CONSTRAINED INTERPOLATION BY PARAMETRIC RATIONAL CUBIC SPLINES

by

LAU BEE FANG

Dissertation submitted in partial fulfillment of the requirements for the degree of Master of Science in Mathematics

June 2010

ACKNOWLEDGEMENTS

I would like to seize this opportunity to express my most sincere gratitude to my supervisor, Dr. Kong Voon Pang who has been very dedicated and patient in giving guidance, advice and suggestions towards the completion of this dissertation. Also, I would like to express my appreciation to all my lectures, the staff of the USM (School Of Mathematical Sciences) and my course mates.

Finally and most importantly, to my parents, my brothers and sisters, thank you very much for all the support and encouragement. Especially, my parents (Lau Kim Joon & Saw Siew Hong) and my second sister (Lau Bee Eng) have been giving me endless support throughout the preparation of this dissertation. I would not have accomplished my goal without all of you. Lastly, I would like to dedicate the dissertation to the mighty God.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTERPOLASI TERKEKANG DENGAN SPLIN KUBIK NISBAH BERPARAMETER

ABSTRAK

Interpolasi terkekang adalah berguna dalam masalah seperti mereka bentuk sebuah lengkung yang perlu dihadkan dalam suatu kawasan tertentu. Dalam disertasi ini, kami membincangkan interpolasi terkekang dengan menggunakan splin kubik nisbah yang diperkenalkan dalam (Goodman et al, 1991). Terdapat dua kaedah pengubahsuaian lengkung disarankan, kaedah yang melibatkan modifikasi pemberat α, β berkaitan dengan titik hujung segmen lengkung dibincangkan dalam disertasi ini. Skim ini memperoleh sebuah G^2 lengkung interpolasi yang terletak di sebelah garis-garis yang diberikan seperti data yang diberikan. Sebagai perkembangan daripada kertas ini, kami akan memperoleh satu skim interpolasi terkekang alternatif dengan menggunakan lengkung kubik nisbah. Pemberat Ω, θ yang berkaitan dengan titik kawalan dalaman diubah suai untuk memperoleh sebuah G^1 lengkung interpolasi yang terletak di sebelah garis-garis yang diberikan seperti data yang diberikan.

ABSTRACT

Constrained interpolation could be useful in problem like designing a curve that must be restricted within a specified region. In this dissertation, we discuss constrained interpolation using rational cubic splines introduced in (Goodman et al, 1991). There are two curve modification methods suggested and the one which involves modification of the weights α , β associated with the end points of the curve segments is discussed in this dissertation. This scheme obtains a G^2 interpolating curve which lies on one side of the given lines as the given data. Extension from this paper, we will derive an alternative constrained interpolation scheme using rational cubic curve. The weights Ω , θ associated with the inner control points are modified to obtain a G^1 interpolating curve which lies on one side of the given lines as the given data.

CHAPTER 1

INTRODUCTION

In the field of Computer Aided Geometric Design (CAGD), there are many methods and approaches in generating curves and surfaces. Rational interpolation is one of the fundamental interpolation concepts. Many interpolation methods use splines which interpolate two consecutive data points to form a curve segment and these curve segments will be joined together to form a smooth interpolating curve. The interpolating curve is shape preserving in the sense that if has the minimal number of inflections consistent with the data.

Constrained interpolation or shape preserving interpolation could be useful in problems like designing a smooth curve that must fit within a specified region such as non-negativity preservation. Non-negativity is a very important aspect of shape. There are many physical situations where entities only have meaning when their values are positive. For Example, in a probability distribution the representation is always positive or when dealing with samples of population, the data is always positive.

The problem of shape preserving interpolation has been considered by a number of authors, for example, in (Goodman, 1988), (Goodman et al, 1991), (Ong & Unsworth, 1992), (Li & Meek, 2006) and (Saifudin et al, 2006).

Goodman (1988) described the interpolation which preserves local convexity and local monotonicity. Particular schemes are then given for interpolation by parametric piecewise polynomials. The interpolating curves can have geometry continuity and they are convex in the region where the data are convex.

