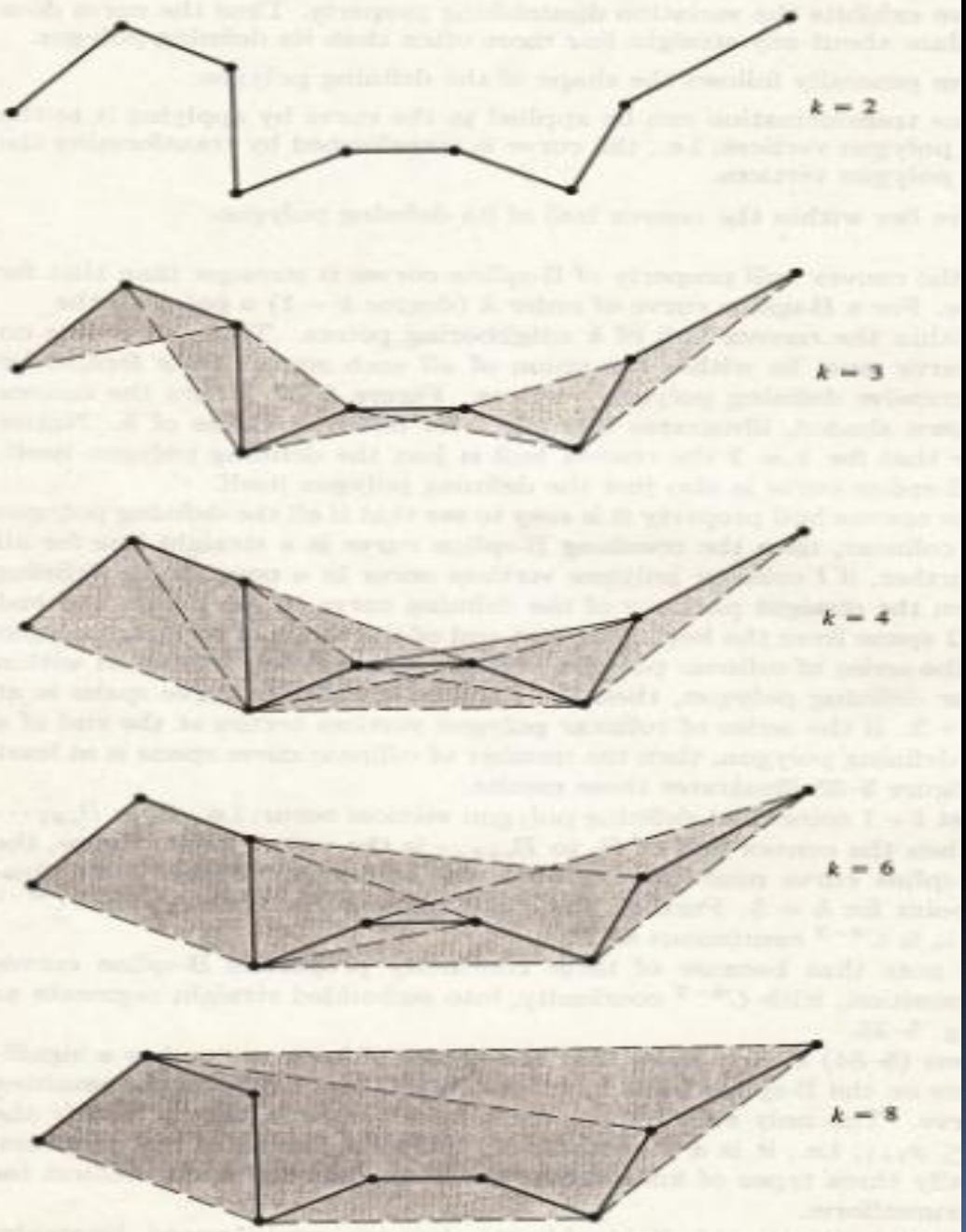


Figure 5–32 Convex hull properties of B-spline curves.

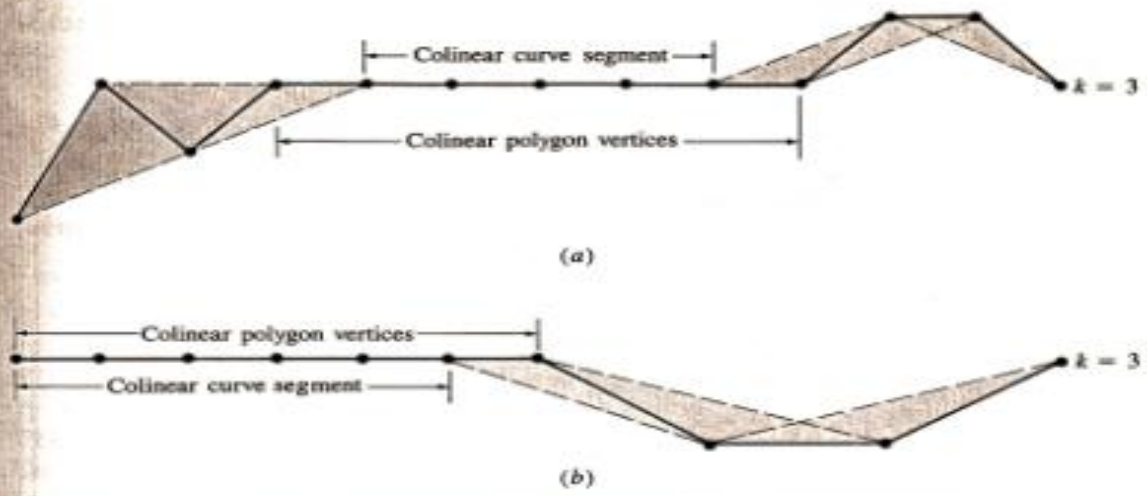


Figure 5–33 B-spline convex hull properties for colinear curve segments. (a) Within the defining polygon vertices; (b) at the end of the defining polygon vertices.

$$[\,-0.2 \quad -0.1 \quad 0 \quad 0.1 \quad 0.2\,]$$

In practice, uniform knot vectors generally begin at zero and are incremented by 1 to some maximum value or are normalized in the range between 0 and 1, i.e., equal decimal intervals, e.g.,

$$[\,0 \quad 0.25 \quad 0.5 \quad 0.75 \quad 1.0\,]$$

For a given order k, uniform knot vectors yield periodic uniform basis functions for which

$$N_{i,k}(t) = N_{i-1,k}(t-1) = N_{i+1,k}(t+1)$$

Thus each basis function is a translate of the other. Figure 5–36 illustrates this.

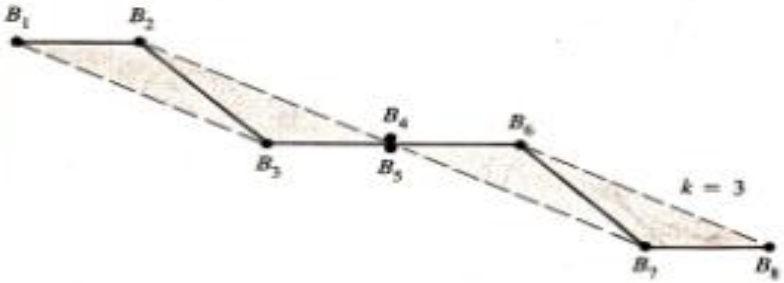


Figure 5–34 Convex hull for coincident polygon vertices, $k = 3$.

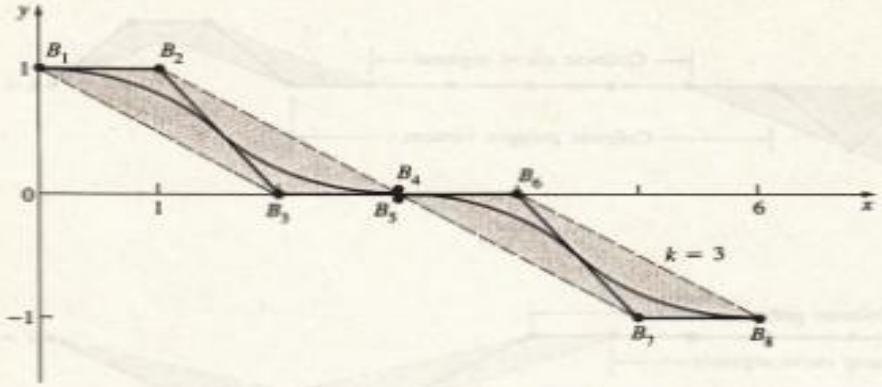


Figure 5–35 Smooth (C^{k-2}) transition into straight segments.

An open uniform knot vector has multiplicity of knot values at the ends equal to the order k of the B-spline basis function. Internal knot values are evenly spaced. Some examples using integer increments are

$k = 2$ [0 0 1 2 3 4 4]
$k = 3$ [0 0 0 1 2 3 3 3]
$k = 4$ [0 0 0 0 1 2 2 2 2]

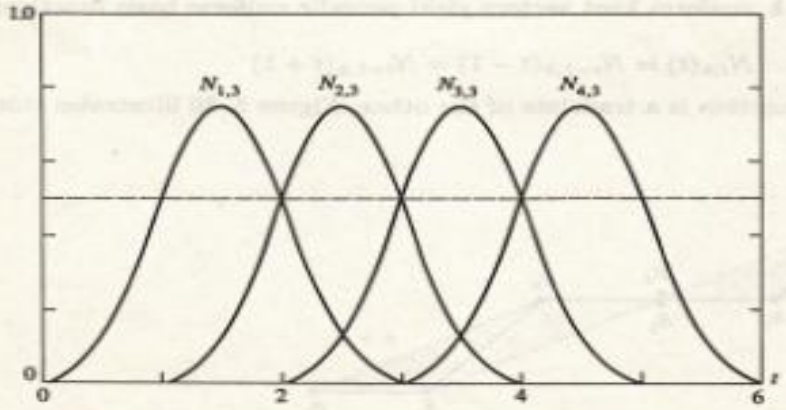


Figure 5–36 Periodic uniform B-spline basis functions, $[X] = [0\ 1\ 2\ 3\ 4\ 5\ 6]$, $n + 1 = 4$, $k = 3$.

A curve for which $P(t_0) = P(t_n)$ is called a closed (or periodic) curve, otherwise it is called open (or non-periodic)

or for normalized increments

$$
\begin{aligned}
k = 2 \quad & [0 \;\; 0 \;\; 1/4 \;\; 1/2 \;\; 3/4 \;\; 1 \;\; 1] \\
k = 3 \quad & [0 \;\; 0 \;\; 0 \;\; 1/3 \;\; 2/3 \;\; 1 \;\; 1 \;\; 1] \\
k = 4 \quad & [0 \;\; 0 \;\; 0 \;\; 0 \;\; 1/2 \;\; 1 \;\; 1 \;\; 1 \;\; 1]
\end{aligned}
$$

Formally, an open uniform knot vector is given by

$$
\begin{aligned}
x_i &= 0 && 1 \le i \le k \\
x_i &= i - k && k+1 \le i \le n+1 \\
x_i &= n - k + 2 && n+2 \le i \le n+k+1
\end{aligned}
$$

The resulting open uniform basis functions yield curves that behave most nearly like Bézier curves. In fact, when the number of defining polygon vertices is equal to the order of the B-spline basis and an open uniform knot vector is used, the B-spline basis reduces to the Bernstein basis. Hence, the resulting B-spline curve is a Bézier curve. In that case, the knot vector is just k zeros followed by k ones. For example, for four polygon vertices the fourth order ($k = 4$) open uniform knot vector is

$$[0 \;\; 0 \;\; 0 \;\; 0 \;\; 1 \;\; 1 \;\; 1 \;\; 1]$$

A cubic Bézier/B-spline curve results. The corresponding open uniform basis functions are shown in Fig. 5–27b. Additional open uniform basis functions are shown in Fig. 5–37.

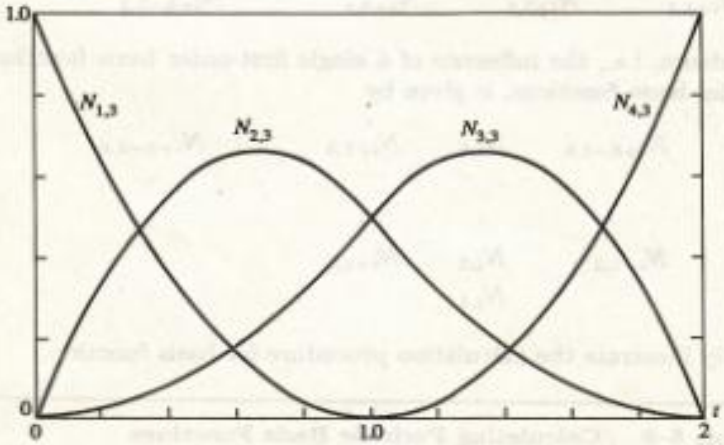


Figure 5–37 Open uniform B-spline basis functions, $[X] = [0 \;\; 0 \;\; 0 \;\; 1 \;\; 2 \;\; 2 \;\; 2]$, $k = 3$, $n + 1 = 4$.

Nonuniform knot vectors may have either unequally spaced and/or multiple internal knot values. They may be periodic or open. Examples are

$$[0 \quad 0 \quad 0 \quad 1 \quad 1 \quad 2 \quad 2 \quad 2]$$

$$[0 \quad 1 \quad 2 \quad 2 \quad 3 \quad 4]$$

$$[0 \quad 0.28 \quad 0.5 \quad 0.72 \quad 1]$$

Figures 5–38b to e show several nonuniform B-spline basis functions for order $k = 3$. The knot vectors used to generate the basis functions all have multiplicity of k equal values at the ends. Figure 5–38a gives the basis functions for an open uniform knot vector for comparison. Notice the symmetry of the basis functions in Figs. 5–38a and b and how that symmetry is lost for the nonuniform basis functions in Figs. 5–38c to e. Notice also that for multiple knot values within the knot vector a cusp occurs in one of the basis functions. Further, in Figs. 5–38d and e notice the shift of the location of the cusp corresponding to the change in location of the multiple knot value in the knot vector.

Because the Cox-deBoor formula (see Eq. 5–84) used to calculate B-spline basis functions is a recursion relation, a basis function of a given order k depends on lower order basis functions down to order 1. For a given basis function $N_{i,k}$ this dependence forms a triangular pattern given by

$$\begin{array}{lllll}
N_{i,k} & & & & \\
N_{i,k-1} & N_{i+1,k-1} & & & \\
N_{i,k-2} & N_{i+1,k-2} & N_{i+2,k-2} & & \\
\cdot & & & \cdot & \\
\cdot & & & \cdot & \cdot \\
N_{i,1} & N_{i+1,1} & N_{i+2,1} & N_{i+3,1} \cdot & N_{i+k-1,1}
\end{array}$$

The inverse dependence, i.e., the influence of a single first-order basis function $N_i, 1$ on higher order basis functions, is given by

$$\begin{array}{lllllll}
N_{i-k+1,k} & \cdot & N_{i+k-1,k} & N_{i,k} & N_{i+1,k} & \cdot & N_{i+k-1,k} \\
\cdot & \cdot & & & & \cdot & \cdot \\
 & \cdot & \cdot & \cdot & \cdot & \cdot & \\
 & & N_{i-1,2} & N_{i,2} & N_{i+1,2} & & \\
 & & & N_{i,1} & & &
\end{array}$$

Examples more fully illustrate the calculation procedure for basis function.

