Malaysian Journal of Mathematical Sciences 3(2): 203-226 (2009)

A Class of Bézier-Like Splines in Smooth Monotone Interpolation

Ng Seou Ling, V.P. Kong and B.H. Ong
School of Mathematical Sciences,
Universiti Sains Malaysia, 11800 Penang, Malaysia
E-mail: seouling118@yahoo.com, kongvp@cs.usm.my,bhong@cs.usm.my

ABSTRACT

We develop a new family of curves, the quartic and quintic Bézier-like curves, and investigate the use of these curves in smooth monotonicity preserving interpolation. These polynomial curves have, besides four control points, two additional parameters for shape control. The effect of the two parameters on the curve is analysed. Conditions on these parameters for the quartic and quintic Bézier-like curves to be monotonic are derived. Based on these conditions, a local C^1 monotonicity preserving quartic Bézier-like spline interpolation scheme is presented. A C^2 monotonicity preserving interpolation scheme is also developed where the optimal quintic Bézier-like spline interpolant is chosen through a constrained minimization of its mean curvature.

Keywords: *Bézier-like spline; Monotonicity preverving;* C^1 *monotone interpolation;* C^2 *monotone interpolation*

INTRODUCTION

In many areas of science and engineering, it is often necessary to obtain a continuous mathematical representation of a finite set of discrete data. It is well known that standard techniques for interpolation are often incapable of reproducing the shape of the data like monotonicity, convexity or positivity and this may destroy the physical interpretation of the phenomenon or the idea of the designer. This has led to considerable interest in shape preserving interpolation problem.

The problem of monotonicity preserving interpolation has been considered by a number of authors. Fritsch and Carlson, (1980) use the representation of Hermite cubic polynomial to derive the necessary and sufficient conditions for a cubic to be monotone on an interval. Delbourgo and Gregory, (1985) develop an explicit representation of a C^1 piecewise rational cubic function which can be used to solve the problem of shape preserving interpolation. Heß and Schmidt, (1994) construct a monotonicity preserving interpolation of discrete data by quintic polynomial splines. They show that monotonicity can be always preserved by quintic C^2-splines.

In (Jamaludin *et al.*, 1996) the authors introduced cubic Bézier-like curves. Based on the cubic Bézier-like basis functions, we develop in this paper a new family of curves, the quartic and quintic Bézier-like curves. These polynomial curves are derived via a linear or quadratic convex combination of two or three cubic Bézier-like curves which have the same control points. With the representation of these curves in the Hermite form, their parameters are grouped in a natural way as ratios thus resulting in two new parameters which are referred to as α and β. The effect of α and β on the quartic and quintic curves is analyzed. They provide local control on the shape of the curve. We present the use of these Bézier-like curves in smooth monotonicity preserving interpolation. For the piecewise quintic curves, the second order continuity at the end points of the adjacent segments of the curves can be attained easily by using these two parameters independently from one another. We derive the monotonicity conditions on these parameters of the quartic Bézier-like curve. Based on these conditions, we construct a local C^1 monotonicity preserving curve interpolation scheme using the quartic Bézier-like splines. We also derive monotonicity conditions for the quintic Bézier-like curve and a C^2 monotonicity preserving curve interpolation scheme is developed where the optimal quintic Bézier-like spline interpolant is chosen through a quadratic programming.

CUBIC BÉZIER-LIKE CURVE AND ITS PROPERTIES

The representation of a cubic Bézier-like curve with parameters $a, b \in \mathbb{R}$ and control points $V_i, i = 0, 1, 2, 3$, $V_i \in \mathbb{R}^n$, n is a positive integer, is defined in (Jamaludin *et al.*, 1996) to be

$$r(t; a,b) = \sum_{i=0}^{3} F_i(t; a,b)\, V_i\,, \qquad 0 \le t \le 1 \tag{1}$$

where the F_i are cubic basis functions with

$$F_0(t; a,b) = (1-t)^2 \left(1 + t(2-a)\right), \quad F_1(t; a,b) = a(1-t)^2\, t\,,$$
$$F_2(t; a,b) = b\,(1-t)\, t^2\,, \quad F_3(t; a,b) = t^2 \left(1 + (1-t)(2-b)\right), \qquad a, b \in \mathbb{R}\,.$$

Note that the basis functions $F_i(t)$ are controlled by two parameters a and b. This property gives us the convenience to change the shape of the curve without changing the control points. Some useful properties of this basis functions are as follows:

(i) *Positivity*
If $0 \le a,\ b \le 3$, then

$$F_i(t;a,b) \ge 0,\ \ 0 \le t \le 1.$$

(ii) *Partition of unity*

$$\sum_{i=0}^{3} F_i(t;a,b) = 1, \qquad 0 \le t \le 1.$$

The piecewise cubic Bézier-like curve r in [1] with control points V_i has the following interpolatory properties at the endpoints:

$$r(0) = V_0, \qquad r(1) = V_3,$$
$$\frac{d}{dt} r(0) = a\,(V_1 - V_0), \qquad \frac{d}{dt} r(1) = b\,(V_3 - V_2).$$

For $n = 2$ or 3, every point on the curve in [1] lies in the convex hull of the control polygon if $0 \le a,\ b \le 3$. This is a consequence of the above properties (i) and (ii) of the basis functions.

Quartic Bézier-like Curve

Let $r_1(t\,;\,a,\,b)$ and $r_2(t\,;\,p,\,q)$ be two pieces of cubic Bézier-like curves with the same control ordinates $(V_i \in \mathbb{R}: 0 \le i \le 3)$ but their parameters may be different, i.e.

$$r_1(t\,;\,a,\,b) = \sum_{i=0}^{3} F_i(t\,;\,a,\,b)\,V_i\,,$$
$$r_2(t\,;\,p,\,q) = \sum_{i=0}^{3} F_i(t\,;\,p,\,q)\,V_i\,, \quad 0 \le t \le 1,$$

where $a,\ q > 0$ and $b,\ p \in \mathbb{R}$. Then we define $Q(t\,;\,a\,,b\,,\,p\,,\,q)$ which is abbreviated as $Q(t)$ by

$$Q(t) = (1-t)\ r_1(t\,;\,a\,,b) + t\ r_2(t\,;\,p\,,q)\,,\ \ 0 \le t \le 1.$$

The quartic Bézier-like curve $Q(t)$ has the interpolatory properties at the endpoints similar to those of the cubic Bézier-like curve in [1]:

$$Q(0) = V_0, \quad Q(1) = V_3,$$
$$\frac{d}{dt}Q(0) = a(V_1 - V_0), \quad \frac{d}{dt}Q(1) = q(V_3 - V_2). \qquad [2]$$