Goodman, et al. (1991) have derived a scheme in constructing a G^2 parametric interpolating curve which lies on the same side of a given set of constraint lines as the data. The method depends on necessary and sufficient conditions for a rational cubic to cross a line. Two methods of modification are discussed, one by changing the weights of the rational cubic segments and the other by the addition of new interpolation points. In (Ong & Unsworth, 1992), a similar scheme is applied to the given functional data where a C^1 non-parametric interpolating curve which lies on the same side of the given constraint lines as the data.

Li & Meek (2006) have derived a smooth, obstacle-avoiding curve from a given obstacle-avoiding polyline path. A method is given to replace that polyline path by a G^2 cubic spline curve that also avoids the obstacles. Saifudin, et al. (2006) use Bézier-like quartic to preserve the shape of positive interpolating curve. The degree of the smoothness is C^1 continuity.

In this dissertation we will discuss the constrained interpolation scheme presented in (Goodman et al, 1991) and derive an alternative scheme using rational cubic curve. We compare these two schemes for non-negativity preserving and constrained interpolation. In (Goodman et al, 1991), given data points in a plane lying on one side of one or more given line, a scheme is derived which generates a planar curve interpolating the given data points and lying on the same side of the given lines as the data points. Necessary and sufficient conditions are derived for a rational cubic to cross a line. A default curve is first generated and the segments which have crossed over any of the given lines are modified. There are two methods suggested and the one which involves modification of the weights associated with the end points of the curve segments is discussed in this dissertation. Extension from this paper, we will derive an alternative constrained interpolation scheme using rational cubic curve whereas the weights associated with the inner Bézier points of the curve segments are manipulated.

In Chapter 2, we will state the definition of rational Bézier curve, some properties of rational Bézier curve and the effect of the weights to rational Bézier curve. The rational Bézier spline curve have been widely implemented in computer aided geometric design (CAGD) specifically for conic section which cannot be represented exactly in the usual Bézier form. The rational Bézier curve adds adjustable weights to provide closer approximations to arbitrary shape. The geometric continuity at the joints of rational Bézier spline is also discussed in this chapter.

Necessary and sufficient conditions to ensure the non-negativity for a rational cubic curve are discussed in chapter 3. The constrained interpolation scheme proposed by Goodman, et al. (1991) will be described in Chapter 4. Rational Cubic Splines are used to generate a G^2 interpolating curve. The weights α and β associated to the end points of curve segment are utilized to ensure the resulting curve lies on one side of the given lines as the given data. Some numerical example are presented graphically by using this constrained interpolation scheme.

In Chapter 5, we derive an alternative constrained interpolation scheme using rational cubic curve. The weights Ω, θ which associated with the inner control points are modified to obtain a G^1 interpolating curve which lies on one side of the given lines as the given data. Some graphical examples are presented to illustrate this G^1 constrained interpolation scheme.

Lastly, Chapter 6 provides the comparison between the two schemes discussed in chapters 4 and 5 and the conclusion.

CHAPTER 2

PARAMETRIC RATIONAL BÉZIER

2.1 Introduction

Rational function is the quotient of two polynomials. It can be represented as

$$R(x) = \frac{Y(x)}{W(x)},$$

$Y(x)$ and $W(x)$ are polynomials. For rational Bézier curve of degree n, it is defined parametrically by

$$R(t) = \frac{P(t)}{W(t)} = \frac{\sum_{i=0}^{n} \omega_i C_i B_i^n(t)}{\sum_{i=0}^{n} \omega_i B_i^n(t)}, \quad 0 \le t \le 1, \tag{2.1}$$

where $B_i^n(t)$ are the nth-degree Bernstein functions. C_i are called Bézier points. They form the Bézier polygon of the rational curve. $\omega_i > 0$ are called the weights.