Example 5–9 Calculating Periodic Basis Functions

Calculate the four third-order ($k = 3$) basis functions $N_{i,3}(t)$, $i = 1, 2, 3, 4$. Here $n+1$, the number of basis functions, is 4. The basis function dependencies

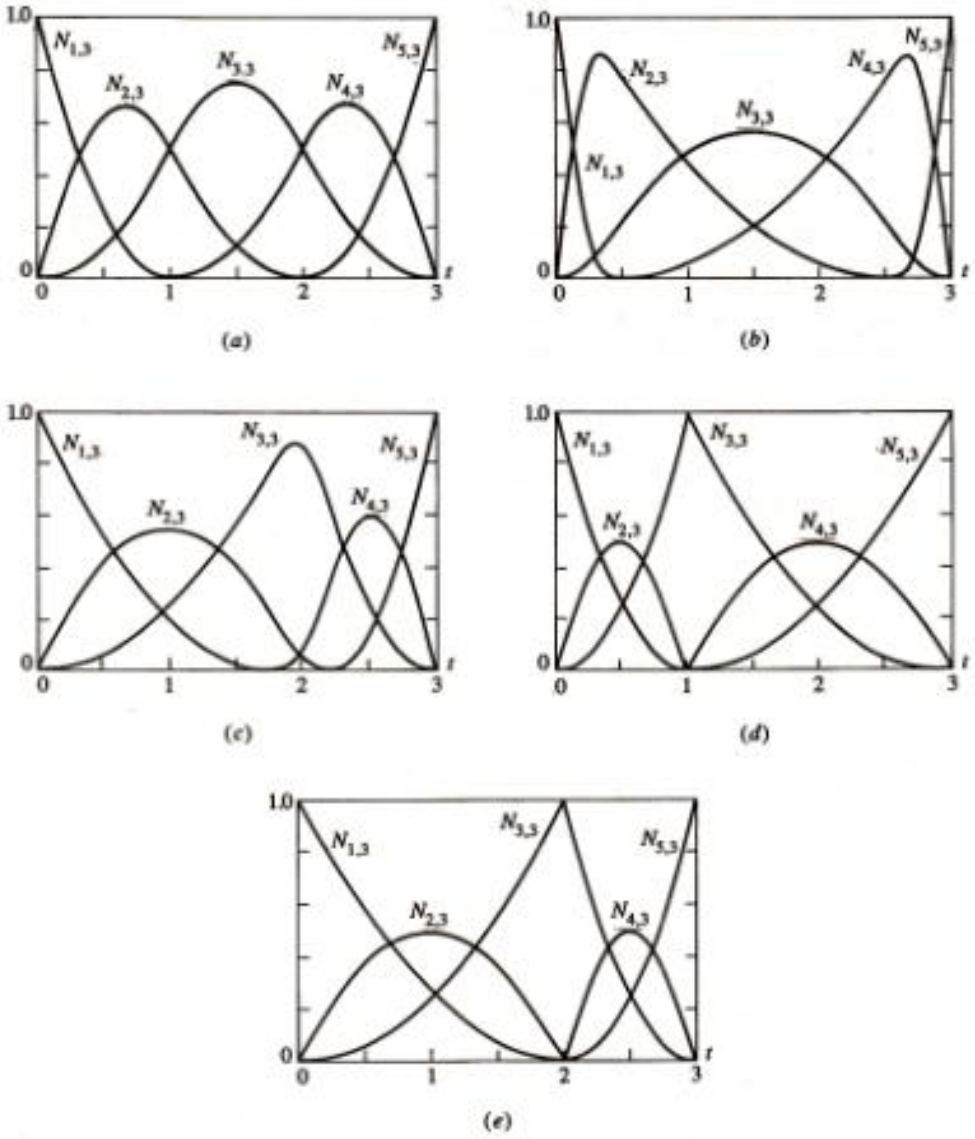


Figure 5–38 Nonuniform basis functions for $n + 1 = 5$, $k = 3$.
(a) $[X] = [0\ \ 0\ \ 0\ \ 1\ \ 2\ \ 3\ \ 3\ \ 3]$;
(b) $[X] = [0\ \ 0\ \ 0\ \ .4\ \ 2.6\ \ 3\ \ 3\ \ 3]$;
(c) $[X] = [0\ \ 0\ \ 0\ \ 1.8\ \ 2.2\ \ 3\ \ 3\ \ 3]$;
(d) $[X] = [0\ \ 0\ \ 0\ \ 1\ \ 1\ \ 3\ \ 3\ \ 3]$;
(e) $[X] = [0\ \ 0\ \ 0\ \ 2\ \ 2\ \ 3\ \ 3\ \ 3]$.

for $N_{i,3}$ are given by the following diagram:

$$\begin{array}{cccccc} N_{1,3} & N_{2,3} & N_{3,3} & N_{4,3} & & \\ N_{1,2} & N_{2,2} & N_{3,2} & N_{4,2} & N_{5,2} & \\ N_{1,1} & N_{2,1} & N_{3,1} & N_{4,1} & N_{5,1} & N_{6,1} \end{array}$$

The inverse dependencies for $i \geq 1$ are given by

$$\begin{array}{cccccc} N_{1,3} & N_{2,3} & N_{3,3} & N_{4,3} & N_{5,3} & N_{6,3} \\ N_{1,2} & N_{2,2} & N_{3,2} & N_{4,2} & N_{5,2} & \\ N_{1,1} & N_{2,1} & N_{3,1} & N_{4,1} & & \end{array}$$

Now, what is the knot vector range needed for this calculation? Equation (5–84) shows that the calculation of $N_{6,1}$ requires knot values x_6 and x_7, while calculation of $N_{1,1}$ requires x_1 and x_2. Thus, knot values from 0 to $n+k$ are required. The number of knot values is thus $n+k+1$. Hence the knot vector for these periodic basis functions is

$$[X] = [0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6]$$

where $x_1 = 0, \cdots, x_7 = 6$. The parameter range is $0 \leq t \leq 6$. Using Eq. (5–84) and the dependency diagram above, the basis functions for various parameter ranges are

$0 \leq t < 1$

$$N_{1,1}(t) = 1; \qquad N_{i,1}(t) = 0, \quad i \neq 1$$
$$N_{1,2}(t) = t; \qquad N_{i,2}(t) = 0, \quad i \neq 1$$
$$N_{1,3}(t) = t^2; \qquad N_{i,3}(t) = 0, \quad i \neq 1$$

$1 \leq t < 2$

$$N_{2,1}(t) = 1; \qquad N_{i,1}(t) = 0, \quad i \neq 2$$
$$N_{1,2}(t) = (1-t); \qquad N_{2,2}(t) = (t-1); \qquad N_{i,2}(t) = 0, \quad i \neq 1,2$$
$$N_{1,3}(t) = \frac{t}{2}(1-t) + \left(\frac{3-t}{2}\right)(t-1);$$
$$N_{2,3}(t) = \frac{(t-1)^2}{2}; \qquad N_{i,3}(t) = 0, \quad i \neq 1,2,3$$

$2 \leq t < 3$

$$N_{3,1}(t) = 1; \qquad N_{i,1}(t) = 0, \quad i \neq 3$$
$$N_{2,2}(t) = (3-t); \qquad N_{3,2}(t) = (t-2); \qquad N_{i,2}(t) = 0, \quad i \neq 2,3$$
$$N_{1,3}(t) = \frac{(3-t)^2}{2};$$
$$N_{2,3}(t) = \frac{(t-1)(3-t)}{2} + \frac{(4-t)(t-2)}{2};$$
$$N_{3,3}(t) = \frac{(t-2)^2}{2}; \qquad N_{i,3}(t) = 0, \quad i \neq 1,2,3$$

$3 \le t < 4$

$N_{4,1}(t) = 1; \quad N_{i,1}(t) = 0, \quad i \neq 4$

$N_{3,2}(t) = (4-t); \quad N_{4,2}(t) = (t-3); \quad N_{i,2}(t) = 0, \quad i \neq 3,4$

$N_{2,3}(t) = \frac{(4-t)^2}{2}; \quad N_{3,3}(t) = \frac{(t-2)(4-t)}{2} + \frac{(5-t)(t-3)}{2};$

$N_{4,3}(t) = \frac{(t-3)^2}{2}; \quad N_{i,3}(t) = 0, \quad i \neq 2,3,4$

$4 \le t < 5$

$N_{5,1}(t) = 1; \quad N_{i,1}(t) = 0, \quad i \neq 5$

$N_{4,2}(t) = (5-t); \quad N_{5,2}(t) = (t-4); \quad N_{i,2}(t) = 0, \quad i \neq 4,5$

$N_{3,3}(t) = \frac{(5-t)^2}{2};$

$N_{4,3}(t) = \frac{(t-3)(5-t)}{2} + \frac{(6-t)(t-4)}{2};$

$N_{i,3}(t) = 0, \quad i \neq 3,4$

$5 \le t < 6$

$N_{6,1}(t) = 1; \quad N_{i,1}(t) = 0, \quad i \neq 6$

$N_{5,2}(t) = (6-t); \quad N_{i,2}(t) = 0, \quad i \neq 5$

$N_{4,3}(t) = \frac{(6-t)^2}{2}; \quad N_{i,3}(t) = 0, \quad i \neq 4$

Note that because of the $<$ sign in the definition of $N_{i,1}$, all basis functions are precisely zero at $t = 6$.

These results are shown in Fig. 5–36 and Fig. 5–39c. Note that each one of the basis functions is a piecewise parabolic (quadratic) curve. Here, three piecewise parabolic segments on the intervals $x_i \rightarrow x_{i+1}$, $x_{i+1} \rightarrow x_{i+2}$, $x_{i+2} \rightarrow x_{i+3}$ are joined together to form each $N_{i,3}$ basis function. Further, note that each of the basis functions is simply a translate of the other.

Using the results of Ex. 5–9, the buildup of the higher order basis functions $N_{i,3}$ from lower order basis functions is easily illustrated. Figure 5–39a shows the first-order basis functions determined in Ex. 5–9, Fig. 5–39b shows the second-order basis functions and Fig. 5–39c repeats the third-order basis functions of Fig. 5–36 for completeness. Notice how the range of nonzero basis function values spreads with increasing order. The basis function is said to provide support on the interval x_i to x_{i+k}.

Examining Fig. 5–36 closely reveals an important property of uniform basis functions. Recalling from Eq. 5–85 that $\sum N_{i,k}(t) = 1$ at any parameter value t shows that a complete set of periodic basis functions for $k = 3$ is defined only in the range $2 \le t \le 4$. Outside of this range the $\sum N_{i,k}(t) \neq 1$. For a uniform knot vector beginning at 0 with integer spacings the usable parameter range is $k - 1 \le t \le (n + k) - (k - 1) = n + 1$. For more general or normalized knot vectors, the reduction in usable parameter range corresponds to the loss of $k - 1$ knot value intervals at each end of the knot vector.

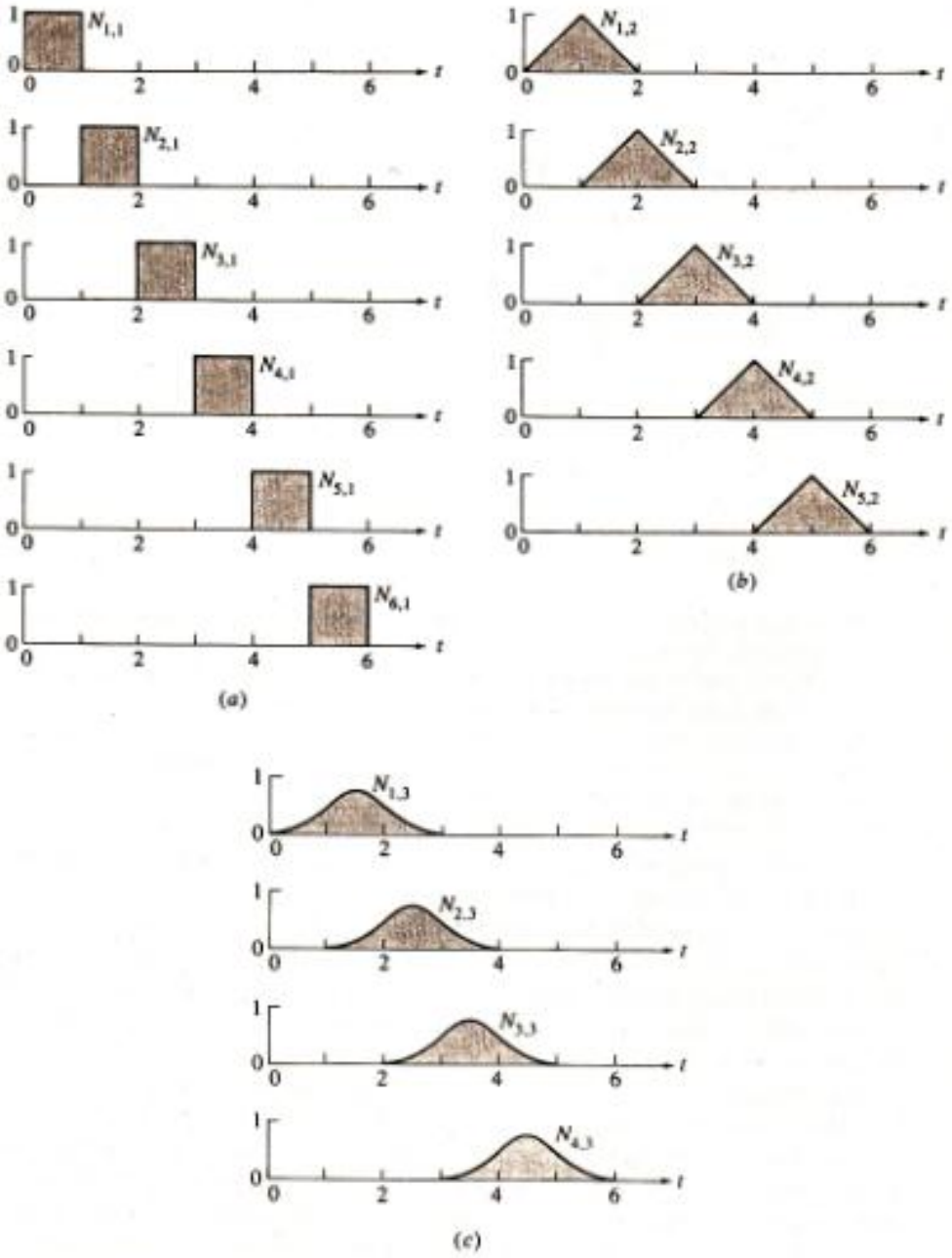


Figure 5–39 Periodic basis function buildup $n + 1 = 4$. (a) $k = 1$; (b) $k = 2$; (c) $k = 3$.

