数理解析研究所講究録
1514巻 2006年 46-52

An answer to an open problem on cubic spiral transition between two circles

鹿児島大学理学部 ズルフィカル ハビブ, 酒井 宦

Zulfiqar Habib * and Manabu Sakai †
Zulfiqar Habib, Kagoshima University, Japan ‡

Abstract

This paper derives a spiral condition for a single Bézier cubic transition curve of G^2 contact, between two circles with one circle inside the other. A spiral is free of local curvature extrema, making spiral design an interesting mathematical problem with importance for both physical and aesthetic applications. In high way, railway route, or satellite path design it is often desirable to have a transition curve from circle to circle. We treat an open problem on planar cubic Bézier spiral segments which are proposed as transition curve elements, examine techniques for curve design using the new results, and derive lower and upper bounds for the distance between two circles.

1 Introduction

Parametric cubic curves are popular in CAD applications because they are of low degree polynomial curves that allow inflection points (where curvature is zero), so they are suitable for the composition of G^2 blending curves. The Bézier form of a parametric cubic curve is usually used in CAD/CAM and CAGD (Computer Aided Geometric Design) applications [3] because of its geometric and numerical properties.

In this paper the term 'spiral' refers to a curved line segment whose curvature varies monotonically with constant sign. Such curves are useful for transition between circles and straight lines. A G^2 point of contact of two curves is a point where the two curves meet and where their unit tangent vectors and signed curvatures match, and a G^3 contact is a point where the derivative of curvatures also match.

In curve and surface design it is often desirable to have a transition curve of G^2 contact, composed of single cubic spiral segment, between two circles. The purpose may be practical, e.g., in highway design, railway route, satellite path, or aesthetic applications [2, 4]. Sudden changes between curves of widely different radii or between long tangents and sharp curves should be avoided by the use of curves of *gradually increasing or decreasing radii* without at the same time introducing an appearance of forced alignment. The importance of this design feature is discussed in [4].

Cubic curves, although smoother, are not always helpful since they might have unwanted cusps, loops, up to two inflection points, and curvature extrema [5, 6, 8, 9]. According to Farin [3], curvature extrema of a fair curve "should only occur where explicitly desired by the designer". B-splines and Bézier curves do not normally allow this. However, it can be accomplished when designing with cubic Bézier spirals. The transition curves of G^2 contact have been considered for joining (i) straight line to circle, (ii) circle to circle with a broken back C transition, (iii) circle to circle with an S transition, (iv) straight line to straight line and (v) circle to circle where one circle lies inside the other with a C oval transition.

Meek & Walton [10, 11, 12] have considered a possibility of cubic and PH (Pythagorean hodograph) quintic splines to join two circles with one circle inside the other. The *fifth* case (v) used in highway

*E-mail: habib@eniac.sci.kagoshima-u.ac.jp
†E-mail: msakai@sci.kagoshima-u.ac.jp, Department of Mathematics and Computer Science
‡Graduate School of Science and Engineering, Korimoto 1-21-35, Kagoshima University, Kagoshima 890-0065, Japan.

design has not been solved completely at this time. Numerical treatment of the case would imply that it does not always seem to have a solution. In their scheme, the beginning point of spiral segment can not be used as a point of G^2 contact with circle because the curvature there is 0 and therefore the scheme is not extendable to 3D non-planar cases.

We treat the unsolved *fifth* case with sufficient conditions for the existence of the spiral transition curve of G^2 or G^3 contact at the larger or smaller circle, and derive lower and upper bounds for the distance between two circles. Finally, we give some simple examples to show how to use theoretical results and draw the cubic Bézier spiral transition curve between two circles.

The organization of our paper is as follows. Next section gives a brief discussion of the notation and conventions for the cubic Bézier spiral with some theoretical background, then description of method followed by numerical examples, and concluding remarks.