Besides that, rational Bézier curve (2.1) also can be rewritten as

$$R(t) = \sum_{i=0}^{n} C_i R_i^n(t), \quad 0 \le t \le 1,$$

where

$$R_i^n(t) = \frac{\omega_i B_i^n(t)}{\sum_{j=0}^{n} \omega_j B_j^n(t)}, \quad i = 0, 1, \ldots., n,$$

are the rational basis functions.

2.2 Effect of Weights

For rational Bézier curve, control points can be used to vary the shape of the curve. Besides that, the weights ω_i can also be used as additional design parameter. It is because each control point in rational Bézier curve is assigned a weight. The weight defines how much does a point "attract" the curve. Thus, the shape only changes if weights of the control points are different. However, in this dissertation we only concentrate on the method by modifying the weights of the rational Bézier curve without changing the control points. We will only consider the rational cubic Bézier curve as in (Goodman et al, 1991) which gives enough degree of curve designing.

The following Figure 2.1 and 2.2 show some examples of the rational cubic Bézier curve

$$R(t)=\frac{\omega_0 AB_0^3(t)+\omega_1 BB_1^3(t)+\omega_2 CB_2^3(t)+\omega_3 DB_3^3(t)}{\omega_0 B_0^3(t)+\omega_1 B_1^3(t)+\omega_2 B_2^3(t)+\omega_3 B_3^3(t)}, \tag{2.2}$$

which yields a family of rational Bézier curves with different weight values. First example is shown in Figure 2.1 where the weight ω_1 is varied as

ω_1=0, ω_1=0.5, ω_1=1, ω_1=2 and ω_1=10.

Observe that increasing weight ω_1 causes the curve to move towards the associated Bézier point B. Similar result holds for ω_2 towards the Bézier point C. Second example is shown in Figure 2.2 where the weight ω_0 is manipulated as

ω_0=0, ω_0=0.5, ω_0=2, ω_0=10 and ω_0=20.

The curves are move toward the associated Bézier point A when we increase the weight ω_0. Similarly, the curve is move toward the Bézier point D as the weight ω_3 increasing.

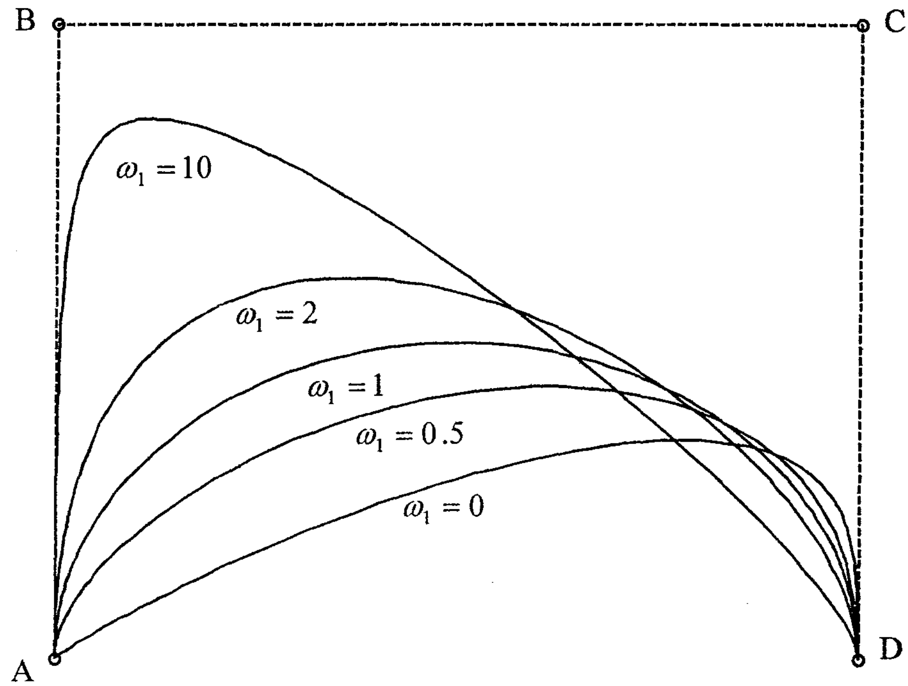

Figure 2.1 Rational cubic Bézier curves with different values ω_1 and the same Bézier polygon.