If m_0 and m_1 are the first order derivatives of Q at 0 and 1 respectively, then by [2] and putting $u = (1-t)$, we obtain

$$Q(t) = V_0 u^4 + 4u^3 t\left(V_0 + \frac{m_0}{4}\right) + 6u^2 t^2\left(\frac{V_3 + V_0}{2} + \frac{p}{6a}m_0 - \frac{b}{6q}m_1\right)$$
$$+ 4ut^3\left(V_3 - \frac{m_1}{4}\right) + V_3 t^4$$

which clearly suggests that the parameters a, b, p, q can be grouped as ratios $\alpha = p/a$, $\beta = b/q$. Thus $Q(t)$ and its first order derivative can be represented as follows:

$$Q(t;\alpha,\beta) = V_0 u^4 + 4u^3 t\left(V_0 + \frac{m_0}{4}\right) + 6u^2 t^2\left(\frac{V_3 + V_0}{2} + \frac{1}{6}(\alpha m_0 - \beta m_1)\right)$$
$$+ 4ut^3\left(V_3 - \frac{m_1}{4}\right) + V_3 t^4$$

[3]

$$\frac{d}{dt}Q(t;\alpha,\beta) = m_0 u^3 + 3u^2 t\left(2\Delta - m_0 + \frac{2}{3}(\alpha m_0 - \beta m_1)\right)$$
$$+ 3ut^2\left(2\Delta - m_1 + \frac{2}{3}(\beta m_1 - \alpha m_0)\right) + m_1 t^3$$

[4]

where $\Delta = V_3 - V_0$.

Quintic Bézier-like Curve

Suppose $r_1(t;a,b)$, $r_2(t;p,q)$ and $r_3(t;f,g)$ are three pieces of cubic Bézier-like function curves with the same control ordinates $(V_i \in \mathbb{R}: 0 \le i \le 3)$ but their parameters are different, i.e.

$$r_1(t\,;a,b)=\sum_{i=0}^{3} F_i(t\,;a,b)V_i\,,$$

$$r_2(t\,;p,q)=\sum_{i=0}^{3} F_i(t\,;p,q)V_i\,,$$

$$r_3(t\,;f,g)=\sum_{i=0}^{3} F_i(t\,;f,g)V_i\,,\quad 0\le t\le 1,$$

where $a, g > 0$ and $b, p, q, f \in \mathbb{R}$.

When these three curves are combined as a convex combination with quadratic polynomial coefficients, we obtain a quintic polynomial curve $R(t\,;a,b,p,q,f,g)$ on $t\in[0,1]$ which is abbreviated as $R(t)$ and defined by

$$R(t)=(1-t)^2 r_1(t\,;a,b)+2t(1-t)r_2(t\,;p,q)+t^2 r_3(t\,;f,g)\,. \qquad [5]$$

The quintic Bézier-like curve $R(t)$ has the interpolatory properties at the endpoints similar to those of the cubic Bézier-like curve in [1]:

$$R(0) = V_0, \qquad R(1) = V_3,$$

$$\frac{d}{dt}R(0) = a(V_1 - V_0), \quad \frac{d}{dt}R(1) = g(V_3 - V_2). \qquad [6]$$

As we have a number of parameters at our disposal, we let $p = a/2$ and $p = q/2$ to obtain a simpler representation for $R(t)$. If in addition, m_0 and m_1 are the first order derivatives of $R(t)$ at 0 and 1 respectively, then by [5] and [6] with $u=(1-t)$, $\Delta = V_3 - V_0$ and grouping the parameters as ratios $\alpha = f/a$, $\beta = b/g$, then $R(t)$ can be represented as

$$R(t;\alpha,\beta)=V_0u^5+(m_0+5V_0)u^4t+(m_0-\beta m_1+7V_0+3V_3)u^3t^2$$
$$+(\alpha m_0-m_1+3V_0+7V_3)u^2t^3+(-m_1+5V_3)ut^4+V_3t^5. \qquad [7]$$

Its derivatives are

$$\frac{d}{dt}R(t;\alpha,\beta)=m_0u^4+2(3\Delta-m_0-\beta m_1)u^3t+3(-m_0+\alpha m_0+4\Delta$$
$$+\beta m_1-m_1)u^2t^2+2(3\Delta-m_1-\alpha m_0)+m_1t^4 \quad [8]$$

and

$$\frac{d^2}{dt^2}R(t;\alpha,\beta)=2(3\Delta-\beta m_1-3m_0)u^3+6(\Delta+\alpha m_0-m_1+2\beta m_1)u^2t$$
$$+6(-\Delta+m_0-2\alpha m_0-\beta m_1)ut^2+2(-3\Delta+\alpha m_0+3m_1)t^3.$$

In particular, $\frac{d^2}{dt^2}R(0;\alpha,\beta)=2(3\Delta-\beta m_1-3m_0)$ and $\frac{d^2}{dt^2}R(1;\alpha,\beta)$ $=2(-3\Delta+\alpha m_0+3m_1)$. Observe that each of these two endpoint derivatives only depends on one of the two parameters, α or β, and this makes it very easy to achieve C^2 continuity between adjacent segments.

Effect of the Parameters α and β on the Quartic and Quintic Curves

We examine the effect of the parameters α and β on the quartic Bézier-like curve and quintic Bézier-like curve. We are only interested in the case where $\alpha\ge 0$ and $\beta\ge 0$. Differentiating [3] and [7] partially with respect to the parameters α and β,

$$\frac{\partial Q}{\partial\alpha}=m_0u^2t^2, \qquad \frac{\partial R}{\partial\alpha}=m_0u^2t^3, \quad [9]$$
$$\frac{\partial Q}{\partial\beta}=-m_1u^2t^2, \qquad \frac{\partial R}{\partial\beta}=-m_1u^3t^2. \quad [10]$$

By [9], we notice that for a fixed $t\in(0,1)$, the effect of α depends on m_0. If m_0 is positive, then $\frac{\partial Q}{\partial\alpha}$ and $\frac{\partial R}{\partial\alpha}$ are also positive and so for any fixed $t\in(0,1)$, $Q(t;\alpha,\beta)$ and $R(t;\alpha,\beta)$ will increase as α is increased. If m_0 is negative, then $Q(t;\alpha,\beta)$ and $R(t;\alpha,\beta)$ will decrease as α increases at any fixed $t\in(0,1)$. Thus the set $\{Q(t;\alpha,\beta),t\in[0,1]:\alpha\ge 0\}$ forms a nested family of curves, i.e. the curves in this family do not intersect one another except at the end points $t=0$ and $t=1$. Similarly $\{R(t;\alpha,\beta), t\in[0,1]:\alpha\geq 0\}$ forms a nested family of curves (see Figures 1 and 4).