Example 5–10 Calculating Open Uniform Basis Functions

Calculate the four ($n = 3$) third-order ($k = 3$) basis functions $N_{i,3}(t)$, $i = 1, 2, 3, 4$ with an open knot vector.

Recalling that formally an open knot vector with integer intervals between internal knot values is given by

$$x_i = 0 \qquad 1 \le i \le k$$
$$x_i = i - k \qquad k+1 \le i \le n+1$$
$$x_i = n - k + 2 \qquad n+2 \le i \le n+k+1$$

The parameter range is $0 \le t \le n - k + 2$, i.e., from zero to the maximum knot value. Again, as in Ex. 5–9, the number of knot values is $n+k+1$. Using integer knot values, the knot vector for the current example is

$$[X] = [0 \quad 0 \quad 0 \quad 1 \quad 2 \quad 2 \quad 2]$$

where $x_1 = 0, \cdots, x_7 = 2$. The parameter range is from $0 \le t \le 2$.

Using Eqs. (5–84) and the dependency diagrams, the basis functions for various parameter ranges are

$0 \le t < 1$

$$N_{3,1}(t) = 1; \qquad N_{i,1}(t) = 0, \quad i \neq 3$$
$$N_{2,2}(t) = 1 - t; \qquad N_{3,2}(t) = t; \qquad N_{i,2}(t) = 0, \quad i \neq 2, 3$$
$$N_{1,3}(t) = (1-t)^2; \qquad N_{2,3}(t) = t(1-t) + \frac{(2-t)}{2}t;$$
$$N_{3,3}(t) = \frac{t^2}{2}; \qquad N_{i,3}(t) = 0, \quad i \neq 1, 2, 3$$

$1 \le t < 2$

$$N_{4,1}(t) = 1; \qquad N_{i,1}(t) = 0, \quad i \neq 4$$
$$N_{3,2}(t) = (2-t); \qquad N_{4,2}(t) = (t-1); \qquad N_{i,2}(t) = 0, \quad i \neq 3, 4$$
$$N_{2,3}(t) = \frac{(2-t)^2}{2}; \qquad N_{3,3}(t) = \frac{t(2-t)}{2} + (2-t)(t-1);$$
$$N_{4,3}(t) = (t-1)^2; \qquad N_{i,3}(t) = 0, \quad i \neq 2, 3, 4$$

These results are shown in Fig. 5–40.

Comparing the results from Ex. 5–10 shown in Fig. 5–40 with those from Ex. 5–9 shown in Fig. 5–39 illustrates that significantly different results are obtained when using periodic uniform or open uniform knot vectors. In particular, note that for open uniform knot vectors a complete set of basis functions is defined for the entire parameter range; i.e., $\sum N_{i,k}(t) = 1$ for all t, $0 \le t \le n - k + 2$. Notice also the reduction in parameter range compared to that for a periodic uniform knot vector.

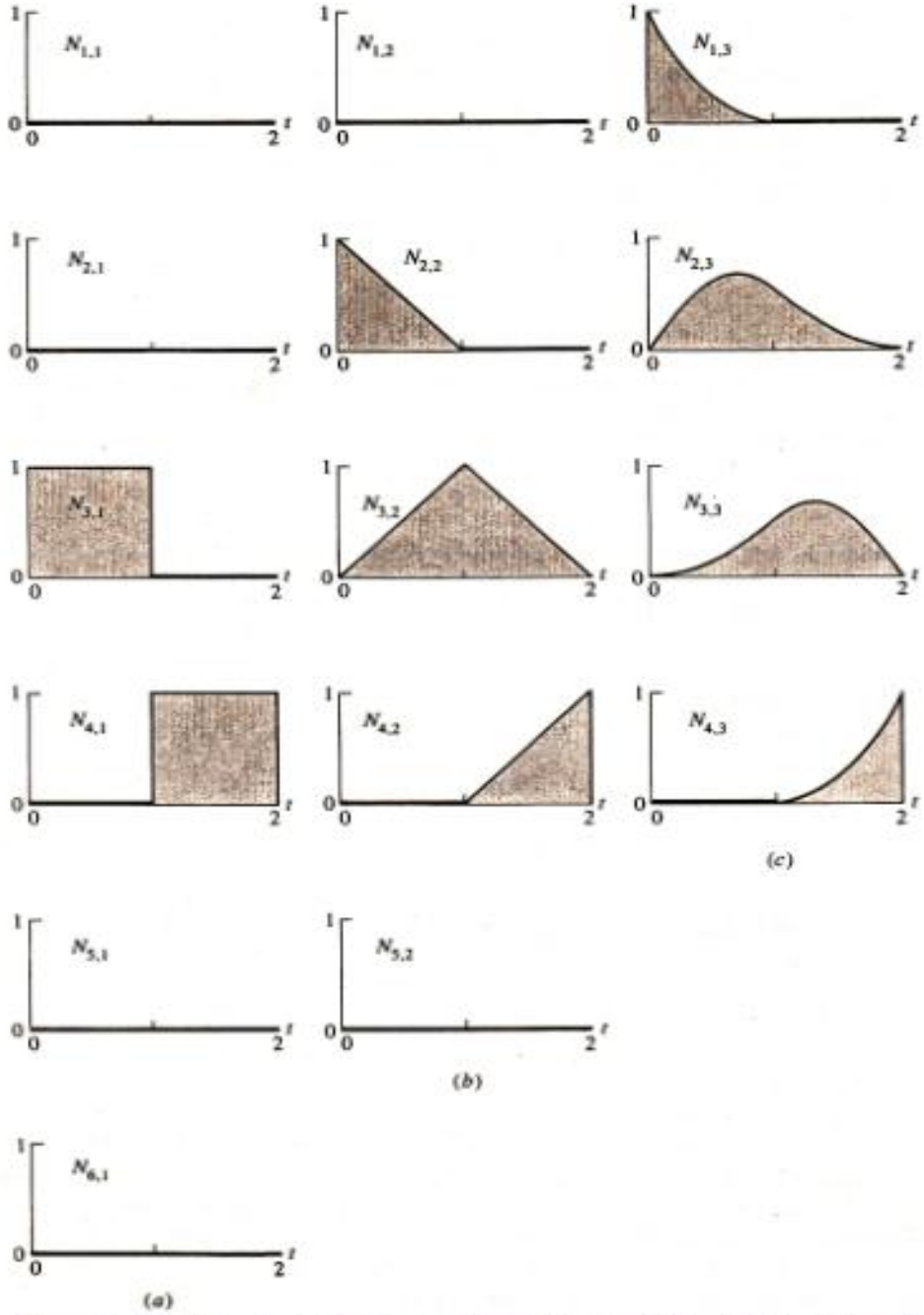


Figure 5–40 Open basis function buildup $n + 1 = 4$. (a) $k = 1$; (b) $k = 2$; (c) $k = 3$.

Example 5–11 Calculating Nonuniform Basis Functions

Calculate the five $(n+4)$ third-order $(k=3)$ basis functions $N_{i,3}(t)$, $i = 1,2,3,4,5$ using the knot vector $[X] = [0\ 0\ 0\ 1\ 1\ 3\ 3\ 3]$ which contains an interior repeated knot value. Using Eqs. (5–84) and the dependency diagrams, the basis functions are

$0 \le t < 1$

$$N_{3,1}(t) = 1; \qquad N_{i,1}(t) = 0, \quad i \neq 2$$
$$N_{2,2}(t) = 1 - t; \qquad N_{3,2}(t) = t; \qquad N_{i,2}(t) = 0, \quad i \neq 2,3$$
$$N_{1,3}(t) = (1-t)^2; \qquad N_{2,3}(t) = t(1-t) + (1-t)t = 2t(1-t);$$
$$N_{3,3}(t) = t^2; \qquad N_{i,3}(t) = 0, \quad i \neq 1,2,3$$

$1 \le t < 1$

$$N_{i,1}(t) = 0, \quad \text{all } i$$
$$N_{i,2}(t) = 0, \quad \text{all } i$$
$$N_{i,3}(t) = 0, \quad \text{all } i$$

Notice specifically that as a consequence of the multiple knot value, $N_{4,1}(t) = 0$ for all t.

$1 \le t < 3$

$$N_{5,1}(t) = 1; \qquad N_{i,1}(t) = 0, \quad i \neq 5$$
$$N_{4,2}(t) = \frac{(3-t)}{2}; \qquad N_{5,2}(t) = \frac{(t-1)}{2}; \qquad N_{i,2}(t) = 0, \quad i \neq 4,5$$
$$N_{3,3}(t) = \frac{(3-t)^2}{4};$$
$$N_{4,3}(t) = \frac{(t-1)(3-t)}{4} + \frac{(3-t)(t-1)}{4} = \frac{(3-t)(t-1)}{2};$$
$$N_{5,3}(t) = \frac{(t-1)^2}{4}; \qquad N_{i,3}(t) = 0, \quad i \neq 3,4,5$$

The results are shown in Fig. 5–38d.

Notice that for each value of t the $\sum N_{i,3}(t) = 1.0$. For example, with $0 \le t < 1$

$$\sum_{i=1}^{5} N_{i,3}(t) = (1-t)^2 + 2t(1-t) + t^2 = 1 - 2t + t^2 + 2t - 2t^2 + t^2 = 1$$

Similarly, for $1 \le t < 3$

$$\sum_{i=1}^{5} N_{i,3}(t) = \frac{1}{4}[(3-t)^2 + 2(3-t)(t-1) + (t-1)^2]$$
$$= \frac{1}{4}[9 - 6t + t^2 - 6 + 8t - 2t^2 + 1 - 2t + t^2]$$
$$= \frac{4}{4} = 1$$

The above discussion shows the significant influence of the choice of knot vector on the shape of the B-spline basis functions and hence on the shape of any resulting B-spline curve.

Because of the flexibility of B-spline basis functions and hence of the resulting B-spline curves, different types of control 'handles' are used to influence the shape of the curve. Control is achieved by:

Changing the type of knot vector and hence basis function: periodic uniform, open uniform or nonuniform.

Changing the order k of the basis function.

Changing the number and position of the defining polygon vertices.

Using multiple polygon vertices.

Using multiple knot values in the knot vector.

These effects are illustrated first with open B-spline curves and then with periodic B-spline curves and finally with nonuniform B-spline curves.

The behavior of an open B-spline curve is in many respects analogous to that of a Bézier curve. In fact, as has previously been mentioned, when the order of a B-spline curve is equal to the number of defining polygon vertices, the B-spline basis reduces to the Bernstein basis. Consequently, the resulting B-spline curve is identical to a Bézier curve. For an open B-spline curve of any order ($k \geq 2$) the first and last points on the curve are coincident with the first and last polygon vertices. Further, the slope of the B-spline curve at the first and last polygon vertices is equal to the slope of the first and last polygon spans.

Figure 5–41 shows three open B-spline curves of different order, each defined by the same four polygon vertices. The fourth-order curve corresponds to the Bézier curve. This curve is a single cubic polynomial segment. The third-order curve is composed of two parabolic curve segments joined at the center of the second span with C^1 continuity. The second-order curve reproduces the defining polygon. It consists of three linear 'curve' segments joined at the second and third polygon vertices with C^0 continuity. Notice that all three curves have the same end slopes, determined by the slope of the first and last spans of the defining polygon. Notice also that as the order of the curve increases, the resulting curve looks less like the defining polygon. Thus, increasing the order 'tightens' or 'smooths' the curve.

Figure 5–42 illustrates the effect of multiple or coincident vertices on the defining polygon. The B-spline curves are all of the order $k = 4$. The lowest curve is defined by four polygon vertices as shown. Here the knot vector is [0 0 0 0 1 1 1 1]. The intermediate curve is defined by five polygon vertices with two coincident vertices at the second polygon vertex, i.e., [3 9]. The knot vector for this curve is [0 0 0 0 1 2 2 2 2]. The highest curve is defined by six polygon vertices with three coincident vertices at [3 9]. The knot vector is [0 0 0 0 1 2 3 3 3 3]. Thus, the defining polygons for the three curves are B_1, B_2, B_3, B_4; B_1, B_2, B_2, B_3, B_4 and $B_1, B_2, B_2, B_2, B_3, B_4$, respectively.

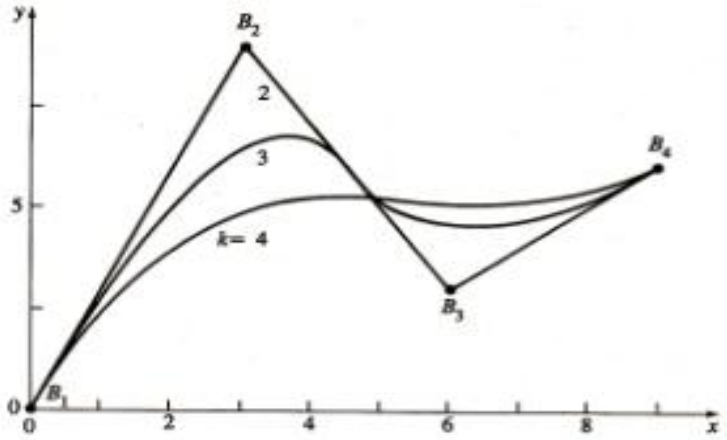


Figure 5–41 Effect of varying order on B-spline curves.

The lowest curve is composed of a single cubic segment. The intermediate curve is composed of two segments joined midway between B_2 and B_3. The highest curve is composed of three segments. The first is from B_1 to B_2, the second from B_2 to midway between B_2 and B_3. The final segment is from midway between B_2 and B_3 to B_4. Notice that as the number of multiple vertices at B_2 increases, the curve is pulled closer to B_2. If the number of multiple vertices is $k - 1$, then a sharp corner or cusp is created. This sharp corner is predicted by the convex hull properties of B-spline curves. Close examination of Fig. 5–42 shows that on both sides of the multiple vertex location a linear segment occurs.