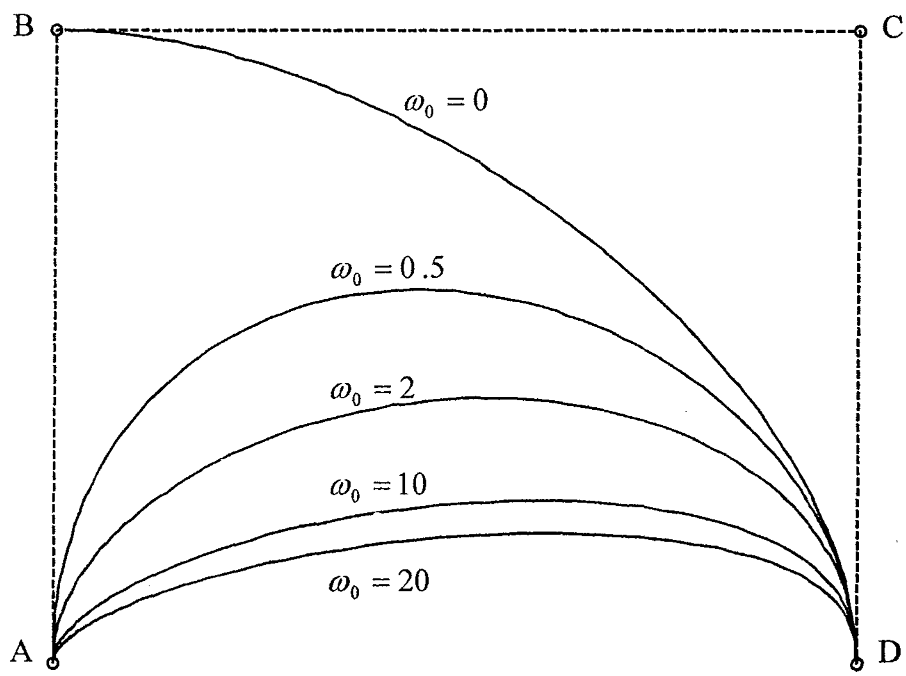

Figure 2.2 Rational cubic Bézier curves with different values ω_0 and the same Bézier polygon.

In general, increasing a weight ω_i causes all points on the curve to move towards the Bézier point, while decreasing the weight cause all points to move away from the Bézier point.

Let $\omega_3 = r\omega_0$ where $r > 0$. Assume $r = 3$, from Figure 2.3, we observe that although the weights ω_0 and ω_3 of each curve are increasing, but the ratio of the weights ω_0 to ω_3 is always the same i.e. 1:3. Thus, the curves are three times more "attract" by the Bézier point D compare with the Bézier point A.