The effect of the parameter β depends on m_1. It is clear from [10] that, for any fixed $t\in(0,1)$, if m_1 is negative, then $Q(t;\alpha,\beta)$ and $R(t;\alpha,\beta)$ increase when β increases. If m_1 is positive, then $Q(t;\alpha,\beta)$ and $R(t;\alpha,\beta)$ will decrease as β increases at any fixed $t\in(0,1)$. Thus $\{Q(t;\alpha,\beta), t\in[0,1]:\beta\geq 0\}$ and $\{R(t;\alpha,\beta), t\in[0,1]:\beta\geq 0\}$ are also nested families of curves (see Figures 2 and 5).

Letting $\alpha=\beta=\mu$ and partially differentiating Q and R with respect to μ, we have

$$\frac{\partial Q}{\partial \mu}=(m_0-m_1)u^2t^2, \qquad [11]$$

$$\frac{\partial R}{\partial \mu}=u^2t^2g(t),$$

where $g(t)=tm_0+(1-t)(-m_1)$.

From [11], for any fixed $t\in(0,1)$, as μ is increased, $Q(t;\alpha,\beta)$ increases if $m_0-m_1>0$ and decreases if $m_0-m_1<0$ (see Figure 3). On the other hand, the sign of $\frac{\partial R}{\partial \mu}$ depends directly on the sign of $g(t)$ which is a linear combination of m_0 and m_1. If $m_0>0$ and $-m_1>0$, then for any fixed $t\in(0,1)$, $g(t)$ is positive. Thus, for any fixed $t\in(0,1)$, $R(t;\mu,\mu)$ increases when μ is increased. If $m_0<0$ and $-m_1<0$, then $g(t)<0$ and so for any fixed $t\in(0,1)$, $R(t;\mu,\mu)$ decreases when μ increases. If $m_0(-m_1)<0$, $g(t)$ will change sign at $t=\frac{m_1}{m_0+m_1}$. Thus, when μ is increased, $R(t;\mu,\mu)$ increases on the interval $\left(0,\frac{m_1}{m_0+m_1}\right)$ and decreases on $\left(\frac{m_1}{m_0+m_1},1\right)$ or vice versa (see Figure 6).

Below are some graphical examples that illustrate the effect of the parameters α, β and μ on the quartic Bézier-like curve $Q(t;\alpha,\beta)$ and quintic Bézier-like curve $R(t;\alpha,\beta)$. In the Figures 1-6, the given data points $P_0, P_1 \in \mathbb{R}^2$ are joined with a dashed line and the given derivatives at the endpoints, $t=0$ and $t=1$ are respectively $m_0 = 3$ and $m_1 = 4$. The curves with ascending parameter values are respectively marked as (i), (ii), (iii), (iv) and (v). The following different values are used for the varying parameter in each of the figures, namely 0, 1, 3, 8 and 20.

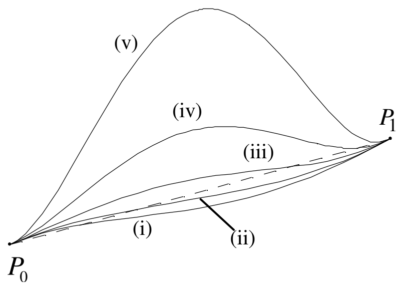

Figure 1: Quartic Bézier-like curves with different parameter values $\alpha = 0, 1, 3, 8$ or 20 while $\beta = 1$.

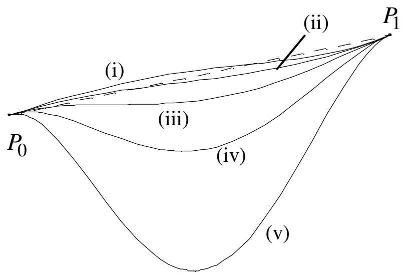

Figure 2: Quartic Bézier-like curves with different parameter values $\beta = 0, 1, 3, 8$ or 20 while $\alpha = 1$.

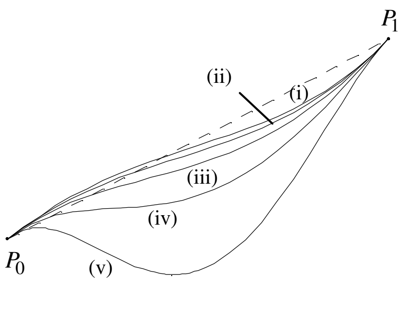

Figure 3: Quartic Bézier-like curves with different parameter values $\mu = 0, 1, 3, 8$ or 20.

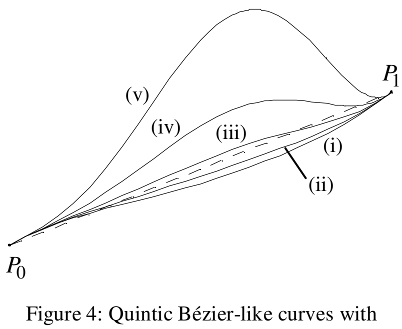

Figure 4: Quintic Bézier-like curves with different parameter values $\alpha = 0, 1, 3, 8$ or 20 while $\beta = 1$.

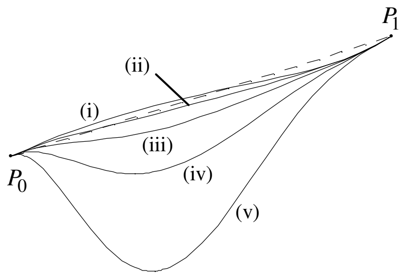

Figure 5: Quintic Bézier-like curves with different parameter values $\beta = 0, 1, 3, 8$ or 20 while $\alpha = 1$.

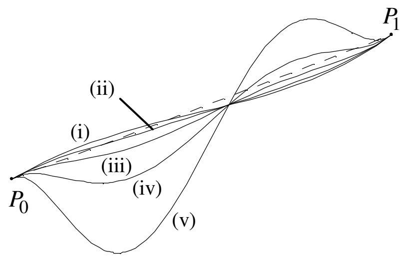

Figure 6: Quintic Bézier-like curves with different parameter values $\mu = 0, 1, 3, 8$ or 20.

C^1 MONOTONICITY PRESERVING INTERPOLATION WITH QUARTIC BÉZIER-LIKE SPLINE

In this section, we derive the monotonicity conditions in terms of the parameters α and β of the quartic Bézier-like curve. By using the monotonicity conditions, we construct a C^1 monotonic quartic Bézier-like spline interpolating scheme.

