MA 323 Geometric Modelling
Course Notes: Day 22
B-Splines and NURBS

David L. Finn

Today, we will discuss B-splines. The name B-spline arises from basis splines. Bezier curves and Bezier spline are a special case of B-splines. The advantage over B-splines is that they are more flexible and allow one to create a wider variety of curves. The disadvantage of B-splines is the complexity in their definition.

The major difference between B-splines and Bezier splines is that the B-splines vertices each are given their own basis functions. Rather than Bezier splines where the basis function (the Bernstein polynomials) are used on different control points.

22.1 Definition of a B-spline

The form of the definition of a B-spline is very much like the form of a Bezier curve, except that it uses different basis functions. Recall a Bezier curve is given by

$$\sum_{i=0}^{n} B_i^n(t)\, p_i.$$

A B-spline is given by

$$\sum_{i=0}^{n} N_{i,k}(t)\, p_i.$$

The difference is the basis function instead of the Bernstein polynomials $B_i^n(t)$ the B-splines use the functions $N_{i,k}(t)$ that are defined by the Mansfield, de Boor, Cox recursion formulas.

Before giving the exact description of the basis functions, we will list some of the properties of the basis functions.

- $N_{i,k}(t) \geq 0$ for all values of t.
- $\sum_{i=0}^{n} N_{i,k}(t) = 1$ for the parameter values of the curve $t_{min} \leq t \leq t_{max}$. Thus, we have a barycentric combination of the control points b_0, b_1, $\cdots$, b_n. One difference is that the interval $t_{min} \leq t \leq t_{max}$ must be determined from the definition of the B-spline.
- $N_{i,k}(t)$ is a piecewise defined polynomial curve of degree $k-1$.
- The domain of each basis function is the entire real line, every $t \in \mathbf{R}$, but each basis function is only nonzero for a finite segment. The exact segment is determined by a knot sequence.

22.2 The B-spline Basis functions

To define the B-spline basis functions, we need a knot sequence U, a nondecreasing sequence of real numbers $u_0 \leq u_1 \leq \cdots \leq u_m$. Next, we define the functions 0th basis functions $N_{i,0}(t)$ by

$$N_{i,0}(t) = \begin{cases} 1 & \text{if } u_i \leq t \leq u_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

then we define recursively the functions $N_{i,k}(t)$ by the recursion formulas

$$N_{i,k}(t) = \frac{t - u_i}{u_{i+k} - u_i} N_{i,k-1}(t) + \frac{u_{i+k+1} - t}{u_{i+k+1} - u_{i+1}} N_{i+1,k-1}(t).$$

To handle the possibility of $0/0$ since the knots are now not necessarily distinct, we set $0/0 = 0$ for the purpose of defining these functions. We list a few more properties of the functions $N_{i,k}(t)$ that are a consequence of the definition.

- $N_{i,k}(t) = 0$ for t outside the interval $[u_i, u_{i+k+1}]$. This property is called the *local support property.*

- In any given interval of the form $u_j \leq t \leq u_{j+1}$ at most $k+1$ of the basis functions $N_{i,k}$ are nonzero, namely the functions $N_{j-k,k}$, $N_{j+1-k,k}$, $\cdots$, $N_{j,k}$.

- The interval for which a B-spline curve is defined is $u_k \leq t \leq u_{m-k}$, for in this interval $\sum_{i=0}^{n} N_{i,k}(t) = 1$.

On the next few pages we provide some figures that display the basis functions generated by the Manfred-de Boor-Cox recursion described above.

Once, you have the basis functions it is easy to create B-spline curves. It is important to note that from the algorithm, in general you need to provide more knots than control points to generate a B-spline.

The 3rd basis functions $N_{i,3}(t)$ for a B-spline with knot sequence $U = [0, 1, 2, 3, 4, 5, 6, 7]$ so i ranges from 0 to 3 is given below.

The 3rd basis functions $N_{i,3}(t)$ for a B-spline with knot sequence $U = [0, 0, 0, 1, 2, 3, 4, 4, 4]$ so i ranges from 0 to 4 is given below.