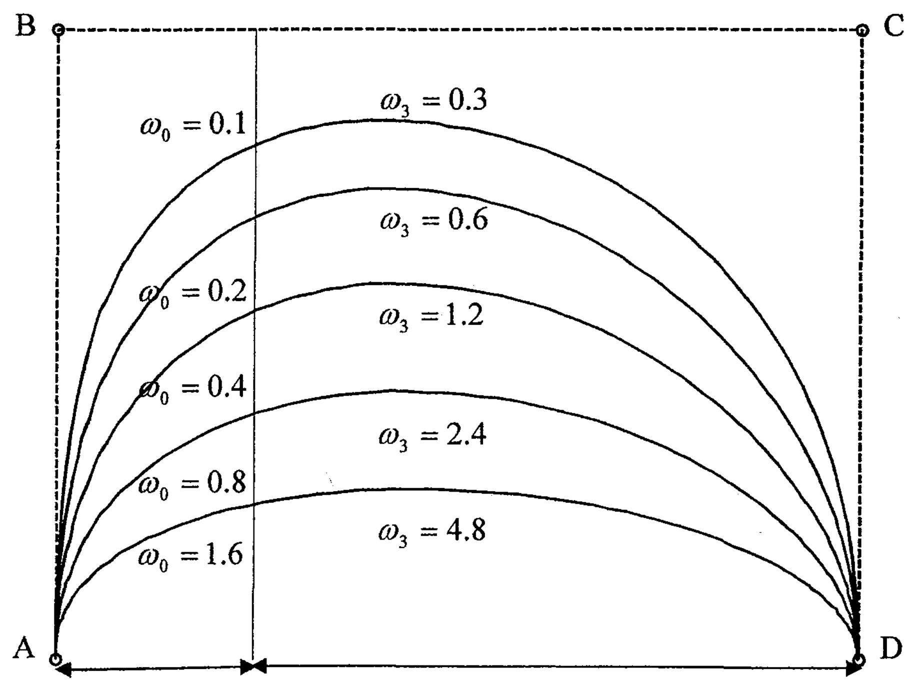

Figure 2.3 Rational cubic Bézier curves with ratio weight ω_0 to ω_3 is 1:3.

If the weights of rational Bézier curve are

$$\omega_0 = \omega_1 = \omega_2 = \omega_3,$$

then

$$R(t) = AB_0^3(t) + BB_1^3(t) + CB_2^3(t) + DB_3^3(t), \quad 0 \le t \le 1.$$

Hence, Bézier curve is a special case of rational Bézier curve.

2.3 Geometric Continuity

It is rare that an application requires the generation of a single curve. Complex curves are often created segment-by-segment. The nature of the complete curve is determined by continuity conditions at each join point. If two curve segments have the same value at a join point, the overall curve is said to exhibit G^0 continuity. It can be written as

$$R^-(z) = R^+(z),$$

where z is a join point. G^1 continuity implies that curve segments have matching values at a join point and first derivatives match in direction, but not in magnitude. Thus G^1 are,

$$R^-(z) = R^+(z) \quad \text{and} \quad R'^-(z) = \psi R'^+(z)$$

where

$$\psi = \frac{|R'^-(z)|}{|R'^+(z)|}.$$

G^2 continuity implies that the unit tangent vector and curvature vary continuously along the curve. It can be written as

$$R^-(z) = R^+(z)\ ,\ R'^-(z) = \psi R'^+(z) \text{ and } R''^-(z) = \psi^2 R''^+(z) + \phi R'^+(z)$$

where $\phi \in \Re$.

CHAPTER 3

NON-NEGATIVE RATIONAL CUBIC BÉZIER CURVE

3.1 Introduction

The rational cubic basis functions, $R_i^3(t)$ has satisfy the properties of non-negativity and partition of unity. $R_i^3(t)$ is non-negative on the interval [0, 1], i.e.

$$R_i^3(t) \geq 0 \quad , \quad 0 \leq t \leq 1, \quad i = 0, 1, \ldots, n.$$

The partition of unity means that the sum of the rational cubic basis functions is one on the interval [0,1], i.e.

$$\sum_{i=0}^{3} R_i^3(t) = 1, \quad 0 \leq t \leq 1.$$

The non-negativity and partition of unity properties lead to two important properties of the rational cubic curve, namely, the convex hull property and the invariance under affine transformation.

3.2 Necessary and Sufficient Conditions for the Polynomial to Cross a Line

Let us first consider the cubic polynomial

$$p(t) = a(1-t)^3 + b3t(1-t)^2 + c3t^2(1-t) + dt^3, \qquad 0 \leq t \leq 1, \qquad a,\ d > 0,$$

as the case for the parametric rational cubic in Section 3.3 can be reduced to this simpler case, as discussed later.

In this section, we would like to determine the necessary and sufficient conditions for determining when the polynomial $p(t) \le 0$ for some $t \in (0,1)$. We observe that the cubic Bézier Bernstein basis functions $B_0^3(t)$, $B_1^3(t)$, $B_2^3(t)$, $B_3^3(t)$ are always positive for $t \in (0,1)$. When both the coefficients b and c are non-negative, then clearly $p(t) > 0$. Thus for $p(t) \le 0$, it is necessary that $b < 0$ and/ or $c < 0$. By using elementary algebra, the following result was obtained (Goodman et al, 1991).