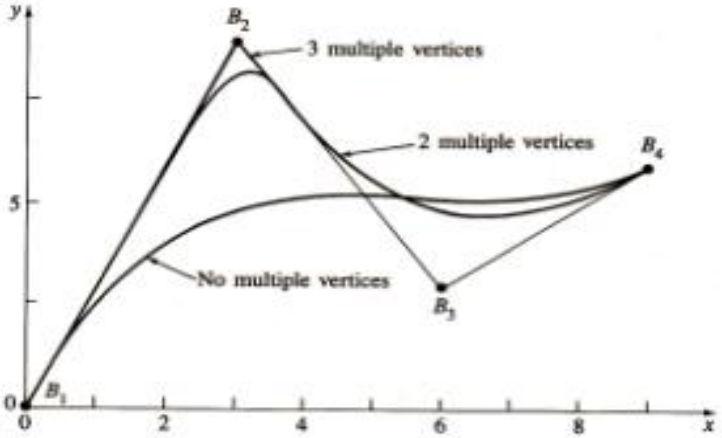


Figure 5–42 Effect of multiple vertices at B_2 on a B-spline curve , $k = 4$.

Although a cusp exists when $k-1$ multiple vertices occur, the C^{k-2} differentiability of the curve is maintained. At first glance, this might seem contradictory. However, a cusp is defined by a zero tangent vector. But a zero tangent vector does not preclude the tangent vector varying continuously. The ability to include sharp corners or cusps within a continuously C^{k-2} differentiable curve is an important characteristic of B-spline curves.

Finally notice that each of the curves has the same slope at the ends.

Figure 5–43 shows three fourth-order B-spline curves. The defining polygons each have eight vertices as indicated. The three curves shown are obtained by moving the polygon vertex B_5 successively to B_5' and B_5''. Note that moving B_5 influences the curve only over a limited region. Specifically, only the curve segments corresponding to the polygon spans B_3B_4, B_4B_5 and B_5B_6, B_6B_7 are affected by the movement of B_5. In general, the curve is affected only over those curve segments corresponding to $\pm k/2$ polygon spans around the displaced point.

A detailed example more fully illustrates the technique for calculating open B-spline curves.

Example 5–12 Calculating an Open B-spline Curve

Consider the same defining polygon used previously in Ex. 5–7 to determine a Bézier curve, i.e., B_1 [1 1], B_2 [2 3], B_3 [4 3], B_4 [3 1]. Calculate both second- and fourth-order B-spline curves.

For $k = 2$ the open knot vector is

[0 0 1 2 3 3]

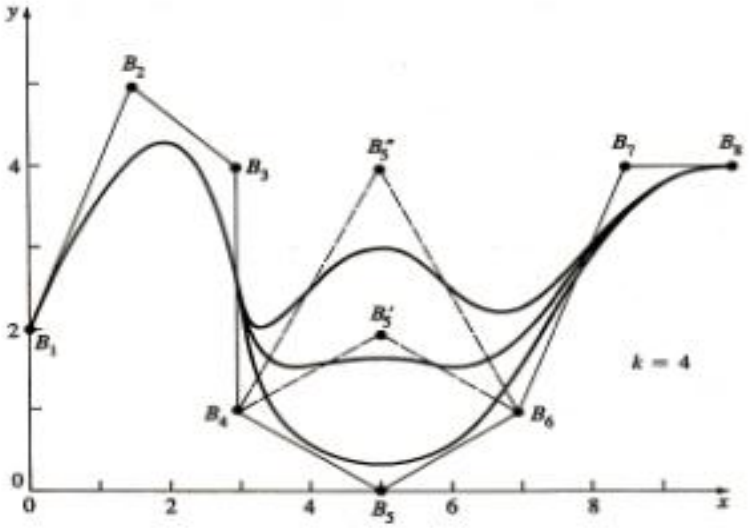


Figure 5–43 Local control of B-spline curves.

where $x_1 = 0,\ x_2 = 0, \cdots, x_6 = 3$. The parameter range is $0 \le t \le 3$. The curve is composed of three linear $(k-1=1)$ segments. For $0 \le t < 3$ the basis functions are:

$0 \le t < 1$

$$N_{2,1}(t) = 1; \qquad N_{i,1}(t) = 0, \quad i \neq 2$$
$$N_{1,2}(t) = 1 - t; \qquad N_{2,2}(t) = t; \qquad N_{i,2}(t) = 0, \quad i \neq 1, 2$$

$1 \le t < 2$

$$N_{3,1}(t) = 1; \qquad N_{i,1}(t) = 0, \quad i \neq 3$$
$$N_{2,2}(t) = 2 - t; \qquad N_{3,2}(t) = (t-1); \qquad N_{i,2}(t) = 0, \quad i \neq 2, 3$$

$2 \le t < 3$

$$N_{4,1}(t) = 1; \qquad N_{i,1}(t) = 0, \quad i \neq 4$$
$$N_{3,2}(t) = (3 - t); \qquad N_{4,2}(t) = (t-2); \qquad N_{i,2}(t) = 0, \quad i \neq 3, 4$$

Using Eq. (5–83) the parametric B-spline curve is

$$P(t) = B_1 N_{1,2}(t) + B_2 N_{2,2}(t) + B_3 N_{3,2}(t) + B_4 N_{4,2}(t)$$

For each of these intervals the curve is given by

$$P(t) = (1-t)B_1 + tB_2 = B_1 + (B_2 - B_1)t \qquad 0 \le t < 1$$
$$P(t) = (2-t)B_2 + (t-1)B_3 = B_2 + (B_3 - B_2)t \qquad 1 \le t < 2$$
$$P(t) = (3-t)B_3 + (t-2)B_4 = B_3 + (B_4 - B_3)t \qquad 2 \le t < 3$$

In each case the result is the equation of the parametric straight line for the polygon span, i.e., the 'curve' is the defining polygon.

The last point on the curve ($t = t_{max} = 3$) requires special consideration. Because of the open right-hand interval in Eq. (5–84a) all the basis functions $N_{i,k}$ at $t = 3$ are zero. Consequently, the last polygon point does not technically lie on the B-spline curve. However, practically it does. Consider $t = 3 - \epsilon$ where ϵ is an infinitesimal value. Letting $\epsilon \to 0$ shows that in the limit the last point on the curve and the last polygon point are coincident. Practically, this result is incorporated by either artificially adding the last polygon point to the curve description or by defining $N(t = t_{max}) = 1.0$.

For $k = 4$ the order of the curve is equal to the number of defining polygon vertices. Thus the B-spline curve reduces to a Bézier curve. The knot vector with $t_{max} = n - k + 2 = 3 - 4 + 2 = 1$ is [0 0 0 0 1 1 1 1]. The basis functions are

$0 \le t < 1$

$$N_{4,1}(t) = 1; \qquad N_{i,1}(t) = 0, \quad i \neq 4$$
$$N_{3,2}(t) = (1-t); \qquad N_{4,2}(t) = t; \qquad N_{i,2}(t) = 0, \quad i \neq 3, 4$$
$$N_{2,3}(t) = (1-t)^2; \qquad N_{3,3}(t) = 2t(1-t);$$
$$N_{4,3}(t) = t^2; \qquad N_{i,3}(t) = 0, \quad i \neq 2, 3, 4$$
$$N_{1,4}(t) = (1-t)^3; \qquad N_{2,4}(t) = t(1-t)^2 + 2t(1-t)^2 = 3t(1-t)^2;$$
$$N_{3,4}(t) = 2t^2(1-t) + (1-t)t^2 = 3t^2(1-t); \qquad N_{4,4}(t) = t^3$$

Using Eq. (5–83) the parametric B-spline is

$$P(t) = B_1 N_{1,4}(t) + B_2 N_{2,4}(t) + B_3 N_{3,4}(t) + B_4 N_{4,4}(t)$$
$$P(t) = (1-t)^3 B_1 + 3t(1-t)^2 B_2 + 3t^2(1-t) B_3 + t^3 B_4$$

Thus, at $t = 0$

$$P(0) = B_1$$

and at $t = 1/2$

$$P\left(\frac{1}{2}\right) = \frac{1}{8}B_1 + \frac{3}{8}B_2 + \frac{3}{8}B_3 + \frac{1}{8}B_4$$

and

$$P\left(\frac{1}{2}\right) = \frac{1}{8}[\,1 \quad 1\,] + \frac{3}{8}[\,2 \quad 3\,] + \frac{3}{8}[\,4 \quad 3\,] + \frac{1}{8}[\,3 \quad 1\,]$$
$$= [\,11/4 \quad 5/2\,]$$

Comparison with Ex. 5–7 shows that the current results are identical. The resulting curve is shown in Fig. 5–28.

Turning now to periodic B-spline curves, Fig. 5–44 shows three periodic B-spline curves of different orders. Each of the curves is defined by the same polygon vertices as the open B-spline curves in Fig. 5–41. For $k = 2$ the B-spline curve again coincides with the defining polygon. However, notice that for periodic B-spline curves for $k > 2$ the first and last points on the B-spline curve do *not* correspond to the first and last defining polygon vertices. Nor in general is the slope at the first and last points the same as that of the first and last defining polygon spans. For $k = 3$ the B-spline curve starts at the midpoint of the first polygon span and ends at the midpoint of the last polygon span as indicated by the arrows. These effects are a result of the reduced parameter range for periodic B-spline basis functions. For $k = 2$ the periodic knot vector is [0 1 2 3 4 5] with a parameter range of $1 \le t \le 4$. For $k = 3$ the periodic knot vector is [0 1 2 3 4 5 6] with parameter range of $2 \le t \le 4$. For $k = 4$ the periodic knot vector is [0 1 2 3 4 5 6 7] with parameter range of $3 \le t \le 4$. Comparing these results with those for open knot vectors in Fig. 5–41 shows that the multiple knot values at the ends of the open knot vector permit the curve to be defined over the full range of parameter values. The effect is to 'pull' the curve out to the ends of the defining polygon.

Here again the fourth-order curve consists of a single cubic polynomial segment, the third-order curve of two parabolic segments joined at the center of the second polygon span with C^1 continuity, and the second-order 'curve' of three linear segments joined at the second and third polygon vertices with C^0 continuity. Notice that again increasing order has a 'smoothing' effect on the curve, but here it also decreases the curve length.

Figure 5–45 illustrates that the effect of multiple vertices in the defining polygon is similar for periodic and open B-spline curves. The small inset shows the details near B_2.

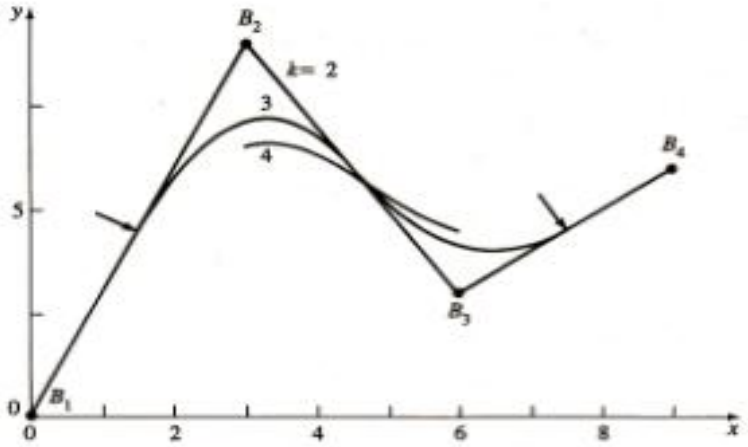


Figure 5–44 Effect of varying order on periodic B-spline curves.

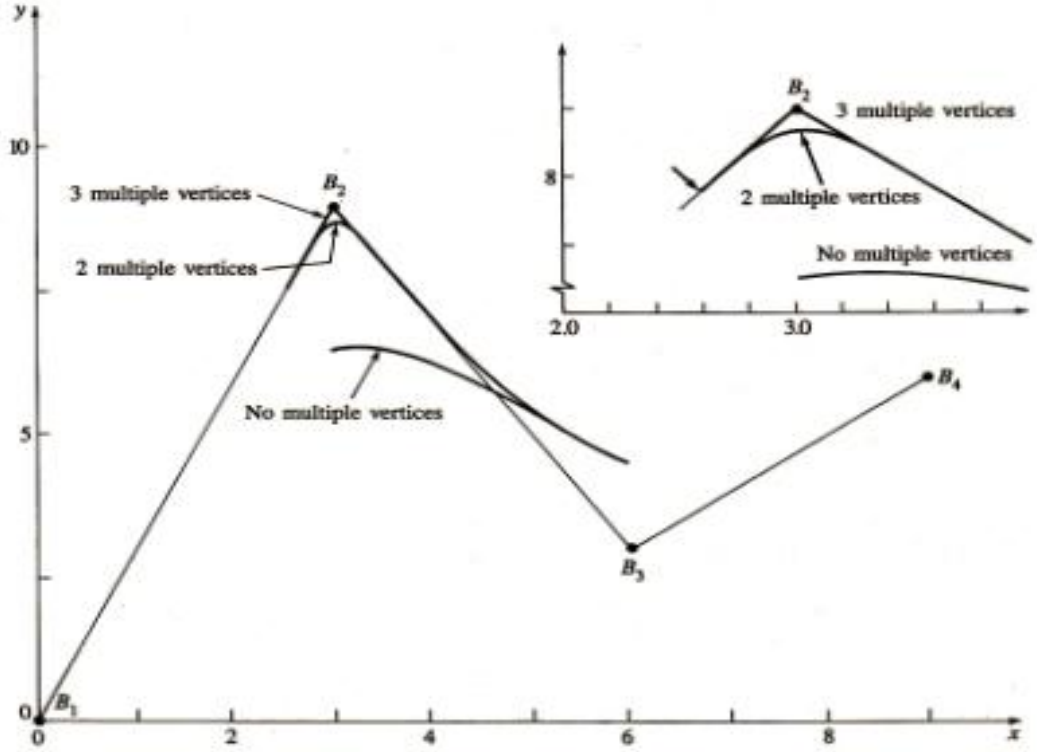


Figure 5–45 Effect of multiple vertices on a periodic B-spline curve, $k = 4$.

Example 5–13 Calculating a Periodic B-spline Curve

Again consider the defining polygon shown in Fig. 5–44. The polygon vertices are $B_1[0 \quad 0]$, $B_2[3 \quad 9]$, $B_3[6 \quad 3]$, $B_4[9 \quad 6]$. Determine the fourth-order ($k = 4$) periodic B-spline curve defined by this polygon.