2 Background, notation and conventions

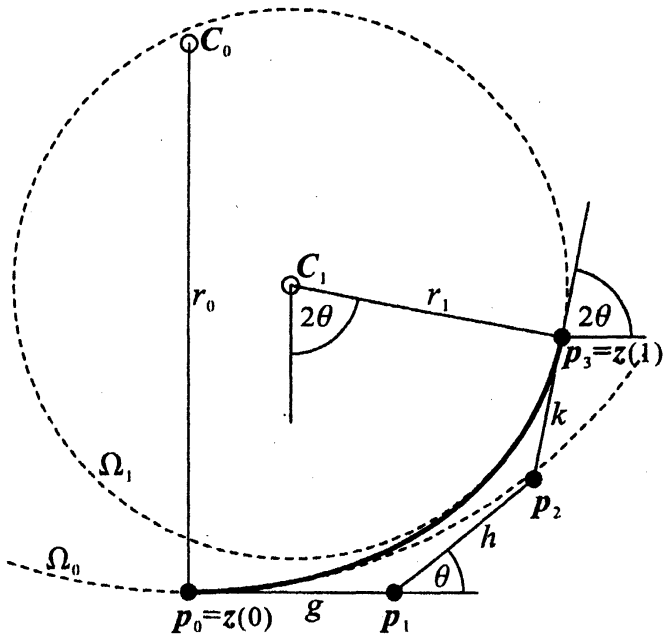


Figure 1: Circle to circle transition with single cubic spiral.

With reference to Figure 1, consider two circles Ω_0, Ω_1 centered at C_0, C_1 with radii r_0, r_1, such that Ω_1 is completely contained inside Ω_0 and $\mu = \sqrt{r_0/r_1}(> 1)$. It is desirable to join the two circles by a single cubic Bézier spiral segment such that both points of contact are G^2 or G^3. A cubic Bézier curve is given by [3] as

$$z(t) = (1-t)^3 p_0 + 3t(1-t)^2 p_1 + 3t^2(1-t)p_2 + t^3 p_3, \quad 0 \le t \le 1. \tag{2.1}$$

Its signed curvature $\kappa(t)$ is given by

$$\kappa(t) = z'(t) \times z''(t) / \|z'(t)\|^3, \tag{2.2}$$

where $\times$ stands for the two-dimensional cross product $(x_0, y_0) \times (x_1, y_1) = x_0 y_1 - x_1 y_0$ and $\|\bullet\|$ means the Euclidean norm. To simplify the analysis, we require

$$\begin{gathered}(i)\ p_0 = (0,0), \quad (ii)\ z'(0) \parallel (1,0), \quad (iii)\ z'(1) \parallel (\cos 2\theta, \sin 2\theta), \\ (iv)\ p_2 - p_1 \parallel (\cos\theta, \sin\theta), \quad (v)\ \kappa(0) = 1/r_0, \quad (vi)\ \kappa(1) = 1/r_1,\end{gathered} \tag{2.3}$$

where $2\theta(\in (0,\pi/2))$ is an anti-clockwise angle from the beginning tangent to the ending one (at the point of contact of the smaller circle Ω_1).

We represent (2.1) in terms of $(g,h,k)(=(\|p_0p_1\|,\|p_1p_2\|,\|p_2p_3\|))$ to obtain

$$\begin{aligned} x(t) &= 3g(1-t)^2t + 3(g+h\cos\theta)(1-t)t^2 + (g+h\cos\theta+k\cos 2\theta)t^3 \\ y(t) &= 3h(\sin\theta)(1-t)t^2 + (h\sin\theta+k\sin 2\theta)t^3. \end{aligned} \tag{2.4}$$

A simple calculation gives the values of curvatures at the end points of the segment

$$\kappa(0) = \frac{2h}{3g^2}\sin\theta\left(=\frac{1}{r_0}\right),\ \kappa(1) = \frac{2h}{3k^2}\sin\theta\left(=\frac{1}{r_1}\right). \tag{2.5}$$