22.3 B-spline curves

A B-spline curve is defined by a knot sequence (a vector) $U = [u_0, u_1, \cdots, u_m]$, a set of control points $[b_0, b_1, \cdots, b_n]$, and an order k (a real number). There is a relation between the number of knots $m+1$, the number of control points $n+1$, and the order of the curve k. The number of control points $n+1$ must equal the number of basis functions $m-k$. [The number of 0th order basis functions $N_{i,0}$ is m, and each time through the recursion there is one less basis function. There are $m-1$ 1st order basis functions, $m-2$ 2nd order basis functions.]

The knot sequence and the order are used to create the basis functions $N_{i,k}(t)$ and provide the domain of the curve. The control points then provide the shape of the curve. The advantage of B-splines is that you can alter the curve either by changing the knot-sequence or the control points. In fact, certain knot sequences are used for different types of curves.

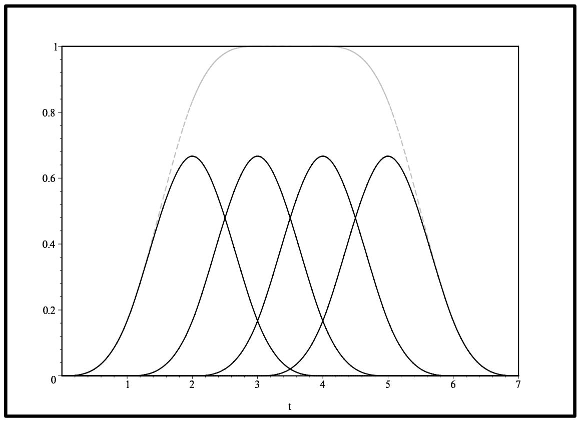

Figure 1: Basis functions for knot sequence $U = [0, 1, 2, 3, 4, 5, 6, 7]$

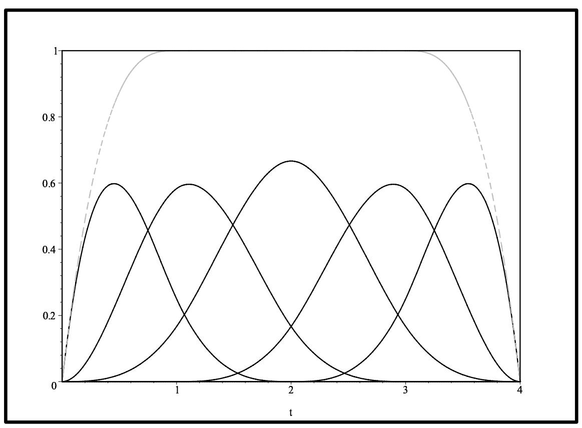

Figure 2: Basis functions for knot sequence $U = [0, 0, 0, 1, 2, 3, 4, 4, 4]$

The knot sequence

$$U = [\underbrace{a \cdots, a}_{k+1}, u_{k+1}, \cdots, u_{m-k-1}, \underbrace{b, \cdots, b}_{k+1}]$$

yields a nonperiodic knot sequence. This type of knot sequence has the property that it does interpolate the end-points. To see this, you have to show that the knot sequence defines a curve on the interval $a \leq t \leq b$, and the basis functions satisfy $N_{0,k}(a) = 1$, $N_{i,k}(a) = 0$ for all $i > 0$ and also $N_{n,k}(b) = 1$, $N_{i,k} = 0$ for all $i < n = m - k - 1$. From these properties, one has that $c(t) = \sum N_{i,k}\, p_i$ satisfies $c(a) = b_0$ and $c(b) = b_n$.

In fact, Bezier curves are a specific type of B-spline with a nonperioidic knot sequence. The knot sequence

$$U = [\underbrace{0, \cdots, 0}_{k+1}, \underbrace{1, \cdots, 1}_{p+1}]$$

defines the Bernstein polynomials of order k. A piecewise Bezier curve of L-segments is given by a knot sequence of the form

$$U = [0, \cdots, 0, 1, \cdots, 1, 2 \cdots, 2, \cdots, L, \cdots, L].$$

In fact, any Bezier spline can be derived as a B-spline with a nonperiodic knot sequence. All you have to do is determine the right type of knot sequence.

Periodic uniform knot sequence

$$U = [0, 1, 2, \cdots, m]$$

yield basis functions that are translates of each other, which means

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

These type of knot vectors are useful for designing periodic curves, i.e. circles. To design a periodic curve, one only needs to repeat the control points at the beginning and end to get smoothness properties to coincide in both the forward and backward directions of the parameter.