For $k = 4$, $[0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7]$ with parameter range $3 \le t < 4$ is the knot vector for the periodic basis functions. The first-order basis functions for this parameter range are (see Eq. 5–84*a*)

$$3 \le t < 4$$

$$N_{4,1}(t) = 1; \qquad N_{i,1}(t) = 0, \quad i \neq 4$$

From Eq. (5–84*b*) the higher order basis functions are then

$$N_{3,2}(t) = (4 - t); \qquad N_{4,2}(t) = (t - 3); \qquad N_{i,2}(t) = 0, \quad i \neq 3, 4$$

$$N_{2,3}(t) = \frac{(4-t)^2}{2}; \qquad N_{3,3}(t) = \frac{(t-2)(4-t)}{2} + \frac{(5-t)(t-3)}{2};$$

$$N_{4,3}(t) = \frac{(t-3)^2}{2}; \qquad N_{i,3}(t) = 0, \quad i \neq 2, 3, 4$$

$$N_{1,4}(t) = \frac{(4-t)^3}{6};$$

$$N_{2,4}(t) = \frac{(t-1)(4-t)^2}{6} + \frac{(5-t)(4-t)(t-2)}{6} + \frac{(5-t)^2(t-3)}{6};$$

$$N_{3,4}(t) = \frac{(t-2)^2(4-t)}{6} + \frac{(t-2)(t-3)(5-t)}{6} + \frac{(6-t)(t-3)^2}{6};$$

$$N_{4,4}(t) = \frac{(t-3)^3}{6}$$

At $t = 3$

$$N_{1,4}(3) = \frac{(4-3)^2}{6} = \frac{1}{6}$$

$$N_{2,4}(3) = \frac{(3-1)(4-3)^2}{6} + \frac{(5-3)(4-3)(3-2)}{6} + \frac{(5-3)^2(3-3)}{6} = \frac{2}{3}$$

$$N_{3,4}(3) = \frac{(3-2)^2(4-3)}{6} + \frac{(3-2)(3-3)(5-3)}{6} + \frac{(6-3)(3-3)^2}{6} = \frac{1}{6}$$

$$N_{4,4}(3) = \frac{(3-3)^3}{6} = 0$$

The point on the B-spline curve at $t = 3$ is thus

$$P(3) = \frac{1}{6}B_1 + \frac{2}{3}B_2 + \frac{1}{6}B_3 + 0B_4$$

$$= \frac{1}{6}[0 \quad 0] + \frac{2}{3}[3 \quad 9] + \frac{1}{6}[6 \quad 3] + 0[9 \quad 6]$$

$$= [3 \quad 6.5]$$

The complete curve is shown in Fig. 5–44.

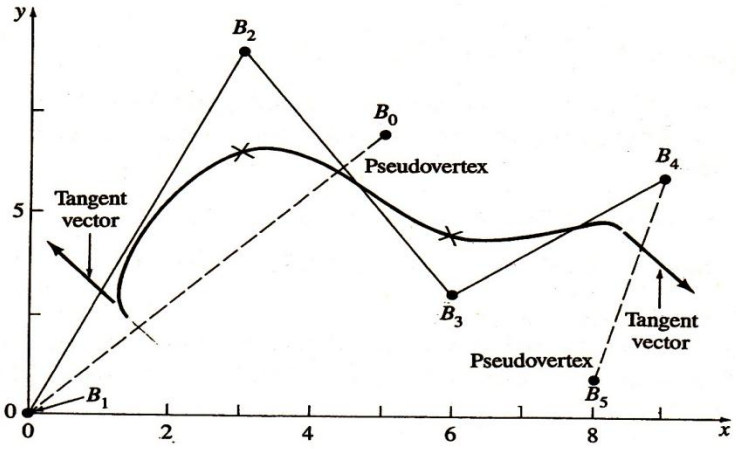


Figure 5–54 Tangent vector control for periodic B-spline curves, $k = 4$.

and

$$B_{n+2} = \frac{1}{\sum_{i=1}^{k-1}(k-1)(k-i-1)N^*_{i,k}}\Big\{(k-1)!\,P''_e - \Big(\sum_{i=1}^{k-2}(k-1)(k-i-1)N^*_{i,2}B_{n-k+4} + \cdots + \sum_{i=1}^{k-2}(k-1)(k-i-1)N^*_{i,k-1}B_{n+1}\Big)\Big\} \qquad n \geq k \tag{5-114}$$

For $k = 4$, Eqs. (5–113) and (5–114) yield

$$B_0 = P''_s + 2B_1 - B_2$$
$$B_{n+2} = P''_e + 2B_{n+1} - B_n$$

Again the start and end points are obtained by substituting these values into Eqs. (5–107) and (5–108). Similarly, the tangent vectors are obtained by using Eqs. (5–103) and (5–104) rewritten in terms of B_0 to B_{n+2}.

5–11 B-SPLINE CURVE FIT

The previous sections discussed the generation of a B-spline curve from its defining polygon. Here, determining a polygon that generates a B-spline curve for a

set of *known* data points is considered. The problem is shown schematically in Fig. 5–55.

If a data point lies on the B-spline curve, then it must satisfy Eq. (5–83). Writing Eq. (5–83) for each of j data points yields

$$\begin{aligned} D_1(t_1) &= N_{1,k}(t_1)B_1 + N_{2,k}(t_1)B_2 + \cdots + N_{n+1,k}(t_1)B_{n+1} \\ D_2(t_2) &= N_{1,k}(t_2)B_1 + N_{2,k}(t_2)B_2 + \cdots + N_{n+1,k}(t_2)B_{n+1} \\ &\vdots \\ D_j(t_j) &= N_{1,k}(t_j)B_1 + N_{2,k}(t_j)B_2 + \cdots + N_{n+1,k}(t_j)B_{n+1} \end{aligned}$$

where $2 \le k \le n+1 \le j$. This system of equations is more compactly written in matrix form as

$$[\,D\,] = [\,N\,][\,B\,] \qquad (5-115)$$

where

$$[\,D\,]^T = [\,D_1(t_1) \quad D_2(t_2) \quad \cdots \quad D_j(t_j)\,]$$

$$[\,B\,]^T = [\,B_1 \quad B_2 \quad \cdots \quad B_{n+1}\,]$$

$$[\,N\,] = \begin{bmatrix} N_{1,k} & \cdots & \cdots & N_{n+1,k}(t_1) \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ N_{1,k}(t_j) & \cdots & \cdots & N_{n+1,k}(t_j) \end{bmatrix}$$

If $2 \le k \le n+1 = j$, then the matrix $[\,N\,]$ is square and the defining polygon is obtained directly by matrix inversion, i.e.,

$$[\,B\,] = [\,N\,]^{-1}[\,D\,] \qquad 2 \le k \le n+1 = j \qquad (5-116)$$

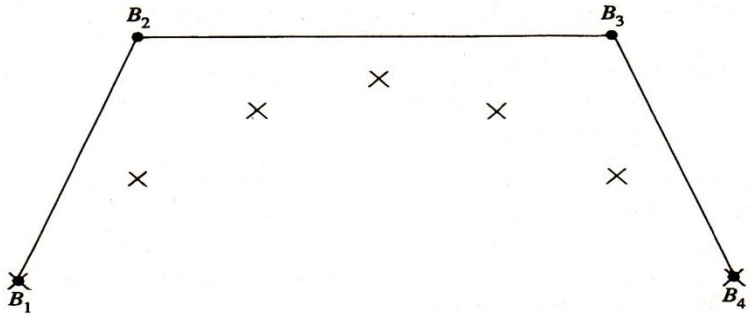


Figure 5–55 Determining a B-spline polygon for a known data set.

In this case, the resulting B-spline curve passes through each data point, i.e., a curve fit is obtained. Although the continuity of the resulting curve is everywhere C^{k-2}, it may not be 'smooth', or 'sweet' or 'fair'. The fitted curve may develop unwanted wiggles or undulations.

A fairer or smoother curve is obtained by specifying fewer defining polygon points than data points, i.e., $2 \leq k \leq n+1 < j$. Here, $[N]$ is no longer square, the problem is overspecified and can only be solved in a mean sense. Recalling that a matrix times its transpose is always square (see Sec. 3–21), the defining polygon vertices for a B-spline curve that fairs or smooths the data are given by

$$[D] = [N][B]$$

$$[N]^T[D] = [N]^T[N][B]$$

and

$$[B] = [[N]^T \quad [N]]^{-1}[N]^T[D] \qquad (5-117)$$

Both of these techniques assume that the matrix $[N]$ is known. Provided that the order of the B-spline basis k, the number of defining polygon points $n+1$, and the parameter value along the curve are known, then the basis functions $N_{i,k}(t_j)$ and hence the matrix $[N]$ can be obtained. Within the restrictions $2 \leq k \leq n+1 \leq j$, the order and number of polygon vertices are arbitrary.

The parameter value t_j for each data point is a measure of the data point's distance along the B-spline curve. One useful approximation for this parameter value uses the chord length between data points. Specifically, for j data points the parameter value at the lth data point is

$$t_1 = 0$$

$$\frac{t_l}{t_{\max}} = \frac{\sum_{s=2}^{l} |D_s - D_{s-1}|}{\sum_{s=2}^{j} |D_s - D_{s-1}|} \qquad l \geq 2$$

The maximum parameter value $t_{\max}$ is usually taken as the maximum value of the knot vector. Similar schemes are mentioned in Refs. (5–15) and (5–22).

For an open uniform knot vector with multiplicity of the knot values at the ends equal to k, a Bézier curve is obtained when $n = k$.

An example illustrates these techniques.

Example 5–17 B-spline Curve Fit

For the five data points D_1 [0 0], D_2 [1.5 2], D_3 [3 2.5], D_4 [4.5 2], D_5 [6 0] determine the third-order ($k = 3$) defining polygons having five and four polygon vertices that generate a B-spline curve 'through' the data points. Use the chord length approximation for the parameter values along the B-spline curve corresponding to the data points.

First determine the chord lengths.

$$D_{21} = |D_2 - D_1| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} = \sqrt{(1.5)^2 + (2)^2} = \sqrt{6.25} = 2.5$$

$$D_{32} = |D_3 - D_2| = \sqrt{(1.5)^2 + (.5)^2} = 1.58$$

$$D_{43} = |D_4 - D_3| = \sqrt{(1.5)^2 + (-.5)^2} = 1.58$$

$$D_{54} = |D_5 - D_4| = \sqrt{(1.5)^2 + (-2)^2} = 2.5$$

and

$$\sum_{s=2}^{5} (D_s - D_{s-1}) = D_{51} = 8.16$$

Thus

$$t_1 = 0$$

$$\frac{t_2}{t_{max}} = \frac{D_{21}}{D_{51}} = \frac{2.5}{8.16} = 0.31$$

$$\frac{t_3}{t_{max}} = \frac{D_{31}}{D_{51}} = \frac{(2.5 + 1.58)}{8.16} = 0.5$$

$$\frac{t_4}{t_{max}} = \frac{D_{41}}{D_{51}} = \frac{(2.5 + 1.58 + 2.5)}{8.16} = 0.81$$

$$\frac{t_5}{t_{max}} = \frac{D_{51}}{D_{51}} = 1$$

For five polygon vertices, the maximum value of the knot vector for a third-order B-spline curve is $n - k + 2 = 4 - 3 + 2 = 3$. n is one less than the number of polygon vertices. The knot vector with multiplicity k at the ends is

$$[\,0 \quad 0 \quad 0 \quad 1 \quad 2 \quad 3 \quad 3 \quad 3\,]$$

With these values Eq. (5–115) becomes

$$[D] = [N][B]$$

$$\begin{bmatrix} 0 & 0 \\ 1.5 & 2 \\ 3 & 2.5 \\ 4.5 & 2 \\ 6 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.007 & 0.571 & 0.422 & 0 & 0 \\ 0 & 0.125 & 0.75 & 0.125 & 0 \\ 0 & 0 & 0.422 & 0.571 & 0.007 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} [B]$$

Solving for $[B]$ yields

$$[B] = [N]^{-1}[D]$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -0.013 & 2.037 & -1.307 & 0.286 & -0.002 \\ 0.003 & -0.387 & 1.769 & -0.387 & 0.003 \\ -0.002 & 0.286 & -1.307 & 2.037 & -0.013 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1.5 & 2 \\ 3 & 2.5 \\ 4.5 & 2 \\ 6 & 0 \end{bmatrix}$$

$$[B] = \begin{bmatrix} 0 & 0 \\ 0.409 & 1.378 \\ 3 & 2.874 \\ 5.591 & 1.377 \\ 6 & 0 \end{bmatrix}$$

Figure 5–56a shows the original data points, the calculated polygon vertices and the resulting curve.

For four polygon vertices, the knot vector with multiplicity k at the ends is

$$[0 \quad 0 \quad 0 \quad 1 \quad 2 \quad 2 \quad 2]$$

$[N]$ becomes

$$[N] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.15 & 0.662 & 0.188 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0.188 & 0.662 & 0.15 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Multiplying by $[N]^T$ and taking the inverse yields

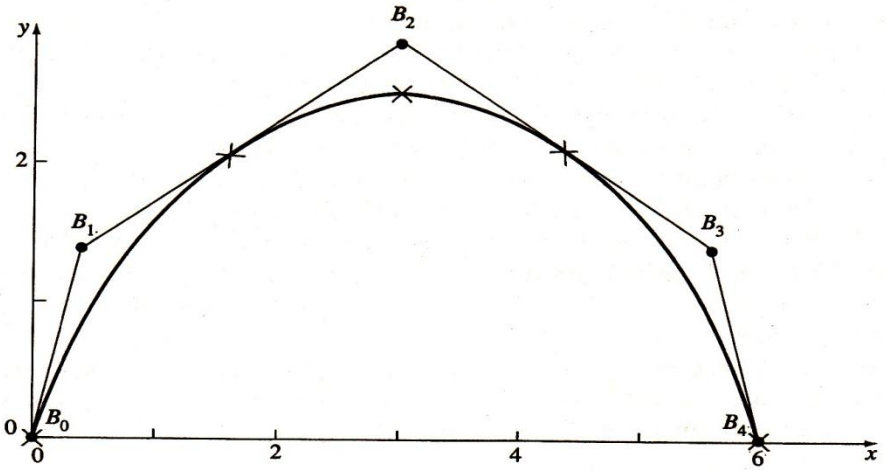
$$[[N]^T[N]]^{-1} = \begin{bmatrix} 0.995 & -0.21 & 0.106 & -0.005 \\ -0.21 & 2.684 & -1.855 & 0.106 \\ 0.106 & -1.855 & 2.684 & -0.21 \\ -0.005 & 0.106 & -0.21 & 0.995 \end{bmatrix}$$

Equation (5–117) then gives

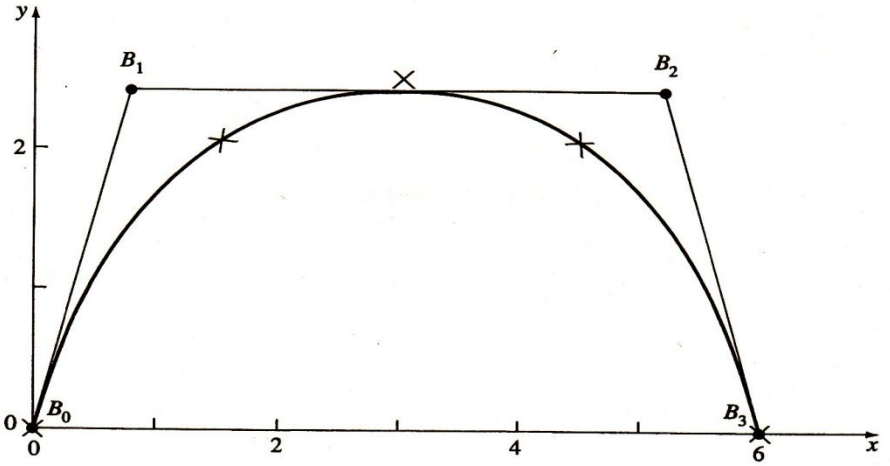
$$[B] = [[N]^T[N]]^{-1}[N]^T[D] = \begin{bmatrix} 0 & 0 \\ 0.788 & 2.414 \\ 5.212 & 2.414 \\ 6 & 0 \end{bmatrix}$$

The original data, the calculated polygon vertices and the resulting curve are shown in Fig. 5–56b. Notice that except at the ends the curve does not pass through the original data points.