To simplify the analysis, we introduce a parameter $p(=\sqrt{h/r_1})$ and obtain from (2.5)

$$(g,k) = pr_1\sqrt{\frac{2}{3}\sin\theta}\,(\mu,1). \tag{2.6}$$

Then, we have a family of cubic transition curves $\boldsymbol{z}(t)$ in (2.1) satisfying (2.3)

$$\begin{aligned} x(t) &= \frac{pr_1t}{3}\left\{3pt(3-2t)\cos\theta + \left((3-3t+t^2)\mu + t^2\cos 2\theta\right)\sqrt{6\sin\theta}\right\} \\ y(t) &= \frac{pr_1t}{3}\left\{3p(3-2t) + 2t\cos\theta\sqrt{6\sin\theta}\right\}\sin\theta. \end{aligned} \tag{2.7}$$

3 Description of method

The outline of our analysis is first to derive the spiral condition on (θ,μ) for the curve (2.7) and then find the lower and upper bounds for the distance between two circles. The results are expressed as Theorems 3.1-3.4.

The derivative of curvature in (2.2) for $t = 1/(1+s)$ is

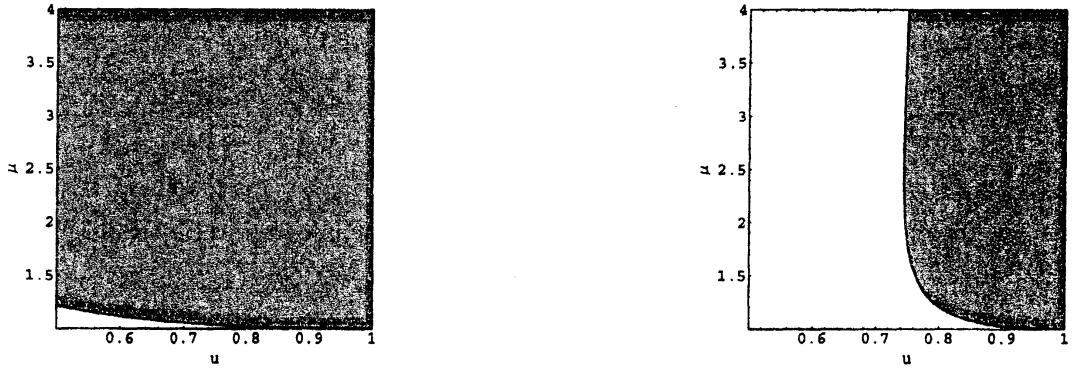
$$\kappa'(t) = \left[\frac{72p^4r_1^4(\sin\theta)^{3/2}}{(1+s)^5\{x'(t)^2+y'(t)^2\}^{5/2}}\right]\sum_{i=0}^{5}c_is^i, \tag{3.1}$$

where c_5, c_0 are quadratic and $c_i, 1\le i\le 4$ are cubic of the forms

$$\begin{aligned} c_5 &= \mu^2\sqrt{\sin\theta}\cos\theta\left(-9p^2+2p\mu\sqrt{6\sin\theta}/\cos\theta+2\mu\sin\theta\right)(=\mu^2c_5(p)), \\ c_4 &= \mu\left(-9\sqrt{6}p^3+15p^2\mu\cos\theta\sqrt{\sin\theta}-4p\mu\sqrt{6}\cos 2\theta\sin\theta+5\mu^2\sqrt{\sin\theta}\sin 2\theta\right) \\ &\quad (=\mu c_4(p)), \\ c_3 &= \mu\left(3\sqrt{6}p^3-48p^2\cos\theta\sqrt{\sin\theta}+2p\mu\sqrt{6}(5+4\cos 2\theta)\sin\theta-2\mu\sqrt{\sin\theta}\sin 4\theta\right) \\ &\quad (=\mu c_3(p)), \\ c_2 &= -3\sqrt{6}p^3+48p^2\mu\cos\theta\sqrt{\sin\theta}-2p\mu\sqrt{6}(5+4\cos 2\theta)\sin\theta+2\mu^2\sqrt{\sin\theta}\sin 4\theta \\ &\quad (=c_2(p)), \\ c_1 &= 9\sqrt{6}p^3-15p^2\cos\theta\sqrt{\sin\theta}+4p\mu\sqrt{6}\cos 2\theta\sin\theta-5\mu\sqrt{\sin\theta}\sin 2\theta\ (=c_1(p)), \\ c_0 &= \sqrt{\sin\theta}\cos\theta\left(9p^2-2p\sqrt{6\sin\theta}/\cos\theta-2\mu\sin\theta\right)(=c_0(p)), \end{aligned} \tag{3.2}$$