Essentially, a knot sequence is used to prescribe the behavior of the B-spline at the end points. They are also used to define the smoothness properties of B-splines, as they control the smoothness of the B-spline at the knot values $t = u_i$.

Some of the properties of B-splines curves are

- Affine invariance. $\mathcal{A}(c(t)) = \sum N_{i,k}\mathcal{A}(p_i)$. This is a property of the barycentric combinations to define the curve.
- Convex hull property. The curve $c(t)$ is contained in the convex hull of the control points. This is a property of barycentric combinations with nonnegative coefficients. The curve actually satisfies a stronger condition as the basis functions have *compact support* (each basis function is nonzero only on a finite interval), so each part of the curve is only effected by a finite number of control points.
- Local control property. Moving a control point p_i only alters the curve near that control point. This is another result of the compact support of the basis functions.
- The smoothness (differentiability) of a B-spline is given by the multiplicity of the interior knots. On each interval $u_i < t < u_{i+1}$ a B-spline curve is infinitely differentiable. At the knots, the curve's smoothness depends on the multiplicity of the knot and the order of the curve. Let u_i have multiplicity r ($u_{i-1} \neq u_i = u_{i+1} = \cdots = u_{i+r-1} \neq u_{i+r}$), then the curve is differentiable at least upto order $k - r$.

22.4 Relation between Number of Points and Knots

There is an important relation between the number of control points, number of knots, and the order of a B-spline curve. Given a knot sequence $U = [u_0, u_1, \cdots, u_m]$, there are $m + 1$ knots. When computing the basis functions for a curve of order k, there are $m - k$ basis functions. [There are m basis functions of order zero, $m - 1$ basis functions of order one, $m - 2$ basis functions of order two, etc.] This is a direct result of the inductive argument for creating the basis functions. To define a B-spline via the basis functions,

$$b(t) = \sum_{i=0}^{n} N_{i,k}(t)\, p_i$$

where $N_{i,k}(t)$ are the basis functions and p_i are the control points. We must have the number of control points $n+1$ equal to the number of basis functions $m-k$.

When creating B splines, curves. It is worth noting that a curve is only defined when the sum of the basis functions equals one, so we have a barycentric representation of the points on the curve from the control points. It is a book-keeping exercise (using the definition of the basis functions) to show that the sum of the basis functions equals one on the interval $u_k \leq t \leq u_{m-k}$. An important observation is that we can only define a curve of order $0 < k < m/2$.

B-spline curves of order 1 curves are just "crooked lines" or a linear interpolant of the control points (as long as no knots are repeated). B-spline curves of order 1 are thus uninteresting. The higher order curves are much more interesting. One other point about the order of the curve, an order k curve is $k-1$ times differentiable in general, meaning that an order 2 curve is in general C^1 and an order 3 curve is in general C^2, at least if no knots are repeated (other than the first and last knots, see below).

22.5 Control Point Interpolation

A B-spline will interpolate control points only if the corresponding basis function has attains the value 1. The construction of the basis functions imply that $0 \leq N_{i,k}(t) \leq 1$. A further examination of the basis functions definition shows that the maximum of the order 1 basis functions is equal to 1 as long as no knots are repeated more than once.

In general, to interpolate the first and last control points p_0 and p_n with a B-spline curve of order k, it is necessary that the first $k+1$ knots are repeated and the last $k+1$ knots are repeated. To interpolate interior control points, you need to repeat a knot k times. The other possibility is to repeat a control point, the sequence of control points can have $p_k = p_{k+1}$. If this is done, one can also force the curve to interpolate a control point.

For instance, a order 2 curve with the knot sequence $U = [0, 0, 0, 1, 2, 2, 2.5, 4, 4, 4]$ will interpolate the first control point and the last control point. The basis functions for this knot sequence are

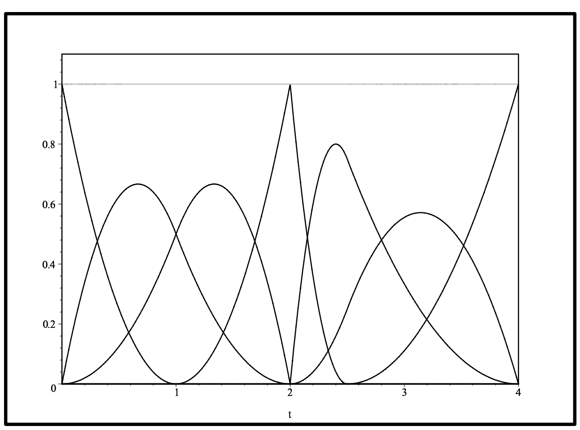

Figure 3: Basis functions for knot sequence $U = [0, 0, 0, 1, 2, 2, 2.5, 4, 4, 4]$