The above fitting technique allows each of the determined defining polygon points for the B-spline curve to be located anywhere in three space. In some design situations it is more useful to constrain the defining polygon points to lie at a particular coordinate value, say x = constant. An example of such a design situation is in fitting B-spline curves to existing ships' lines. Rogers and Fog (Ref. 5–26) have developed such a technique for both curves and surfaces. Essentially, the technique iterates the parameter value of the fixed coordinate until the value on the B-spline curve at the assumed parameter value calculated with the defining polygons obtained using the above fitting technique is within some specified amount of the fixed value, i.e., $|x_{\text{fixed}} - x_{\text{calc}}| \leq$ error. The resulting fit is less accurate but more convenient for subsequent modification.



(a)



(b)

Figure 5–56 Results for Ex. 5–17. (a) Five polygon vertices; (b) four polygon vertices.

5–12 B-SPLINE CURVE SUBDIVISION

The flexibility of a Bézier curve is increased by raising the degree of the defining polynomial curve by adding an additional vertex to the defining polygon (see Sec. 5–8). The flexibility of a B-spline curve can also be increased by raising the order of the defining B-spline basis and hence of the defining polynomial segments. Cohen et al. (Ref. 5–27) provide both the theory and an algorithm for degree raising of B-spline curves.

$$[X] = [0 \quad 0 \quad 0 \quad 1 \quad 2 \quad 2 \quad 2]$$

with $x_1 = 0$, $x_2 = 0, \cdots, x_7 = 2$. The new knot vector is

$$[Y] = [0 \quad 0 \quad 0 \quad 1 \quad 1 \quad 2 \quad 2 \quad 2]$$

with $y_1 = 0$, $y_2 = 0, \cdots, y_8 = 2$. There are five new polygon vertices $C_1 \cdots C_5$.

The nonzero $\alpha_{i,j}^k$'s required to determine the C_j 's are

$$k = 1$$
$$\alpha_{3,1}^1 = \alpha_{3,2}^1 = \alpha_{3,3}^1 = \alpha_{3,4}^1 = \alpha_{4,5}^1 = 1$$
$$k = 2$$
$$\alpha_{2,1}^2 = \alpha_{2,2}^2 = \alpha_{3,3}^2 = \alpha_{3,4}^2 = \alpha_{4,5}^2 = 1$$
$$k = 3$$
$$\alpha_{1,1}^3 = \alpha_{2,2}^3 = \alpha_{3,4}^3 = \alpha_{4,5}^3 = 1, \quad \alpha_{3,3}^3 = \alpha_{2,3}^3 = \frac{1}{2}$$

The new polygon vertices are

$$C_1 = \alpha_{3,1}^3 B_1 = B_1 = [0 \quad 0]$$
$$C_2 = \alpha_{3,2}^3 B_2 = B_2 = [1 \quad 1]$$
$$C_3 = \alpha_{2,3}^3 B_2 + \alpha_{3,3}^3 B_3 = \frac{1}{2}(B_2 + B_3) = \frac{1}{2}([1 \quad 1] + [2 \quad 1]) = \left[\frac{3}{2} \quad 1\right]$$
$$C_4 = \alpha_{3,4}^3 B_3 = B_3 = [2 \quad 1]$$
$$C_5 = \alpha_{4,5}^3 B_4 = B_4 = [3 \quad 0]$$

If C_3 is moved to coincide with C_2, i.e., $C_2 = C_3 = [1 \quad 1]$, both a double vertex and a double knot value corresponding to $C_2 = C_3$ exist. The resulting B-spline curve for $k = 3$ has a cusp or sharp corner at $C_2 = C_3$ (see Prob. 5–30).

5–13 RATIONAL B-SPLINE CURVES

Rational curve and surface descriptions were first introduced into the computer graphics literature by Coons (Ref. 5–32). Rational forms of the cubic spline and Bézier curves previously discussed in this chapter are well known in the literature (see Refs. 5–33 to 5–37). Rational forms of the conic sections are also well known (see Ref. 5–38). Both because of space limitations and because they form a unifying foundation, the current discussion is limited to rational B-spline curves. Rational B-splines provide a single precise mathematical form capable of

representing the common analytical shapes—lines, planes, conic curves including circles, free-form curves, quadric and sculptured surfaces—used in computer graphics and computer aided design.

Versprille (Ref. 5–39) was the first to discuss rational B-splines. The seminal papers by Tiller (Ref. 5–40) and Piegl and Tiller (Ref. 5–41) form the basis of the current discussion. Interestingly enough, nonuniform rational B-splines (NURBS) have been an Initial Graphics Exchange Specification (IGES) standard since 1983 (see Ref. 5–42). IGES is the standard for the interchange of design information between various computer aided design systems and between computer aided design and computer aided manufacturing systems. Rational B-splines have been incorporated into a number of geometric modeling systems. They have also been implemented in hardware (VLSI or microcode) by a number of graphics workstation manufacturers.

A rational B-spline curve is the projection of a nonrational (polynomial) B-spline curve defined in four-dimensional (4D) homogeneous coordinate space back into three-dimensional (3D) physical space. Specifically,

$$P(t) = \sum_{i=1}^{n+1} B_i^h N_{i,k}(t) \qquad (5-121)$$

where the B_i^h's are the 4D homogeneous defining polygon vertices for the nonrational 4D B-spline curve. $N_{i,k}(t)$ is the nonrational B-spline basis function previously given in Eq. (5–84).

Projecting back into three-dimensional space by dividing through by the homogeneous coordinate yields the rational B-spline curve

$$P(t) = \frac{\sum_{i=1}^{n+1} B_i h_i N_{i,k}(t)}{\sum_{i=1}^{n+1} h_i N_{i,k}(t)} = \sum_{i=1}^{n+1} B_i R_{i,k}(t) \qquad (5-122)$$

where the B_i's are the 3D defining polygon vertices for the rational B-spline curve and the

$$R_{i,k}(t) = \frac{h_i N_{i,k}(t)}{\sum_{i=1}^{n+1} h_i N_{i,k}(t)} \qquad (5-123)$$

are the rational B-spline basis functions. Here, $h_i \geq 0$ for all values of i.†

As can be seen from Eqs. (5–121) to (5–123), rational B-spline basis functions and curves are a generalization of nonrational B-spline basis functions and curves. They carry forward nearly all the analytic and geometric characteristics of their nonrational B-spline counterparts. In particular:

Each rational basis function is positive or zero for all parameter values, i.e., $R_{i,k} \geq 0$.

† Note that rational B-spline basis functions for $h_i < 0$ are valid (see Ref. 5–39) but are not convenient in terms of the current discussion.

The sum of the rational B-spline basis functions for any parameter value t is one, i.e.,

$$\sum_{i=1}^{n+1} R_{i,k}(t) \equiv 1 \qquad (5-124)$$

Except for $k = 1$, each rational basis function has precisely one maximum.

A rational B-spline curve of order k (degree $k-1$) is C^{k-2} continuous everywhere.

The maximum order of the rational B-spline curve is equal to the number of defining polygon vertices.

A rational B-spline curve exhibits the variation diminishing property.

A rational B-spline curve generally follows the shape of the defining polygon.

A rational B-spline curve lies within the union of convex hulls formed by k successive defining polygon vertices.

Any *projective* transformation is applied to a rational B-spline curve by applying it to the defining polygon vertices; i.e., the curve is invariant with respect to a *projective* transformation. Note that this is a stronger condition than that for a nonrational B-spline which is only invariant with respect to an *affine* transformation.

From Eqs. (5–85) and (5–123) it is clear that when all $h_i = 1$, $R_{i,k}(t) = N_{i,k}(t)$. Thus, nonrational B-spline basis functions and curves are included as a special case of rational B-spline basis functions and curves. Further, it is easy to show that an open rational B-spline curve with order equal to the number of defining polygon vertices is a rational Bézier curve. For the case of all $h_i = 1$, the rational Bézier curve reduces to a nonrational Bézier curve. Thus, both rational and nonrational Bézier curves are included as special cases of rational B-spline curves.

Since rational B-splines are a four-dimensional generalization of nonrational B-splines, algorithms for degree-raising (see Ref. 5–27 and Ex. 6–18), subdivision (see Sec. 5–12 and Refs. 5–28 to 5–31) and curve fitting (see Sec. 5–11) of nonrational B-spline curves are valid for rational B-splines simply by applying them to the 4D defining polygon vertices.

Open uniform, periodic uniform and nonuniform knot vectors can be used to generate rational B-spline basis functions and rational B-spline curves.

In Eqs. (5–122) and (5–123) the homogeneous coordinates h_i (occasionally called weights) provide additional blending capability. $h = 1$ is called the affine space. By convention it corresponds to physical space. The effect of the homogeneous coordinates h on the rational B-spline basis functions is shown in Fig. 5–57. Here, an open uniform knot vector [0 0 0 1 2 3 3 3] ($n+1 = 5, k = 3$) is used with a homogeneous coordinate vector $h_i = 1$, $i \neq 3$. Values of h_3 range from 0 to 5. The rational B-spline basis functions shown in Fig. 5–57c with $h = 1$ are identical to the corresponding nonrational B-spline basis functions. The rational B-spline curve for $h_3 = 1$, shown in Fig. 5–58, is also identical

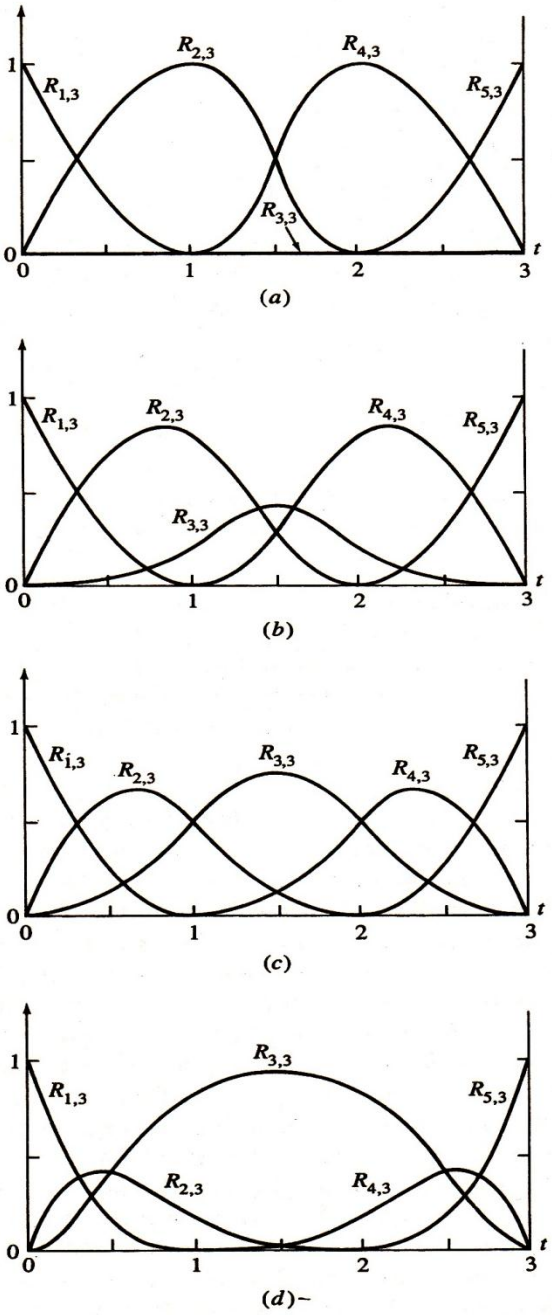


Figure 5–57 Rational B-spline basis functions for $n + 1 = 5$, $k = 3$ with open knot vector $[X] = [0\ \ 0\ \ 0\ \ 1\ \ 2\ \ 3\ \ 3\ \ 3]$, $[H] = [1\ \ 1\ \ h_3\ \ 1\ \ 1]$. (a) $h_3 = 0$; (b) $h_3 = 0.25$; (c) $h_3 = 1.0$; (d) $h_3 = 5.0$.

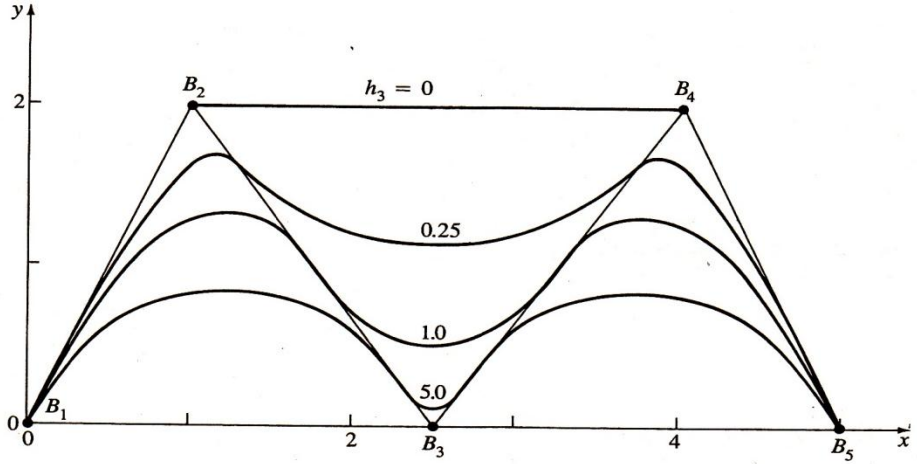


Figure 5–58 Rational B-spline curves for $n + 1 = 5$, $k = 3$ with open knot vector $[X] = [0 \quad 0 \quad 0 \quad 1 \quad 2 \quad 3 \quad 3 \quad 3]$ and $[H] = [1 \quad 1 \quad h_3 \quad 1 \quad 1]$.

with the corresponding nonrational B-spline curve. Notice that for $h_3 = 0$ (see Fig. 5–57a) $R_{3,3} = 0$ everywhere. Thus, the corresponding polygon vertex, B_3, effectively has no influence on the shape of the corresponding B-spline curve. This effect is shown in Fig. 5–58 where the defining polygon vertices B_2 and B_4 are connected by a straight line. Figure 5–57 also shows that as h_3 increases $R_{3,3}$ also increases; but, as a consequence of Eq. (5–124), $R_{2,3}$ and $R_{4,3}$ decrease. The effects on the corresponding rational B-spline curves are shown in Fig. 5–58. Note, in particular, that as h_3 increases the curve is pulled closer to B_3. Hence, as mentioned previously, the homogeneous coordinates provide additional blending capability. Similar characteristics are exhibited for the fourth-order ($k = 4$) rational B-spline basis functions and curves shown in Figs. 5–59 and 5–60, respectively. However, for the higher order curve shown in Fig. 5–60 note that for $h_3 = 0$ the curve does not degenerate to a straight line between B_2 and B_4.