Theorem 3.1: Let

$$p(t) = a(1-t)^3 + b3t(1-t)^2 + c3t^2(1-t) + dt^3, \qquad 0 \le t \le 1,$$

where $a, d > 0$, $b < 0$ and/ or $c < 0$. If $p(t) < 0$ for some $t \in (0,1)$ (respectively, $p(t) = 0$ for only one point in (0,1)), then

$$c^2 > bd, \qquad b^2 > ac \tag{3.1}$$

and

$$3b^2c^2 + 6abcd - 4(ac^3 + b^3d) - a^2d^2 > 0 \quad \text{(respectively=0).} \tag{3.2}$$

Moreover, if (3.2) holds, then (3.1) also holds and $p(t) < 0$ for some $t \in (0,1)$ (respectively, $p(t) = 0$ for only one point in (0,1)).

We shall now apply the above theorem to determine when the parametric cubic Bézier polynomial curve

$$p(t) = A(1-t)^3 + B3t(1-t)^2 + C3t^2(1-t) + Dt^3, \quad 0 \le t \le 1,$$

cross over a given line, where *A*, *B*, *C*, *D* are points in the plane with *A*, *D* on the same side of a given line.

When the given line is the x-axis, the situation is reduced to a scalar problem. Thus Theorem 3.1 applied with $a = A_y$, $b = B_y$, $c = C_y$, $d = D_y$, where A_y, B_y, C_y and D_y are the y-coordinates of the points A, B, C, D respectively. Observe that in this case A_y, B_y, C_y and D_y are also the signed distances of the points A, B, C, D from the x-axis.

For the case of an arbitrary line, the necessary and sufficient conditions are the same as in Theorem 3.1 except that now a, b, c, d are the signed distances of the points A, B, C, D from the given line. The distances of the control points from the given line can be found using the concept of the coordinate geometry. If a point and line are given respectively as $P(p_x, p_y)$ and $ex + fy + g = 0$, the distance ℓ of the point P from the line is

$$\ell = \frac{| ep_x + fp_y + g |}{\sqrt{e^2 + f^2}}.$$

Then the signed distance of point from the line have positive distances if the point lies on the same side of the line as A and D the given points, otherwise it is negative.

3.3 Conditions for a Rational Cubic Curve to Cross a line

Let now consider the parametric rational cubic Bézier curve

$$R(t) = \frac{A\alpha(1-t)^3 + Bt(1-t)^2 + Ct^2(1-t) + D\beta t^3}{\alpha(1-t)^3 + t(1-t)^2 + t^2(1-t) + \beta t^3}, \quad 0 \le t \le 1, \tag{3.3}$$

where $\alpha, \beta > 0$ and $A, B, C, D \in \Re^2$. To make it more precise and also for easier reference later, we have written it in the following corollary which is on analogue of Theorem 3.1.

Corollary 3.1: Let $R(t)$ be given by (3.3), where $\alpha, \beta > 0$, also $A, B, C, D \in \Re^2$, and A, D lie on one side of a given line while B and/ or C lie on the opposite side of the line. If $R(t)$ crosses the given line (respectively, touches the line at only one point), then

$$c^2 > 3\beta bd \qquad \text{and} \qquad b^2 > 3\alpha ac \tag{3.4}$$

and

$$b^2c^2 + 18\alpha\beta abcd - 4(\alpha ac^3 + \beta b^3 d) - 27\alpha^2\beta^2 a^2 d^2 > 0, \tag{3.5}$$

where a, b, c, d are signed distances of the points A, B, C, D from the given line. Moreover, if (3.5) holds, the (3.4) also holds and $R(t)$ crosses over the line (respectively, touches the given line at only one point).