From the plot, one sees that there are seven basis functions and the basis functions $N_{0,2}$, $N_{3,2}$ and $N_{6,2}$ have maximum values equal to one, so that the control points p_0, p_3 and p_6 are interpolated. Below is an example of a B-spline curve generated with this knot sequence.

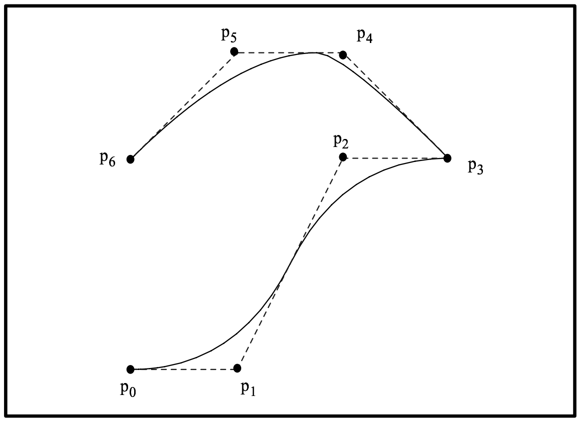

Figure 4: A B-spline using the knot sequence $U = [0, 0, 0, 1, 2, 2, 2.5, 4, 4, 4]$

Forcing a B-spline to interpolate a control point (other than the first and last control point) has an immediate effect on the smoothness of a B-spline curve. Repeating knots decreases the smoothness of the curve by one order of differentiability for each time an interior knot is repeated. For instance, the order 2 curve generated from the knot sequence $U = [0, 0, 0, 1, 2, 2, 2.5, 4, 4, 4]$ above interpolated the control point p_3 with the affect that unless the control points are arranged correctly the curve will have a corner at p_3. It will not be smooth.

It is very important to understand that the knot sequence determines the basis functions once an order has been chosen for the basis functions. Using the same knot sequence $U = [0, 0, 0, 1, 2, 2, 2.5, 4, 4, 4]$, an order 3 curve has the basis functions below so the curve is only defined on the interval $1 \leq t \leq 2.5$, and we have no control points interpolated as none of the basis functions achieves the maximum value 1.

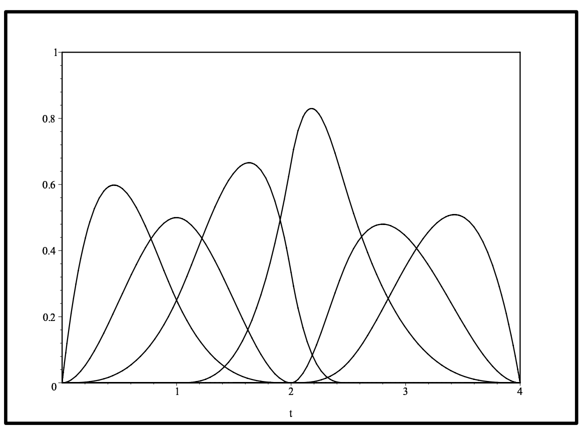

Figure 5: Cubic basis functions for the knot sequence $U = [0, 0, 0, 1, 2, 2, 2.5, 4, 4, 4]$

An example of a cubic B-Spline with the knot sequence $U = [0, 0, 0, 1, 2, 2, 2.5, 4, 4, 4]$ is given below.