It is convenient to define

$$(u,d) = \left(\cos^2\theta, \sqrt{\sin\theta}/\cos\theta\right),\quad (H_0,H_1) = \left(\sqrt{1+3u\mu}, \sqrt{\mu^2+3u\mu}\right),$$



(a) $c_4(p_0) \geq 0$ (b) $c_4(p_5) \geq 0$

Figure 2: Regions described by Theorem 3.1 (left) and Theorem 3.2 (right).

where $1/2 < u < 1$. Let c_0, $c_5 \geq 0$, i.e., $p_0 \leq p \leq p_5$ where

$$(p_0, p_5) = \frac{d}{3}\sqrt{\frac{2}{3}}\,(1 + H_0, \mu + H_1)\,.$$

Depending on the value of p ($= p_0$ or p_5), the transition curve (2.1) has G^3 contact either with smaller or larger circle, respectively, because the derivative of curvature there is 0. Both of these cases are discussed in the following sections. When drawing curves interactively it is often desirable to know the valid regions for placing the values of u and μ to ensure required features. Figure 2 determine regions in which to place (u, μ) for the more general cubic Bézier spiral segment with G^3 contact on one side and G^2 on the other side.

3.1 Case 1. The transition curve of G^3 contact at the smaller circle

For $p = p_0$, we have the following values of $c_i(p)$, $i = 1, \ldots, 4$ from (3.2).

$$(c_1(p), c_2(p), c_3(p), c_4(p)) = \frac{2d^3}{27}\,(3d_1, d_2, d_3, 3d_4)\,, \tag{3.3}$$

where

$$\begin{aligned}
d_1 &= 2\left\{3u(4u-1)\mu - 5u + 4\right\} H_0 - 2\left\{3u(6u-1)\mu + 5u - 4\right\},\\
d_2 &= 4\left\{9u^2(1+6u)\mu^2 - 3u(12u-5)\mu - 2\right\} - 8\left\{9u(2u-1)\mu + 1\right\} H_0,\\
d_3 &= 4\left\{9u(1+3u-6u^2)\mu - 24u + 2\right\} + 8\left\{3u(1+6u)\mu - 12u + 1\right\} H_0,\\
d_4 &= 4\left\{15u^2\mu^2 + u(1-6u)\mu - 2\right\} - 8\left\{u(3u-2)\mu + 1\right\} H_0.
\end{aligned}$$

Theorem 3.1. *If $c_4(p_0) \geq 0$ (i.e., (u, μ) belongs to the gray region in Figure 2(a)), then the cubic curve $z(t)$ $(= (x(t), y(t)))$ of the form (2.7) of G^3 contact at the smaller circle joining the two circles with one circle Ω_1 inside the other circle Ω_0 is a spiral.*