Figure 5–61 shows periodic uniform basis functions for $n + 1 = 5$, $k = 3$ for a knot vector $[X] = [0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7]$ and homogeneous coordinate vector $[H] = [1 \quad 1 \quad h_3 \quad 1 \quad 1]$ with $0 \leq h_3 \leq 5$. Here, as for nonrational B-spline basis functions, the usable parameter range is $2 \leq t \leq 5$. Only this parameter range is shown in Fig. 5–61. Again, the rational B-spline basis functions for $h_3 = 1$ are identical to the corresponding nonrational basis functions. However, note that for $h_3 \neq 1$ the basis functions are no longer periodic and hence no longer translates of each other. Figure 5–62 shows the corresponding rational B-spline curves. Notice that the end points of all the curves are coincident.

Figures 5–63 and 5–64 show the corresponding rational B-spline fourth-order ($k = 4$) basis functions and curves. Here, notice that the start and end points of the curves lie along a straight line.

Recalling the $(t_{\max} - \epsilon)_{\epsilon \to 0}$ argument of Ex. 5–12, evaluation of Eqs. (5–122) and (5–123) at the ends of the curve shows that the first and last points on

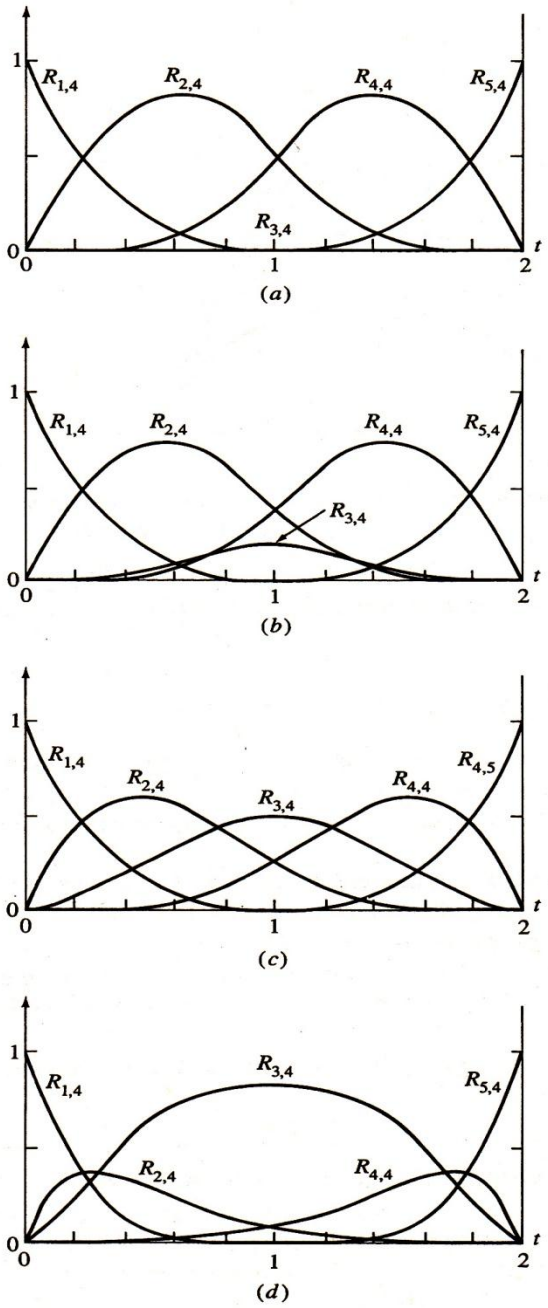


Figure 5–59 Rational B-spline basis functions for $n + 1 = 5$, $k = 4$ with open knot vector $[X] = [0\ \ 0\ \ 0\ \ 0\ \ 1\ \ 2\ \ 2\ \ 2\ \ 2]$, $[H] = [1\ \ 1\ \ h_3\ \ 1\ \ 1]$. (a) $h_3 = 0$; (b) $h_3 = 0.25$; (c) $h_3 = 1.0$; (d) $h_3 = 5.0$.

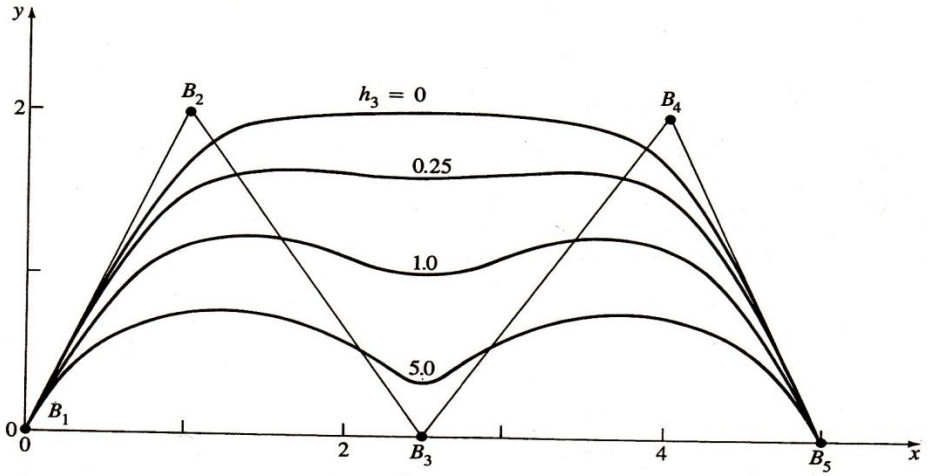


Figure 5–60 Rational B-spline curves for $n + 1 = 5$, $k = 4$ with open knot vector $[X] = [0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 2 \quad 2 \quad 2 \quad 2]$, $[H] = [1 \quad 1 \quad h_3 \quad 1 \quad 1]$.

an open rational B-spline curve are coincident with the first and last defining polygon vertices. Specifically,

$$P(0) = B_1 \qquad \text{and} \qquad P(t_{\max}) = P(n - k + 2) = B_{n+1}$$

Figure 5–65 shows that the effect of moving a single polygon vertex is similar to the results for nonrational B-splines. Here, $[H] = [1 \quad 1 \quad 0.25 \quad 1 \quad 1]$. If $h_3 = 0$, moving B_3 has no effect on the curve. As the value of h_3 increases the effect of moving B_3 increases.

Figure 5–66 shows the effect of multiple coincident vertices at B_3 on a fourth-order rational B-spline curve. Note that, like their nonrational counterparts, $k - 1$ coincident vertices yield a sharp corner, or cusp. Further, since multiple coincident vertices yield spans of zero length, the existence of the sharp corner or cusp is independent of the values of $h_i \geq 0$ corresponding to the multiple vertices (see Prob. 5–33).

An example more fully illustrates the procedure for calculation of rational B-splines.

Example 5–20 Calculation of Open Rational B-spline Curves

Consider the defining polygon given by the vertices $B_1 [0 \quad 1]$, $B_2 [1 \quad 2]$, $B_3 [2.5 \quad 0]$, $B_4 [4 \quad 2]$, $B_5 [5 \quad 0]$. Determine the point at $t = 3/2$ for the third-order ($k = 3$) open rational B-spline curve with homogeneous vectors given by $[H] = [1 \quad 1 \quad h_3 \quad 1 \quad 1]$, $h_3 = 0, 1/4, 1, 5$.

The knot vector is $[0 \quad 0 \quad 0 \quad 1 \quad 2 \quad 3 \quad 3 \quad 3]$. The parameter range is $0 \leq t \leq 3$. The curves are composed of three piecewise rational quadratics, one for each of the interior intervals in the knot vector.

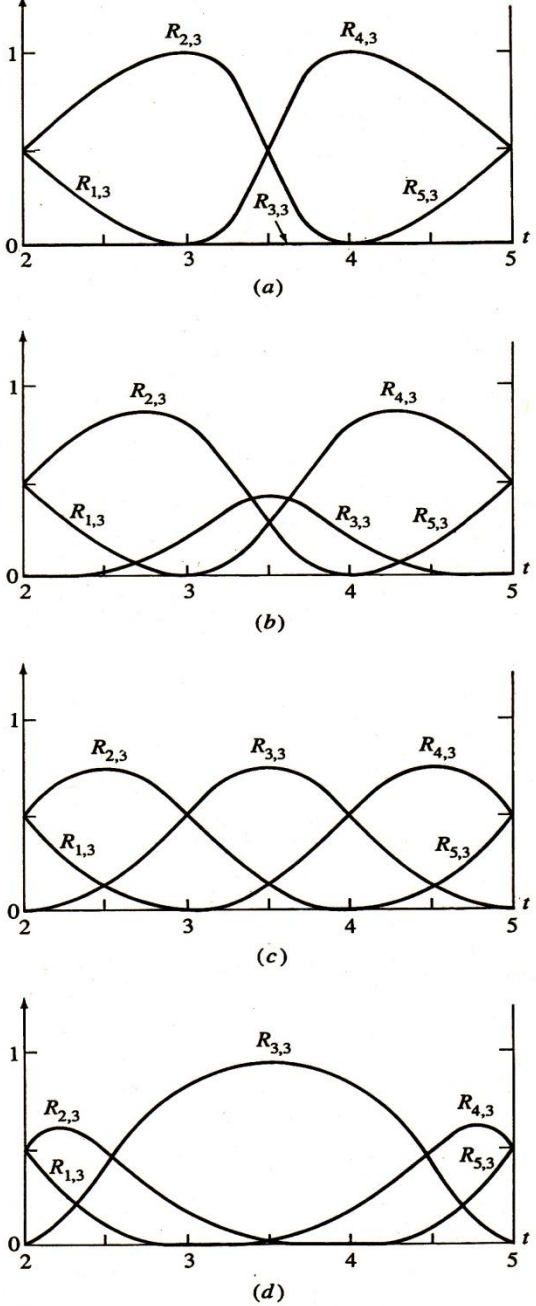


Figure 5–61 Rational B-spline basis functions for $n + 1 = 5$, $k = 3$ with periodic knot vector $[X] = [0\ \ 1\ \ 2\ \ 3\ \ 4\ \ 5\ \ 6\ \ 7]$ and $[H] = [1\ \ 1\ \ h_3\ \ 1\ \ 1]$. (a) $h_3 = 0$; (b) $h_3 = 0.25$; (c) $h_3 = 1.0$; (d) $h_3 = 5.0$.

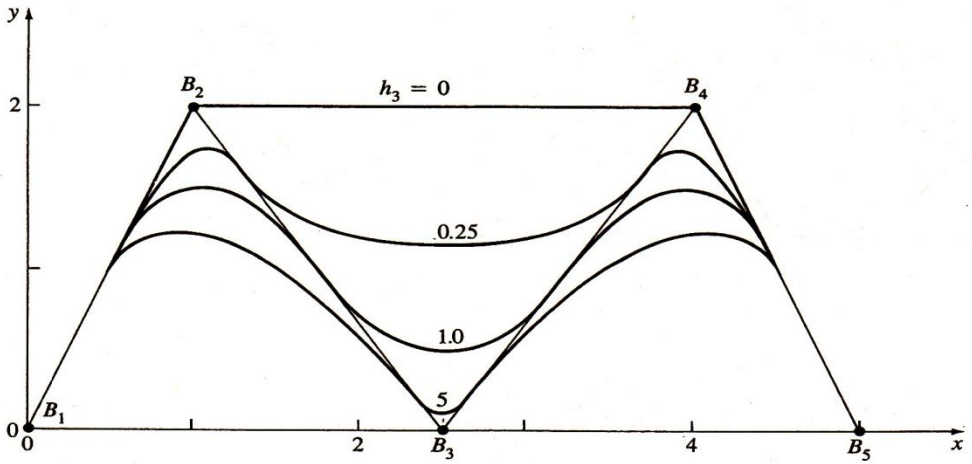


Figure 5–62 Rational B-spline curves for $n+1=5$, $k=4$ with periodic knot vector $[X]=[0\ \ 1\ \ 2\ \ 3\ \ 4\ \ 5\ \ 6\ \ 7]$ and $[H]=[1\ \ 1\ \ h_3\ \ 1\ \ 1]$.

Using Eq. (5–84) on the interval $1 \le t < 2$, the nonrational B-spline basis functions are

$$1 \le t < 2$$

$$N_{4,1}(t) = 1; \qquad N_{i,1}(t) = 0, \quad i \neq 4$$

$$N_{3,2}(t) = (2-t); \qquad N_{4,2}(t) = (t-1); \qquad N_{i,2}(t) = 0, \quad i \neq 3,4$$

$$N_{2,3}(t) = \frac{(2-t)^2}{2}; \qquad N_{3,3}(t) = \frac{t(2-t)}{2} + \frac{(3-t)(t-1)}{2};$$

$$N_{4,3}(t) = \frac{(t-1)^2}{2}; \qquad N_{i,3}(t) = 0, \quad i \neq 2,3,4$$

From Eq. (5–123) and these results, after first determining the denominator,

$$S = \sum_{i=1}^{n+1} h_i N_{i,k}(t) = h_1 N_{1,3}(t) + h_2 N_{2,3}(t) + h_4 N_{4,3}(t) + h_5 N_{5,3}(t)$$

$$= h_2 N_{2,3}(t) + h_4 N_{4,3}(t)$$

$$= \frac{(2-t)^2}{2} + \frac{(t-1)^2}{2} = \frac{2t^2 - 6t + 5}{2}$$

the rational B-spline basis functions are

$$1 \le t < 2$$

$$h_3 = 0$$

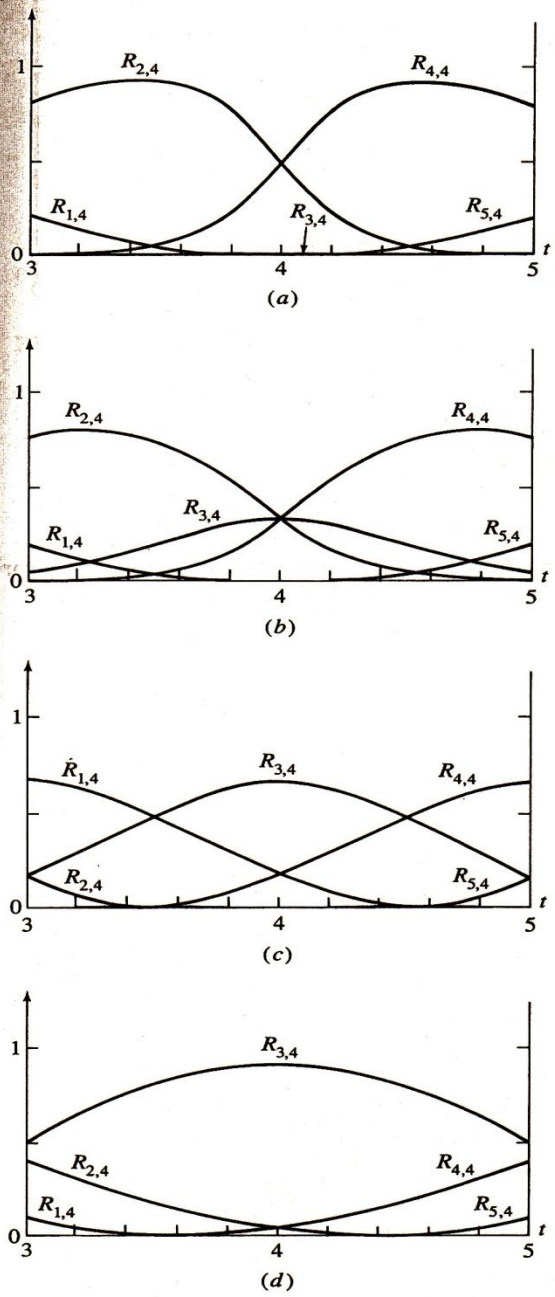


Figure 5–63 Rational B-spline basis functions for $n + 1 = 5$, $k = 4$ with periodic knot vector $[X] = [0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8]$ and $[H] = [1 \quad 1 \quad h_3 \quad 1 \quad 1]$. (a) $h_3 = 0$; (b) $h_3 = 0.25$; (c) $h_3 = 1.0$; (d) $h_3 = 5.0$.