Let $\{ P_i = (x_i, y_i) : 0 \le i \le n \}$ be a set of monotonic increasing data so that $x_0 < x_1 < \cdots < x_n$ and $y_0 \le y_1 \le \cdots \le y_n$. The derivatives d_i at x_i are estimated by using the Fritsch and Butland's method (1984), i.e. for $1 \le i \le n-1$,

$$d_i = \begin{cases} \dfrac{2\Delta_{i-1}\Delta_i}{\Delta_{i-1} + \Delta_i}, & \text{if } \Delta_{i-1}\Delta_i > 0 \\ 0, & \text{otherwise} \end{cases}$$

where $\Delta_i = \dfrac{y_{i+1} - y_i}{h_i}$ and $h_i = x_{i+1} - x_i$, $0 \le i \le n-1$. The derivative d_0 at the first data point may be defined as

$$d_0 = \min\left\{ \max\{0, d_0^*\}, 2\Delta_0 \right\}$$

where $d_0^* = \begin{cases} \left(1+\frac{\Delta_0}{\Delta_1}\right)\Delta_0 - \frac{\Delta_0}{\Delta_1}\frac{y_2-y_0}{x_2-x_0}, & \text{if } \Delta_1 \neq 0 \\ 0 & \text{otherwise} \end{cases}$.

The derivative d_n is defined similarly. Clearly we have $0 \le d_i, d_{i+1} \le 2\Delta_i$, $0 \le i \le n-1$.

Monotonicity Preserving Conditions

Let us consider the quartic Bézier-like curve $Q(t;\alpha,\beta)$ in [3]. For $\frac{dQ}{dt} \ge 0$ on $[0,1]$, it is necessary that the derivatives m_0 and m_1 are non-negative. By the method used in the estimation of derivatives, we have

$$0 \le m_0, m_1 \le 2\Delta$$

where $\Delta = V_3 - V_0 > 0$.

Since $u, t \ge 0$ for $t \in [0,1]$, it follows that if the coefficients of the cubic Bernstein basis functions in [4] are non-negative, then $\frac{d}{dt}Q(t;\alpha,\beta) \ge 0$. Thus, the sufficient conditions for $Q(t;\alpha,\beta)$ to be monotonic on $[0,1]$ are

$$A = m_0 \ge 0,$$

$$B = 2\Delta - m_0 + \frac{2}{3}(\alpha m_0 - \beta m_1) \ge 0, \qquad [12]$$

$$C = 2\Delta - m_1 + \frac{2}{3}(\beta m_1 - \alpha m_0) \ge 0, \qquad [13]$$

$$D = m_1 \ge 0.$$

When the above sufficient conditions are not met, there are four cases to be considered.

Case 1: $A > 0, D > 0, B < 0$ and/or $C < 0$

The following theorem quoted from (Goodman *et al.*, 1991) gives us the conditions for $\frac{dQ}{dt} \ge 0$ on $[0,1]$.

Theorem 1: Let

$$P(t) = \overline{A}(1-t)^3 + 3\overline{B}t(1-t)^2 + 3\overline{C}t^2(1-t) + \overline{D}t^3, \qquad 0 \le t \le 1,$$

where $\overline{A}, \overline{D} > 0$, and $\overline{B} < 0$ and/or $\overline{C} < 0$.

Then $P(t) > 0$ for $t \in [0,1]$ [resp. $P(t) = 0$ for only one point in $(0,1)$] if and only if

$$3\overline{B}^2\,\overline{C}^2 + 6\overline{A}\,\overline{B}\,\overline{C}\,\overline{D} - 4\left(\overline{A}\,\overline{C}^3 + \overline{B}^3\,\overline{D}\right) - \overline{A}^2\,\overline{D}^2 < 0 \quad \text{[resp. = 0]}. \qquad [14]$$

Let us denote $\Phi(\overline{A}, \overline{B}, \overline{C}, \overline{D}) = 3\overline{B}^2\,\overline{C}^2 + 6\overline{A}\,\overline{B}\,\overline{C}\,\overline{D} - 4(\overline{A}\,\overline{C}^3 + \overline{B}^3\,\overline{D}) - \overline{A}^2\,\overline{D}^2$. Observe that $\frac{d}{dt}Q(t;\alpha,\beta)$ in [4], has the same form as the cubic polynomial in the theorem above. In this case, $A > 0, D > 0, B < 0$ and/or $C < 0$. We can check whether the curve segment Q is monotonically increasing by using the condition in [14]. If $\Phi(A, B, C, D) < 0$, then we can use the default values for α and β to construct a monotonic curve, else we scale both α and β by using a scalar factor, $\lambda \in [0, 1)$, so that the new values of B and C which are respectively

$$B^* = 2\Delta - m_0 + \frac{2}{3}\lambda\,(\alpha m_0 - \beta m_1) \qquad \text{and}$$
$$C^* = 2\Delta - m_1 + \frac{2}{3}\lambda(\beta m_1 - \alpha m_0)$$

satisfy $\Phi(A, B^*, C^*, D) = 0$.

Case 2: $A = 0$ and $D > 0$

By substituting $A = 0$ (i.e. $m_0 = 0$) in [12] and [13], we obtain

$$B = 2\Delta - \frac{2}{3}\beta m_1, \qquad C = 2\Delta - m_1 + \frac{2}{3}\beta m_1.$$

Since $0 \le m_1 \le 2\Delta$, for any non-negative β, we have $C \ge 0$. For $\frac{dQ}{dt}$ to be non-negative, it is necessary that $B \ge 0$ since $A = 0$. We need to scale the β by a scalar factor $\lambda \in [0, 1)$ so that

$$B = 2\Delta - \frac{2}{3}\lambda\beta m_1 \ge 0.$$

Case 3: $A > 0$ and $D = 0$

This case is similar to case (ii) above. We just need to scale α.

Case 4: $A = 0$ and $D = 0$

Since the two endpoint derivatives are zero, $A = m_0 = 0$ and $D = m_1 = 0$, $B = 2\Delta \ge 0$ and $C = 2\Delta \ge 0$. Thus $\frac{dQ}{dt} \ge 0$ for any $\alpha, \beta \ge 0$.