Proof. First note that d_1, $d_2 \geq 0$ where (i) $a - b \geq 0$ comes from $a \geq 0$, $a^2 - b^2 \geq 0$ and (ii) the quadratic $q(\mu) \geq 0$ for $\mu > 1$ comes from $q''(\mu)$, $q'(1)$ and $q(1) \geq 0$. Let

$$\mu_1 = \frac{-1 + 18u + 135u^2 - 108u^3 + 108u^4 - (1+3u)(1-6u)\sqrt{1 + 34u + 157u^2 - 84u^3 + 36u^4}}{8u(1+6u)^2}, \tag{3.4}$$

$$\mu_2 = \frac{1 + 2u + 6u^2 + 2(1+u)\sqrt{19 - 12u + 9u^2}}{25u}.$$

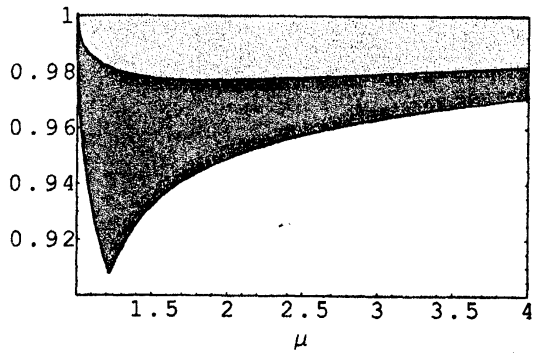


Figure 3: Graphs of $m_1(\mu)$ (all gray) and $m_2(\mu)$ (only light gray) with respect to μ.

Then $d_3 \geq 0$ if $\mu \geq \mu_1$ and $d_4 \geq 0$ if $\mu \geq \mu_2$. *Mathematica* helps us to show that $\mu_2 \geq \mu_1$. □

3.1.1 Upper and lower bounds for the distance between two circles

It is often required to find the transition curve for given distance between two circles. Here we derive the condition on the magnitude of $\|C_0 - C_1\|$ for the existence of the circle to circle transition curve (2.1) of G^2 contact. Since $C_0 = (0, r_0)$ and $C_1 = z(1) - r_1(\sin(2\theta), -\cos(2\theta))$, *Mathematica* helps us to obtain $C_1 - C_0 = r_1(g_1(p_0), g_2(p_0))$ where

$$\begin{aligned} g_1(p) &= p^2 \cos\theta + p(\mu + \cos 2\theta)\sqrt{\frac{2\sin\theta}{3}} - \sin 2\theta, \\ g_2(p) &= p^2 \sin\theta + p\sqrt{\frac{2\sin\theta}{3}} \sin 2\theta - \mu^2 + \cos 2\theta. \end{aligned} \tag{3.5}$$

Since

$$\|C_0 - C_1\| = r_1\sqrt{g_1^2(p_0) + g_2^2(p_0)}\,(= r_1 h(p_0)) = \left\{\frac{h(p_0)}{\mu^2 - 1}\right\}(r_0 - r_1). \tag{3.6}$$

A change in $u(= \cos^2\theta)$ for each μ satisfying $c_4(p_0) = 0$ gives an attainable range of $m_1(\mu)(= h(p_0)/(\mu^2 - 1))$ shown by bold curve in Figure 3 and an upper bound is 1 $(= \lim_{u\to 1} m_1(\mu))$.

Theorem 3.2. *The cubic curve $z(t)(= (x(t), y(t))$ of the form (2.7) of G^3 contact at the smaller circle joining the two circles with one circle Ω_1 inside the other circle Ω_0 exists if $m_1(\mu)(r_0 - r_1) \leq \|C_0 - C_1\| < (r_0 - r_1)$.*

Proof. This result follows immediately from the above discussion. □

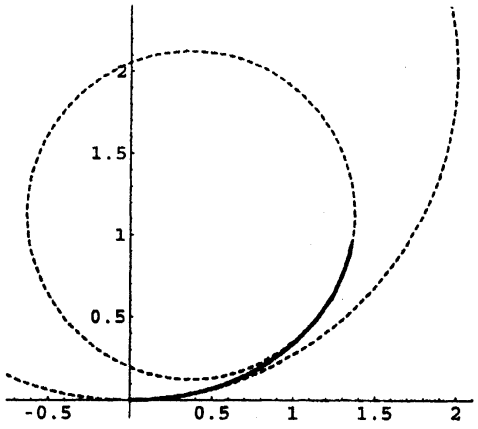
3.2 Case 2. The transition curve of G^3 contact at the larger circle

For $p = p_5$, we have the following values of $c_i(p)$, $i = 1, \ldots, 4$ from (3.2).