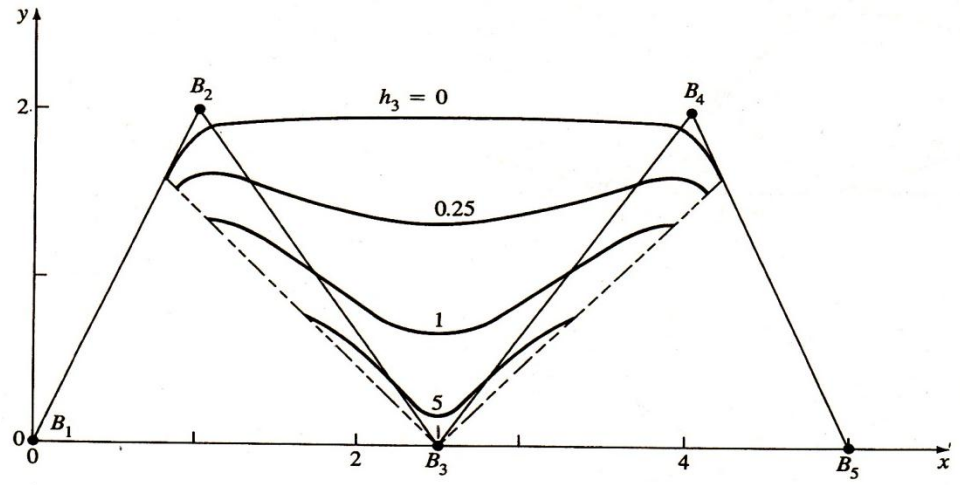


Figure 5–64 Rational B-spline curves for $n+1=5$, $k=4$ with periodic knot vector $[X] = [0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8]$, $[H] = [1 \quad 1 \quad h_3 \quad 1 \quad 1]$.

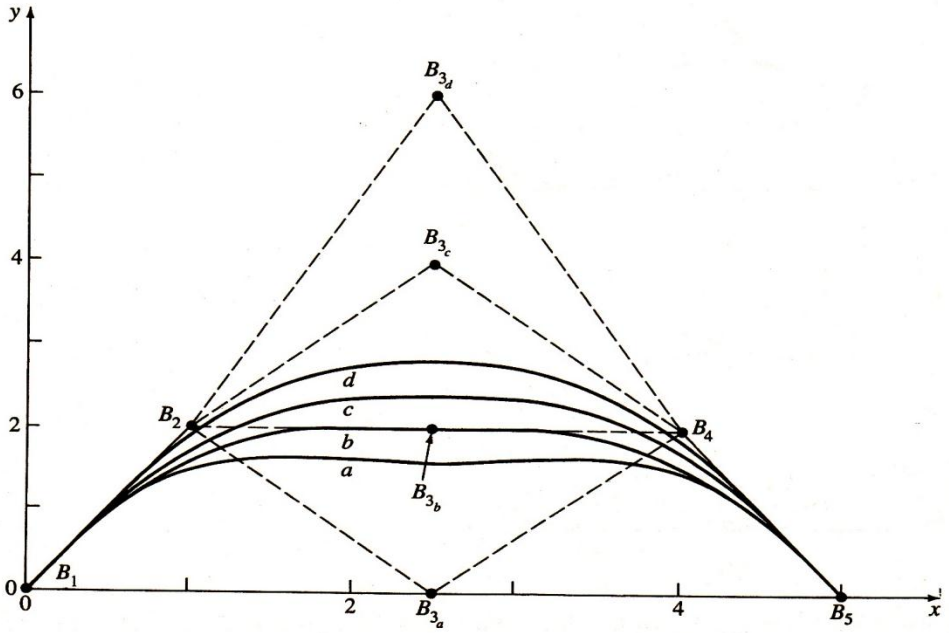


Figure 5–65 Effect of moving a single polygon vertex on a rational B-spline curve, $n+1=5$, $k=4$, $[H] = [1 \quad 1 \quad 0.25 \quad 1 \quad 1]$.

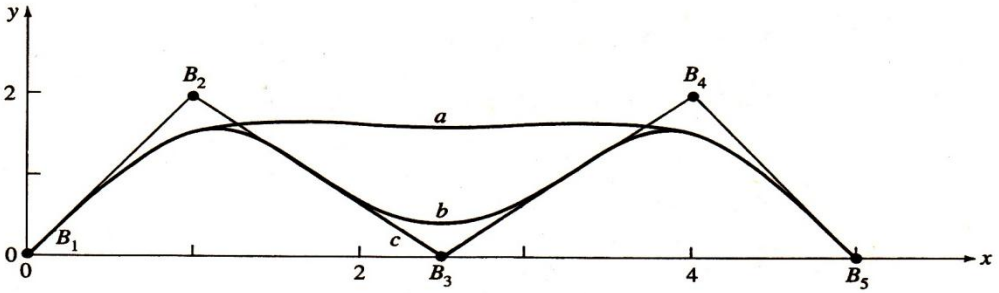


Figure 5–66 Effect of multiple vertices at B_3 on a rational B-spline curve, $n+1=5$, $k=4$. (a) Single vertex $[H] = [1 \quad 1 \quad 0.25 \quad 1 \quad 1]$; (b) double vertex $[H] = [1 \quad 1 \quad 0.25 \quad 0.25 \quad 1 \quad 1]$; (c) triple vertex $[H] = [1 \quad 1 \quad 0.25 \quad 0.25 \quad 0.25 \quad 1 \quad 1]$.

$$R_{1,3}(t) = 0$$

$$R_{2,3}(t) = \frac{h_2 N_{2,3}(t)}{S} = \frac{(2-t)^2}{2t^2 - 6t + 5}$$

$$R_{3,3}(t) = 0$$

$$R_{4,3}(t) = \frac{h_4 N_{4,3}(t)}{S} = \frac{(t-1)^2}{2t^2 - 6t + 5}$$

$$R_{5,3}(t) = 0$$

$h_3 = 1/4$

$$S = h_2 N_{2,3}(t) + h_3 N_{3,3}(t) + h_4 N_{4,3}(t)$$

$$= \frac{(2-t)^2}{2} + \frac{t(2-t)}{8} + \frac{(3-t)(t-1)}{8} + \frac{(t-1)^2}{2}$$

$$= \frac{6t^2 - 18t + 17}{8}$$

$$R_{1,3}(t) = 0$$

$$R_{2,3}(t) = \frac{4(2-t)^2}{6t^2 - 18t + 17}$$

$$R_{3,3}(t) = \frac{t(2-t) + (3-t)(t-1)}{6t^2 - 18t + 17} = \frac{-2t^2 + 6t - 3}{6t^2 - 18t + 17}$$

$$R_{4,3}(t) = \frac{4(t-1)^2}{6t^2 - 18t + 17}$$

$$R_{5,3}(t) = 0$$

$h_3 = 1$

$$S = 1$$

$$R_{1,3}(t) = 0$$

$$R_{2,3}(t) = N_{2,3}(t) = \frac{(2-t)^2}{2}$$

$$R_{3,3}(t) = N_{3,3}(t) = \frac{t(2-t)}{2} + \frac{(3-t)(t-1)}{2}$$

$$R_{4,3}(t) = N_{4,3}(t) = \frac{(t-1)^2}{2}$$

$$R_{5,3}(t) = 0$$

$h_3 = 5$

$$S = \frac{(2-t)^2}{2} + \frac{5t(2-t)}{2} + \frac{5(3-t)(t-1)}{2} + \frac{(t-1)^2}{2}$$

$$= -4t^2 + 12t - 5$$

$$R_{1,3}(t) = 0$$

$$R_{2,3}(t) = \frac{(2-t)^2}{2(-4t^2+12t-5)}$$

$$R_{3,3}(t) = \frac{5t(2-t) + 5(3-t)(t-1)}{2(-4t^2+12t-5)} = \frac{5(-2t^2+6t-3)}{2(-4t^2+12t-5)}$$

$$R_{4,3}(t) = \frac{(t-1)^2}{2(-4t^2+12t-5)}$$

$$R_{5,3}(t) = 0$$

Complete results are shown in Fig. 5–57.

Evaluating these results at $t = 3/2$ yields

$h_3 = 0$: $R_{1,3}(3/2) = 0$; $R_{2,3}(3/2) = \frac{1}{2}$; $R_{3,3}(3/2) = 0$; $R_{4,3}(3/2) = \frac{1}{2}$; $R_{5,3}(3/2) = 0$

$h_3 = \frac{1}{4}$: $R_{1,3}(3/2) = 0$; $R_{2,3}(3/2) = \frac{2}{7}$; $R_{3,3}(3/2) = \frac{3}{7}$; $R_{4,3}(3/2) = \frac{2}{7}$; $R_{5,3}(3/2) = 0$

$h_3 = 1$: $R_{1,3}(3/2) = 0$; $R_{2,3}(3/2) = \frac{1}{8}$; $R_{3,3}(3/2) = \frac{3}{4}$; $R_{4,3}(3/2) = \frac{1}{8}$; $R_{5,3}(3/2) = 0$

$h_3 = 5$: $R_{1,3}(3/2) = 0$; $R_{2,3}(3/2) = \frac{1}{32}$; $R_{3,3}(3/2) = \frac{15}{16}$; $R_{4,3}(3/2) = \frac{1}{32}$; $R_{5,3}(3/2) = 0$

The corresponding points on the rational B-spline curves are

$h_3 = 0$: $P(3/2) = \frac{1}{2}[1 \quad 2] + \frac{1}{2}[4 \quad 2] = \left[\frac{5}{2} \quad 2\right]$

$h_3 = \frac{1}{4}$: $P(3/2) = \frac{2}{7}[1 \quad 2] + \frac{3}{7}\left[\frac{5}{2} \quad 0\right] + \frac{2}{7}[4 \quad 2] = \left[\frac{5}{2} \quad \frac{8}{7}\right]$

$$h_3 = 1: \qquad P(3/2) = \frac{1}{8}[1 \quad 2] + \frac{3}{4}\left[\frac{5}{2} \quad 0\right] + \frac{1}{8}[4 \quad 2] = \left[\frac{5}{2} \quad \frac{1}{2}\right]$$

$$h_3 = 5: \qquad P(3/2) = \frac{1}{32}[1 \quad 2] + \frac{15}{16}\left[\frac{5}{2} \quad 0\right] + \frac{1}{32}[4 \quad 2] = \left[\frac{5}{2} \quad \frac{1}{8}\right]$$

Complete results are shown in Fig. 5–58.

The derivatives of rational B-spline curves are obtained by formal differentiation of Eqs. (5–122) and (5–123). Specifically,

$$P'(t) = \sum_{i=1}^{n+1} B_i R'_{i,k}(t) \tag{5-125}$$

with

$$R'_{i,k}(t) = \frac{h_i N'_{i,k}(t)}{\sum_{i=1}^{n+1} h_i N_{i,k}} - \frac{h_i N_{i,k} \sum_{i=1}^{n+1} h_i N'_{i,k}}{\left(\sum_{i=1}^{n+1} h_i N_{i,k}\right)^2} \tag{5-126}$$

Evaluating these results at $t = 0$ and $t = n - k + 2$ yields

$$P'(0) = (k-1)\frac{h_2}{h_1}(B_2 - B_1) \tag{5-127}$$

$$P'(n-k+2) = (k-1)\frac{h_n}{h_{n+1}}(B_{n+1} - B_n) \tag{5-128}$$

which shows that the direction of the slope is along the first and last polygon spans, respectively.

Higher order derivatives are obtained in a similar manner (see Probs. 5–35 and 5–36).

A simple example illustrates these results.

Example 5–21 Derivatives of Open Rational B-spline Curves

Consider the defining polygon previously used in Ex. 5–16. The polygon vertices were $B_1[1 \quad 1]$, $B_2[2 \quad 3]$, $B_3[4 \quad 3]$, $B_4[3 \quad 1]$. Determine the first derivative of the second order rational B-spline curve ($k = 2$) with $[H] = [1 \quad 1/2 \quad 1 \quad 1]$.

The knot vector is $[X] = [0 \quad 0 \quad 1 \quad 2 \quad 3 \quad 3]$. The parameter range is $0 \le t \le 3$. From Eq. (5–125) the first derivative is

$$P'(t) = B_1 R'_{1,2}(t) + B_2 R'_{2,2}(t) + B_3 R'_{3,2}(t) + B_4 R'_{4,3}(t)$$

From Eqs. (5–12) and (5–16) the nonrational basis functions and their derivatives are

$$0 \le t < 1$$

$$N_{1,2}(t) = 1 - t; \qquad N_{2,2}(t) = t; \qquad N_{i,2}(t) = 0, \quad i \ne 1,2$$

$$N'_{1,2}(t) = -1; \qquad N'_{2,2}(t) = 1; \qquad N'_{i,2}(t) = 0, \quad i \ne 1,2$$