C^1 Monotonicity Preserving Interpolation

Let $\{P_i = (x_i, y_i) : 0 \le i \le n\}$ be a set of monotonic increasing data so that $x_0 < x_1 < \cdots < x_n$ and $y_0 \le y_1 \le \cdots \le y_n$. (The case of a monotonic decreasing set of data can be treated in a similar manner). We shall now construct a C^1 monotonic quartic Bézier-like spline interpolating curve. The derivatives d_i at x_i are estimated by using Fritsch and Butland's method as described earlier. Between each consecutive two data points (x_i, y_i) and (x_{i+1}, y_{i+1}), if $y_i = y_{i+1}$, then a straight line segment is constructed joining the two data points. Otherwise a piece of quartic Bézier-like curve $Q_i(t;\alpha_i, \beta_i)$, $i = 0, \ldots, n-1$ of the form in [3] is constructed on $[x_i, x_{i+1}]$ interpolating the data values y_i, y_{i+1} and the derivatives d_i, d_{i+1} at the endpoints, namely

$$\begin{aligned} Q_i(t;\alpha_i,\beta_i) = y_i u^4 + 4u^3 t\left(y_i + \frac{h_i d_i}{4}\right) + 6u^2 t^2\left(\frac{y_i + y_{i+1}}{2} + \frac{1}{6}\alpha_i h_i d_i\right. \\ \left. - \frac{1}{6}\beta_i h_i d_{i+1}\right) + 4ut^3\left(y_{i+1} - \frac{h_i d_{i+1}}{4}\right) + y_{i+1} t^4. \end{aligned} \quad [15]$$

The construction of any curve segment is done locally. On each interval $[x_i\,, x_{i+1}]$, we determine the values for α_i and β_i to construct a C^1 monotonic quartic Bézier-like curve segment $Q_i(t;\alpha_i\,,\beta_i)$ in [15]. Let the initial values of α_i and β_i be 3. Other non-negative values may be used for α_i and β_i, but $\alpha_i = \beta_i = 3$ are chosen as we have observed that this choice gives a visually pleasing curve in general. If these default values of α_i and β_i give a monotonic curve segment, then the curve $Q_i(t;\alpha_i\,,\beta_i)$ is fixed otherwise the values of α_i and β_i will be determined by scaling the initial values with some suitable factor $\lambda_i \in [0, 1)$ as described in section titled Monotonicity Preserving Conditions. The other curve segments are generated analogously.

C^2 MONOTONICITY PRESERVING INTERPOLATION WITH QUINTIC BÉZIER-LIKE SPLINE

Consider a strictly increasing set of data $\{P_i = (x_i\,, y_i) : 0 \le i \le n\}$ so that $x_0 < x_1 < \cdots < x_n$ and $y_0 < y_1 < \cdots < y_n$, $h_i = x_{i+1} - x_i$, $i = 0, \cdots, n-1$. (The case of a strictly decreasing set of data can be treated in a similar manner). We will derive the C^2 and monotonicity conditions for the quintic Bézier-like spline to produce a C^2 monotonically increasing quintic Bézier-like spline curve for the monotonic data.

The derivatives d_i at x_i are estimated by using Fritsch and Butland's method. Between each consecutive two data points $(x_i\,, y_i)$ and $(x_{i+1}\,, y_{i+1})$, a piece of quintic Bézier-like curve $R_i(t;\alpha_i\,,\beta_i)$ of the form [7],

$$\begin{aligned} R_i(t;\alpha_i\,,\beta_i) = {} & y_i u^5 + (h_i d_i + 5 y_i) u^4 t + (h_i d_i - \beta_i h_i d_{i+1} + 7 y_i + 3 y_{i+1}) u^3 t^2 \\ & + (\alpha_i h_i d_i - h_i d_{i+1} + 3 y_i + 7 y_{i+1}) u^2 t^3 \\ & + (-h_i d_{i+1} + 5 y_{i+1}) u t^4 + y_{i+1} t^5, \qquad t \in [0,1], \end{aligned} \qquad [16]$$

is constructed on $[x_i, x_{i+1}]$ interpolating the data values $y_i\,, y_{i+1}$ and the derivatives $d_i\,, d_{i+1}$ at the endpoints.

C^2 Condition

Consider the second order derivatives at the joints of two adjacent curve segments,

$$\frac{d^2}{dt^2} R_i\,(t\,;\alpha_i\,,\beta_i\,) = 2(3\Delta_i\,h_i - \beta_i\,h_i\,d_{i+1} - 3h_i\,d_i)u^3$$
$$+6(\Delta_i\,h_i + \alpha_i\,h_i\,d_i - h_i\,d_{i+1} + 2\beta_i\,h_i\,d_{i+1})u^2\,t$$
$$+6(-\Delta_i\,h_i + h_i\,d_i - 2\alpha_i\,h_i\,d_i - \beta_i\,h_i\,d_{i+1})u\,t^2$$
$$+2(-3\Delta_i\,h_i + \alpha_i\,h_i\,d_i + 3h_i\,d_{i+1})t^3$$

where $\Delta_i = (y_{i+1} - y_i)/h_i$. For $1 \le i \le n-1$, the C^2 continuity requirement at x_i is

$$\frac{1}{(h_{i-1})^2}\frac{d}{dt^2}R_{i-1}(1;\alpha_{i-1},\beta_{i-1}) = \frac{1}{(h_i)^2}\frac{d}{dt^2}R_i(0,\alpha_i,\beta_i).$$

So we require for $1 \le i \le n-1$,

$$3\Delta_{i-1}\,h_i \;+\; 3\Delta_i\,h_{i-1} \;-\; 3d_i\left(h_{i-1} + h_i\right) \;=\; \alpha_{i-1}\,h_i\,d_{i-1} + \beta_i\,h_{i-1}\,d_{i+1}. \qquad [17]$$

We now show that it is always possible to find non-negative values for α_{i-1} and β_i, $i = 1, 2, 3, \ldots, n-1$, such that [17] is satisfied. For simplicity, let us first consider $\alpha_{i-1} = \beta_i = k_i$. Then from [17], we have

$$k_i = \frac{3[\,\Delta_{i-1}\,h_i + \Delta_i\,h_{i-1} - d_i\,(h_{i-1} + h_i\,)\,]}{h_i\,d_{i-1} + h_{i-1}\,d_{i+1}}, \qquad i = 1, \ldots, n-1. \qquad [18]$$

However, k_i obtained from [18] may be negative. The value of k_i can be ensured to be non-negative by decreasing the estimated value of d_i. Therefore we need to consider two cases.

Case 1: $3\Delta_{i-1} h_i + 3\Delta_i h_{i-1} - 3d_i \left(h_{i-1} + h_i \right) \geq 0$

In this case, k_i in [18] is non-negative. We retain the value of d_i which was estimated by the Fritsch and Butland's method.