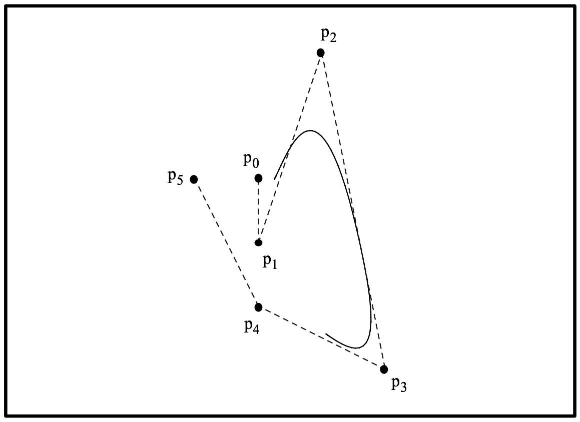

Figure 6: A B-spline using the knot sequence $U = [0, 0, 0, 1, 2, 2, 2.5, 4, 4, 4]$

22.6 Uniform Periodic Knot Sequences

First, we considered a uniform periodic knot sequence. In these knot sequences, the basis functions are all translates of a basic function. For instance, consider the knot sequence $U = [0, 1, 2, 3, 4, 5, 6, 7, 8]$ with $k = 3$. The basis functions are

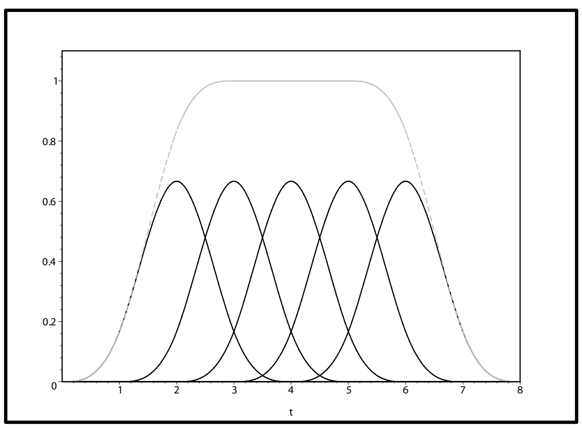

Figure 7: Basis functions for the knot sequence $U = [0, 1, 2, 3, 4, 5, 6, 7, 8]$

which are all translates of the function
A B-spline created with these basis functions is shown below.

A main use of uniform periodic knot sequences is for creating periodic curves by repeating the control points. The knot sequence $U = [0, 1, 2, 3, 4, 5, 6, 7, 8]$ with the control points $[p_0, p_1, p_2, p_3, p_0, p_1]$ creates a "smooth" curve, that appears periodic. This is very useful for creating approximations to circles and other periodic curves, see diagram below of an order 2 curve.

One can use nonuniform periodic knot sequences one just has to repeat the pattern of the knots at the beginning of the knot sequence at the end of the knot sequence. For instance, the sequence

$$U = [0, 1, 5, 7, 10, 12, 15, 15, 18, 20, 21, 25, 27, 30]$$

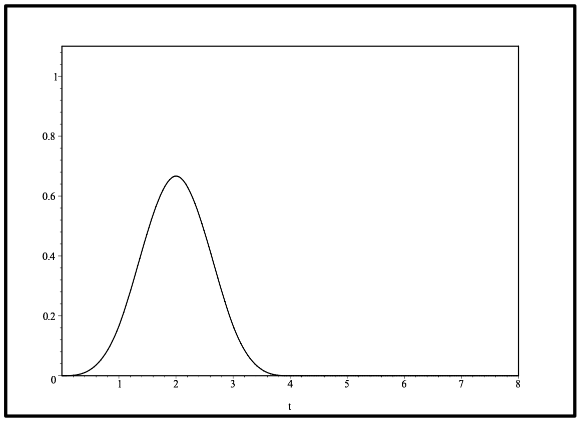

Figure 8: Base Basis function for the knot sequence $U = [0, 1, 2, 3, 4, 5, 6, 7, 8]$

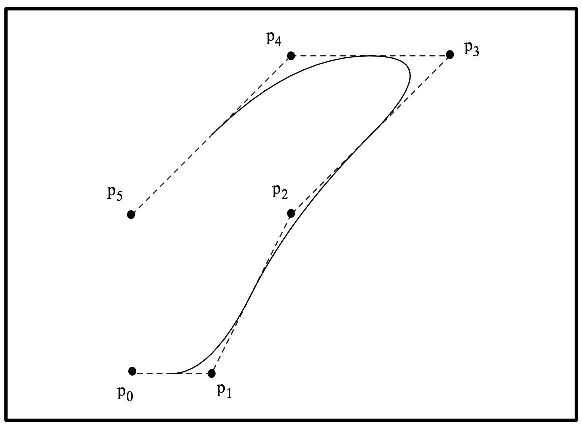

Figure 9: A B-spline using the knot sequence $U = [0, 1, 2, 3, 4, 5, 6, 7, 8]$

is a nonuniform periodic knot sequence. Applying this knot sequence to a sequence of control points $[p_0, p_1, p_2, p_3, p_4, p_5, p_6, p_7, p8, p_0, p_1]$ repeating the first two control points generates a periodic curve of order 2. The trick is to repeat enough of the knot sequence to translate k basis functions. The first k basis functions need to be repeated as the last k basis functions. For a quadratic curve, you need to repeat the spacing of the first 5 knots in the last 5 knots, see basis functions below.
An example of such a periodic curve is given below. Notice the curve passes through p_6. This is because 15 is repeated twice as a knot.