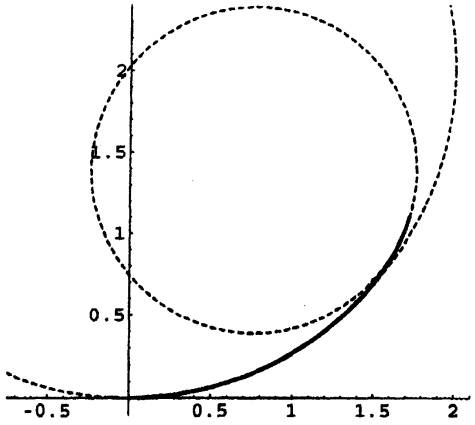
$$(c_1(p), c_2(p), c_3(p), c_4(p)) = \frac{2\mu d^3}{27}\left(3\bar{d}_1, \bar{d}_2, \bar{d}_3, 3\bar{d}_4\right), \tag{3.7}$$

where

$$\begin{aligned} \bar{d}_1 &= 8\{\mu + u(3u-2)\}H_1 + 4\{2\mu^2 + u(6u-1)\mu - 15u^2\}, \\ \bar{d}_2 &= 4\mu\{2(12u-1)\mu + 9u(6u^2 - 3u - 1)\} + 8\{(12u-1)\mu - 3u(6u+1)\}H_1, \\ \bar{d}_3 &= 8\{\mu + 9u(2u-1)\}H_1 + 4\{2\mu^2 + 3u(12u-5)\mu - 9u^2(6u+1)\}, \\ \bar{d}_4 &= 2\mu\{(5u-4)\mu + 3u(6u-1)\} + 2\{(5u-4)\mu - 3u(4u-1)\}H_1. \end{aligned}$$



(a) Case 1: G^3 contact at the smaller circle



(b) Case 2: G^3 contact at the larger circle

Figure 4: G^2 cubic Bézier spiral transition between two circles.

Theorem 3.3. *If $c_4(p_5) \geq 0$ (i.e., (u, μ) belongs to the gray region in Figure 2(b)), then the cubic curve $\boldsymbol{z}(t)(= (x(t), y(t))$ of the form (2.7) of G^3 contact at the larger circle joining the two circles with one circle Ω_1 inside the other circle Ω_0 is a spiral.*

Proof. As in the proof of Theorem 3.1, $\bar{d}_1, \bar{d}_2 \geq 0$. By a simple calculation, $\bar{d}_3 \geq 0$ is equivalent to

$$\mu \geq \frac{3u(1+6u)^2}{13 - 90u + 144u^2 + 4(6u-1)\sqrt{10 - 30u + 36u^2}} \ (= \mu_3).$$

Similarly, $\bar{d}_4 \geq 0$ is equivalent to

$$\mu \geq \frac{3u(1-4u)^2}{4 - 23u + 30u^2 + 2(5u-1)\sqrt{4 - 21u + 21u^2}} \ (= \mu_4), \ \frac{21+\sqrt{105}}{42} \leq u < 1$$

$$\mu \leq \frac{(2u-1)(15u-4) + 2(5u-1)\sqrt{4 - 21u + 21u^2}}{5(4-5u)}, \ \frac{21+\sqrt{105}}{42} \leq u < \frac{4}{5}.$$

Mathematica helps us to show that $\mu_4 \geq \mu_3$. □

3.2.1 Upper and lower bounds for the distance between two circles

Similarly as in Section 3.1.1, we obtain an attainable range of $h(p_5)/(\mu^2 - 1)(= m_2(\mu))$ shown by normal curve in Figure 3 and an upper bound being 1.