Case 2: $3\Delta_{i-1} h_i + 3\Delta_i h_{i-1} - 3d_i \left(h_{i-1} + h_i \right) < 0$

In this case, the derivative d_i, which is $\dfrac{2\,\Delta_{i-1}\,\Delta_i}{\Delta_{i-1} + \Delta_i}$, obtained by Fritsch and Butland's method is too big. Thus we scale the derivative d_i by redefining it as:

$$d_i = \omega_i \, \frac{\Delta_{i-1}\,\Delta_i}{\Delta_{i-1} + \Delta_i}$$

where

$$\omega_i = \left(\frac{\Delta_{i-1}\, h_i + \Delta_i\, h_{i-1}}{h_{i-1} + h_i} \right) \left(\frac{\Delta_{i-1} + \Delta_i}{\Delta_{i-1}\,\Delta_i} \right). \qquad [19]$$

Observe that [17] is then satisfied with $k_i = 0$.

The value k_i in [18] for case 1 or $k_i = 0$ for case 2 may be used as the initial value of α_{i-1} and β_i, but they may be not good enough to produce a visually pleasing C^2 continuous curve. So these initial values may be improved through a constrained optimization process with an objective function that reflects the desired 'fairness' on the curve.

We minimize the mean curvature of the interpolating quintic which is

$$\int_0^1 \left(\frac{d^2}{dt^2} R_i(t;\alpha_i,\beta_i) \right)^2 dt$$
$$= \frac{4h_i^2}{35} \left[\begin{array}{l} 105\Delta_i^2 + (42 - 7\alpha_i + 3\alpha_i^2) d_i^2 + (42 - 7\beta_i + 3\beta_i^2) d_{i+1}^2 \\ \qquad + d_i \left(d_{i+1} (7\alpha_i - \alpha_i \beta_i + 7\beta_i + 21) - 105\Delta_i \right) - 105\Delta_i\, d_{i+1} \end{array} \right].$$

Thus the optimization problem is formulated as:

$$\text{minimize} \sum_{i=0}^{n-1}\left[\int_0^1 \left(\frac{d^2}{dt^2}R_i(t;\alpha_i,\beta_i)\right)^2 dt\right] \qquad [20]$$

subject to the C^2 constraints [17] and $\alpha_i \geq 0$, $\beta_i \geq 0$ for $0 \leq i \leq n-1$.

The variables in the optimization problem are the α_i's and β_i's only. So our objective function is a quadratic polynomial subject to the C^2 conditions in [17] which are linear constraints in α_i and β_i. This is a quadratic programming which can be solved uniquely and effectively. The initial values for α_i and β_i obtained as described above are used with the constrained optimizer 'fmincon' in MATLAB (Optimization Toolbox, 2000) to solve the minimization problem.

Monotonicity Preserving Conditions

We next derive the monotonicity preserving conditions for a piece of quintic curve of the form [7] on $[0,1]$. Consider $\frac{d}{dt}R(t;\alpha,\beta)$ in [8]. Since $u,t \geq 0$ for $t \in [0,1]$, it follows that if the coefficients of the quartic Bernstein basis functions are non-negative, then $\frac{d}{dt}R(t;\alpha,\beta) \geq 0$. Thus, the sufficient conditions for $R(t;\alpha,\beta)$ to be monotonic on $[0,1]$ can be formulated as follows:

$$\begin{aligned}
&m_0 \geq 0, \\
&m_1 \geq 0, \\
&3\Delta - m_0 - \beta m_1 \geq 0, \\
&3\Delta - m_1 - \alpha m_0 \geq 0, \\
&-m_0 + \alpha m_0 + 4\Delta + \beta m_1 - m_1 \geq 0
\end{aligned} \qquad [21]$$

where $\Delta = V_3 - V_0$.

C^2 Monotonicity Preserving Interpolation

With the derivatives $d_i, d_{i+1} \ge 0$, the sufficient conditions for the monotonicity of $R_i(t; \alpha_i, \beta_i)$ in [16] on $[x_i, x_{i+1}]$, $i = 0, \ldots, n-1$ from [21] are:

$$3\Delta_i - d_i - \beta_i d_{i+1} \ \ge\ 0, \qquad [22]$$
$$3\Delta_i - d_{i+1} - \alpha_i d_i \ \ge\ 0, \qquad [23]$$
$$-d_i + \alpha_i d_i + 4\Delta_i + \beta_i d_{i+1} - d_{i+1} \ \ge\ 0. \qquad [24]$$

As $d_i \le 2\Delta_{i-1}$ and $d_i \le 2\Delta_i$, we observe that [24] is automatically true. Moreover,

$$3\Delta_i - d_i \ \ge\ 0 \quad \text{and} \quad 3\Delta_i - d_{i+1} \ge\ 0.$$

So for $d_i, d_{i+1} > 0$, [22] and [23] can be achieved by taking the value of β_i and α_i sufficiently small, namely $0 \le \beta_i \le \beta_i^*$ and $0 \le \alpha_i \le \alpha_i^*$ where

$$\beta_i^* = \frac{3\Delta_i - d_i}{d_{i+1}} \ge 0, \qquad [25]$$
$$\alpha_i^* = \frac{3\Delta_i - d_{i+1}}{d_i} \ge 0. \qquad [26]$$

In the remainder of this section we shall restrict our attention to two adjacent curve segments, i.e. $R_{i-1}(t; \alpha_{i-1}, \beta_{i-1})$ on $[x_{i-1}, x_i]$ and $R_i(t; \alpha_i, \beta_i)$ on $[x_i, x_{i+1}]$ with $d_j > 0$, $i-1 \le j \le i+1$. We shall show that α_j and β_j can be determined so that $R_j(t; \alpha_j, \beta_j)$, $j = i-1, i$ are monotonically increasing and the composite curve is C^2 at x_i.

For clarity, let us write concurrently the sufficient monotonicity preserving conditions for $R_{i-1}(t; \alpha_{i-1}, \beta_{i-1})$ and $R_i(t; \alpha_i, \beta_i)$ which are

$$3\Delta_{i-1} - d_{i-1} - \beta_{i-1} d_i \ \ge\ 0,$$
$$3\Delta_{i-1} - d_i - \alpha_{i-1} d_{i-1} \ \ge\ 0, \qquad [27]$$
$$3\Delta_i - d_i - \beta_i d_{i+1} \ \ge\ 0, \qquad [28]$$
$$3\Delta_i - d_{i+1} - \alpha_i d_i \ \ge\ 0$$

and the C^2 continuity condition at x_i from [17], i.e.

$$3\Delta_{i-1} h_i + 3\Delta_i h_{i-1} - 3d_i(h_{i-1}+h_i) = \alpha_{i-1} h_i d_{i-1} + \beta_i h_{i-1} d_{i+1}. \qquad [29]$$

We shall show that α_{i-1} and β_i can be chosen so that conditions [27]-[29] are satisfied. α_i and β_{i-1} can be determined by using similar arguments on the corresponding adjacent curve segments. We have shown earlier that [27] and [28] can be achieved by taking the values of α_{i-1} and β_i sufficiently small by assuming $0 \le \alpha_{i-1} \le \alpha^*_{i-1}$ and $0 \le \beta_i \le \beta^*_i$ where α^*_{i-1} and β^*_i are as defined in [26] and [25] respectively. Let us take $\alpha_{i-1} = \alpha^*_{i-1}$ and $\beta_i = \beta^*_i$. However these values of α_{i-1} and β_i may not satisfy the C^2 continuity condition in [29]. Here, there are two cases as in section titled C^2 Condition.