22.7 Nonperiodic Knot Sequences

The important features of nonperiodic knot sequences are described above by the interpolation properties of B-spline curves and repeating knots in the knot sequence. We only have to state what we mean by a nonperiodic knot sequence. A nonperiodic knot sequence is given by anything that is not a periodic knot sequence. The most useful nonperiodic knot sequence are given by repeating the first and last knots $k+1$ times so that the first and last control points are interpolated. Other choices of repeated knots can be used to prescribed the tangent vector but leave the endpoints of the curve free. The use of the word nonperi-

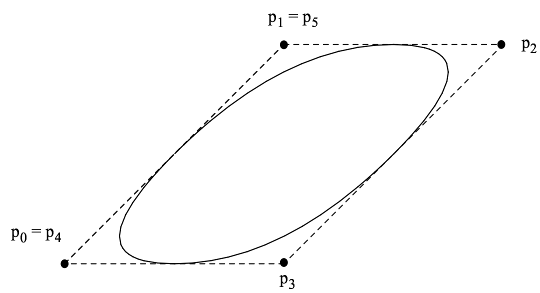

Figure 10: Periodic quadratic curve using the knot sequence $U = [0, 1, 2, 3, 4, 5, 6, 7, 8]$

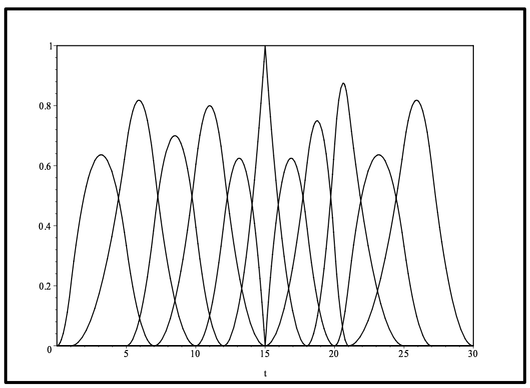

Figure 11: Quadratic basis functions for nonuniform periodic knot sequence

odic is used to represent the fact that if you repeat the last k the control points with the first k control points the curve will not appear to represent a periodic curve.

It useful to understand that one does not normally directly manipulate the knot sequence but rather specifies the type of behavior desired at certain points. For instance, there are specific choices of knot sequences to prescribe the tangent line at an endpoint of a curve and ensure the curvature is zero at this point. The knot sequence allows one to specify physical properties of the curve at a point defined as the endpoint of the curve or an interpolated point. We will not go into these properties in detail. If you are interested, consult the text by Farin CAGD.

22.8 NURBS

A NURB is a non-uniform rational B-Spline. They are rational B-Splines, meaning that they are plane curves defined by perspectively projecting a spatial B-Spline onto the plane $x = 1$. This combines the ability to control the knot sequence to achieve desired interpolation properties with the ability to control the weights of the control points in rational curves to adjust the shape of the curve. The fine control allowed by manipulating control points, knots and weights has made NURBS the industrial standard for design.

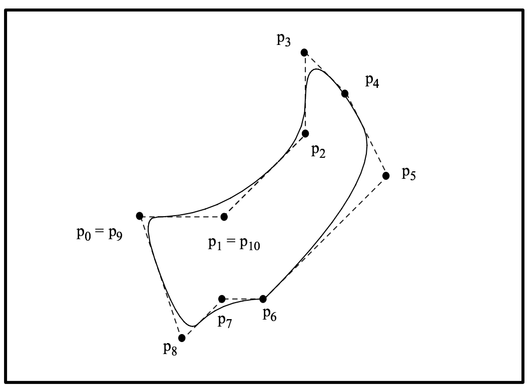

Figure 12: Periodic quadratic curve using the nonuniform periodic knot sequence