Theorem 3.4. *The cubic curve $\boldsymbol{z}(t)(= (x(t), y(t))$ of the form (2.7) of G^3 contact at the larger circle joining the two circles with one circle Ω_1 inside the other circle Ω_0 exists if $m_2(\mu)(r_0 - r_1) \leq \|\boldsymbol{C}_0 - \boldsymbol{C}_1\| < (r_0 - r_1)$.*

Proof. This result follows immediately from the above discussion. □

4 Numerical examples

For example, we consider the conditions $\|\boldsymbol{C}_0 - \boldsymbol{C}_1\| = m(r_0 - r_1)$ for $(r_0, r_1) = (2, 1)$.

Case 1 Figure 4(a) gives the transition curve for $m = 0.95$ and then $\theta \approx 0.689104$.

Case 2 Figure 4(b) is the one for $m = 0.98$ and then $\theta \approx 0.64172$.

5 Conclusion

The cubic Bézier spiral is a reasonable alternative to the clothoid [1] for applications such as highway, railway or satellite path design because it is of low degree; being polynomial, unlike the clothoid, it is a special case of a NURBS curve. Since straight line segments and circular arcs also have NURBS representations [7], curves designed using a combination of cubic Bézier spirals, circular arcs, and straight line segments can be represented entirely by NURBS.

We discussed the fifth unsolved case of Walton & Meek [10, 11, 12] and completed the analysis for most of the practical applications. We proved the existence of cubic spiral transition curve(s) of G^2 contact between two circles depending on the given distance between their centers. Our method uses a parametric value that ranges from 0 to 1, and the curvatures at the end points are nonzero and therefore they can be used as the points of G^2 or G^3 contact and the scheme is extendable to 3D non-planar cases. The algebra required to prove uniqueness seems unwieldy and has not been carried out at this time. Numerical examples are given to prove our analysis and to show how to plot the cubic Bézier spiral transition curve from circle to circle where one circle lies inside the other, using our theoretical results.

Acknowledgment

This work is supported by Japan Society for the Promotion of Science (JSPS/FF1/315-P04034).

References

[1] Baass, K. G. 1984. The use of clothoid templates in highway design. *Transportation Forum* 1:47–52.

[2] Burchard, H., J. Ayers, W. Frey, and N. Sapidis. Designing fair curves and surfaces chapter Approximation with aesthetic constraints. 1993.

[3] Farin, G. 1997. *Curves and Surfaces for Computer Aided Geometric Design: A Practical Guide.* NewYork : Academic Press 4th edition.

[4] Gibreel, G. M., S. M. Easa, Y. Hassan, and I. A. El-Dimeery 1999. State of the art of highway geometric design consistency. *ASCE Journal of Transportation Engineering* 125(4):305–313.

[5] Habib, Z. and M. Sakai 2002. G^2 two-point Hermite rational cubic interpolation. *International Journal of Computer Mathematics* 79(11):1225–1231.

[6] Habib, Z., M. Sarfraz, and M. Sakai 2005. Rational cubic spline interpolation with shape control. *Computers & Graphics* 29(4):594–605.

[7] Piegl, L. and W. Tiller 1995. *The NURBS book.* : Springer.

[8] Sakai, M. 1999. Inflection points and singularities on planar rational cubic curve segments. *Computer Aided Geometric Design* 16:149–156.

[9] Sakai, M. 2001. Osculatory interpolation. *Computer Aided Geometric Design* 18:739–750.

[10] Walton, D. J. and D. S. Meek 1996. A planar cubic Bézier spiral. *Computational and Applied Mathematics* 72:85–100.

[11] Walton, D. J. and D. S. Meek 1998. Planar G^2 curve design with spiral segments. *Computer Aided Design* 30:529–538.

[12] Walton, D. J., D. S. Meek, and J. M. Ali 2003. Planar G^2 transition curves composed of cubic Bézier spiral segments. *Computational and Applied Mathematics* 157(2):453–476.