Case 1: $3\Delta_{i-1} h_i + 3\Delta_i h_{i-1} - 3d_i(h_{i-1}+h_i) \ge 0$

In this case, the value for d_i estimated by the Fritsch and Butland's method is retained. In order to make equation [29] true, we have to choose the values of α_{i-1} and β_i appropriately. The values α^*_{i-1} and β^*_i determined as in [26] and [25] may not satisfy [29]. If it is so, we can scale the values of α^*_{i-1} and β^*_i by a scalar, $\lambda_i \in [0, 1]$. Let

$$\alpha_{i-1} = \lambda_i \alpha^*_{i-1}, \quad \beta_i = \lambda_i \beta^*_i. \qquad [30]$$

We will get the values of α_{i-1} and β_i which fulfill the C^2 condition [29] and monotonicity conditions [27], [28] by substituting [30] into [29] to obtain

$$\lambda_i = \frac{3(\Delta_{i-1} h_i + \Delta_i h_{i-1}) - 3d_i(h_{i-1}+h_i)}{\alpha^*_{i-1} h_i d_{i-1} + \beta^*_i h_{i-1} d_{i+1}} \ge 0.$$

Moreover, $\lambda_i \le 1$ as shown below. From [26] and [25], we obtain

$$3(\Delta_{i-1} h_i + \Delta_i h_{i-1}) - d_i(h_{i-1}+h_i) = \alpha^*_{i-1} h_i d_{i-1} + \beta^*_i h_{i-1} d_{i+1}.$$

So

$$3\left(\Delta_{i-1} h_i + \Delta_i h_{i-1}\right) - 3d_i\left(h_{i-1} + h_i\right) \le \alpha_{i-1}^* h_i d_{i-1} + \beta_i^* h_{i-1} d_{i+1} .$$

Hence $\lambda_i \le 1$. This ensures that while α_{i-1}^* and β_i^* are being scaled by λ_i in order to satisfy [29], the resulting values for α_{i-1} and β_i still satisfy [27] and [28].

Case 2: $3\Delta_{i-1} h_i + 3\Delta_i h_{i-1} - 3d_i\left(h_{i-1} + h_i\right) < 0$

In this case, the derivative d_i obtained by the Fritsch and Butland's method is too big. We shall scale down d_i accordingly by using the scalar α_i in [19]. With the modified derivative, [27]-[29] can be satisfied by choosing $\alpha_{i-1} = 0$ and $\beta_i = 0$ as our simple initial values.

The initial estimated derivatives d_0 and d_n may be 0. However, these two cases can be similarly treated.

As a result of the discussion above, we can get the initial values for α_{i-1} and β_i to satisfy the C^2 and monotonicity conditions on $[x_{i-1}, x_i]$ and $[x_i, x_{i+1}]$. We observe that β_0 on $[x_0, x_1]$ and α_{n-1} on $[x_{n-1}, x_n]$ will not be restricted by the C^2 condition. Hence, we can choose non-negative values for both of them which are sufficiently small to satisfy the monotonicity conditions.

So far we have only proved that the existence of α_i and β_i, $0 \le i \le n-1$ for the C^2 and monotonicity conditions. But in general, the solution set obtained as above is not necessarily good. So we will use this solution set as the initial values to minimize the mean curvature in [20] subject to the continuity and monotonicity linear constraints which are [17] for $i = 1, 2, \ldots, n-1$, [22] and [23] for $i = 0, 1, 2, \ldots, n-1$ and the conditions $\alpha_i \ge 0$, $\beta_i \ge 0$ to get optimal values for α_i and β_i, $0 \le i \le n-1$.

GRAPHICAL EXAMPLES

We describe some graphical examples to illustrate the schemes presented in sections titled C^1 Monotonicity Preserving Interpolation with Quartic Bézier-Like Spline and C^2 Monotonicity Preserving Interpolation with Quintic Bézier-Like Spline. The symbol '•' shown in all of the examples represents the data point of the interpolating curve.

The quartic Bézier-like spline curve shown in Figure 7(a) is generated by using the default values of $\alpha_i = 3$ and $\beta_i = 3$. Note that this curve is not monotonic. After applying the monotonicity conditions, the resulting curve which is shown in Figure 7(b) is now a C^1 monotonic increasing curve.

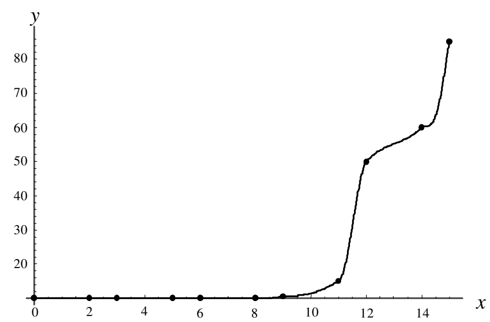

Figure 7: (a) C^1 quartic Bézier-like interpolant.

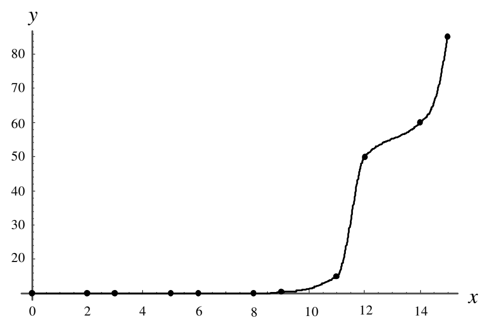

Figure 7: (b) C^1 monotonicity preserving quartic Bézier-like interpolant.

The default values $\alpha_i = 3$ and $\beta_i = 3$ are used to construct a C^1 quartic spline curve. The curve in Figure 8(a) has a "dip" in the last curve segment. Thus, we have to impose the monotonicity conditions to ensure the interpolating curve increases monotonically as shown in Figure 8(b).

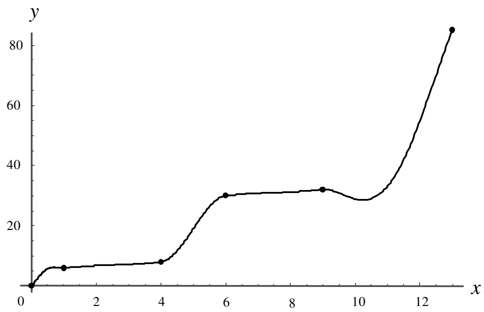

Figure 8: (a) C^1 quartic Bézier-like interpolant.

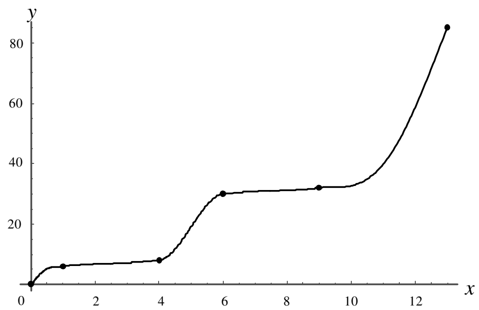

Figure 8: (b) C^1 monotonicity preserving quartic Bézier-like interpolant.

The C^2 piecewise quintic interpolant generated by using the C^2 interpolation scheme is shown in Figure 9(a). Note that the curve is smooth. However it is not monotonic although the data is monotonic increasing. There are unwanted "wiggles" on the curve. When we impose the C^2 and monotonicity conditions on the interpolating curve, the resulting curve which is shown in Figure 9(b) is now indeed a smooth monotonic increasing curve.

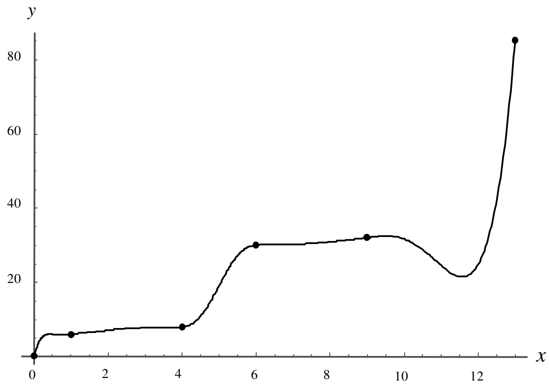

Figure 9: (a) C^2 quintic Bézier-like interpolant.

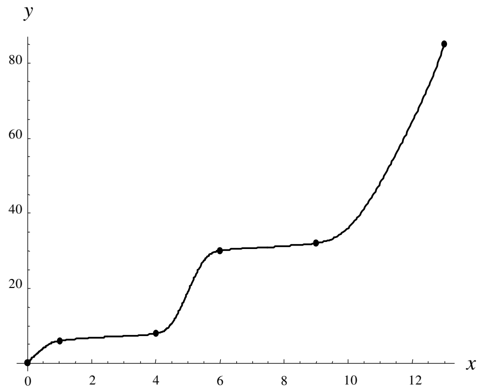

Figure 9: (b) C^2 monotonicity preserving quintic Bézier-like interpolant.

In Figure 10(a), the C^2 curve without the monotonicity preserving conditions has a "dip" in the last curve segment and its fourth segment is not monotonic. After applying the monotonicity conditions on the interpolating curve, the resulting curve increases monotonically as shown in Figure 10(b).

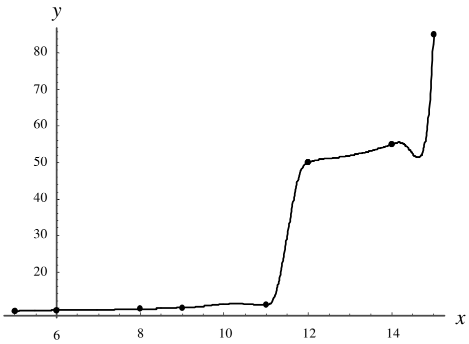

Figure 10: (a) C^2 quintic Bézier-like interpolant.

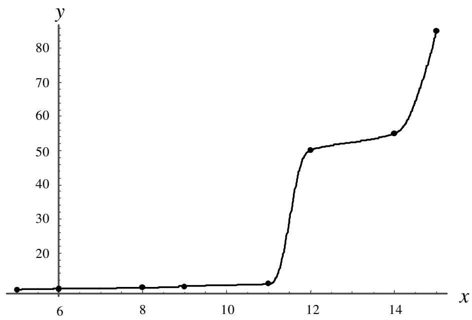

Figure 10: (b) C^2 monotonicity preserving quintic Bézier-like interpolant.

CONCLUSION

In this paper, we have presented a new family of Bézier-like curves namely the quartic Bézier-like curves and quintic Bézier-like curves. Though of different degrees, both the quartic and quintic have four control points like a cubic Bézier and two additional parameters for the control of shape and smoothness. We use the former type to construct a C^1 monotonicity preserving curve and the latter in C^2 monotonicity preserving interpolation. The implementation of a C^1 monotonic quartic Bézier-like spline curve interpolation is easier than the C^2 monotonic quintic Bézier-like spline curve interpolation since the former scheme is a local scheme. Thus any changes to a curve segment would not affect the whole curve.

Though the latter scheme is a global scheme, it produces interpolant which have a higher degree of smoothness. From both of the schemes that we have constructed, we can conclude that the parameters which are introduced in the Bézier-like curves allow us flexible shape control and it is very helpful for curve design.

ACKNOWLEDGEMENTS

The support of the FRGS Grant 203/PMATHS/671078 is acknowledged.

REFERENCES

Delbourgo, R. and Gregory, J.A. 1985. Shape preserving piecewise rational interpolation. *SIAM J. Sci. Stat. Comput*. **6**: 967-976.

Fritsch, F.N. and Butland, J. 1984. A method for constructing local monotone piecewise cubic interpolants. *Siam J. Sci. Stat. Comput.* **5**: 300-304.

Fritsch, F.N. and Carlson, R.E. 1980. Monotone piecewise cubic interpolation. *SIAM J. Numer. Anal.* **17**: 238-246.

Goodman,T.N.T., Ong, B.H. and Unsworth, K. 1991. Constrained interpolation using rational cubic splines, in *NURBS for Curve and Surface Design, ed. Gerald Farin (Philadelphia : SIAM)*, p.59-74.

Heß, Walter and Schmidt, Jochen W. 1994. Positive quartic, monotone quintic C^2-spline interpolation in one and two dimensions. *Journal of Computational and Applied Mathematics,* **55**: 51-67.

Jamaludin, M.A., Said, H.B. and Majid, A.A. 1996. Shape control of parametric cubic curves. In *Proceedings of SPIE – The International Society for Optical Engineering*, **2644**:128-133.

Optimization Toolbox User's Guide for Use with Matlab, The Maths Works, Inc. 2000.