To test whether the parametric rational cubic Bézier curve $R(t)$ crosses over the given line, we just have to check the value of (3.5). If

$$b^2c^2 + 18\alpha\beta abcd - 4(\alpha ac^3 + \beta b^3 d) - 27\alpha^2\beta^2 a^2 d^2 > 0,$$

then $R(t)$ has crossed the given line and it can be modified by scaling its weights α and β by a factor <1 in order to make $R(t) \geq 0$. These can be done by the property discussed in Section 2.2 and the details will be given in following chapter.

Besides that, we also consider the parametric rational cubic Bézier curve as

$$Q(t) = \frac{A(1-t)^3 + B\Omega 3t(1-t)^2 + C\theta 3t^2(1-t) + Dt^3}{(1-t)^3 + \Omega 3t(1-t)^2 + \theta 3t^2(1-t) + t^3}, \qquad 0 \leq t \leq 1, \tag{3.6}$$

where $\Omega, \theta > 0$ and $A, B, C, D \in \Re^2$.

Corollary 3.2: Let $Q(t)$ be given by (3.6), where $\Omega, \theta > 0$, also $A, B, C, D \in \Re^2$, and A, D lie on one side of a given line while B and/ or C lie on the opposite side of the line. If $Q(t)$ crosses the given line (respectively, touches the line at only one point), then

$$c^2 > \frac{\Omega bd}{\theta^2} \quad \text{and} \quad b^2 > \frac{\theta ac}{\Omega^2} \tag{3.7}$$

and

$$3\Omega^2\theta^2 b^2 c^2 + 6\Omega\theta abcd - 4(\theta^3 ac^3 + \Omega^3 b^3 d) - a^2 d^2 > 0, \tag{3.8}$$

where a, b, c, d are signed distances of the points A, B, C, D from the given line.

Inequality (3.8) is used to check whether the parametric rational cubic Bézier curve $Q(t)$ crosses over the given line. If (3.8) holds then $Q(t)$ has crosses the given line and it can be modified by scaling its weights Ω and θ by a factor less than 1 in order to make $Q(t) \geq 0$. The details will be discussed in Chapter 5.

CHAPTER 4

CONSTRAINED INTERPOLATION USING RATIONAL CUBIC CURVE

4.1 Introduction

Suppose planar data points are given that all lie on one side of one or more given lines. Goodman, et al. (1991) developed an interpolation scheme for generating a curve which interpolates these data and also lies on the same side of each given lines. Such an interpolation scheme is said to be local. The scheme is based on the piecewise rational cubic scheme described in (Goodman, 1988). In addition to the shape preserving property, this rational scheme has a number of other desirable properties. It allows for the reproduction of conics, stable, invariant under a rotation of the coordinate axes or change in scale, and has in general G^2 continuity. The continuity may only be G^1 or possibly C^0, in certain situation involving collinear data. The curve is then constructed by piecing together parametric rational cubic and straight line segments.

4.2 Estimation of Curvatures and Tangents

Before we generate the parametric rational cubic, other than the given data points we need to estimate the curvatures and tangents which determine the inner control points B, C and the weights α, β as the initial values.

4.2.1 Estimation of Curvatures

Let us consider again the parametric rational cubic in (3.3) where $\alpha, \beta > 0$ and $A,\ B,\ C,\ D \in \Re^2$. From (3.3), differentiate $R(t)$ and we have

$$R'(t) = \frac{W(t)P'(t) - P(t)W'(t)}{W^2(t)},$$

where

$$P'(t) = -3A\alpha(1-t)^2 + B(1-4t+3t^2) + C(2t-3t^2) + 3D\beta t^2,$$

$$W'(t) = -3\alpha(1-t)^2 + (1-4t+3t^2) + (2t-3t^2) + 3\beta t^2.$$

Clearly,

$$R(0) = A,$$

$$R(1) = D,$$

$$R'(0) = \frac{(B-A)}{\alpha}, \tag{4.1}$$

$$R'(1) = \frac{(D-C)}{\beta}, \tag{4.2}$$

Now we shall derive the second order derivative of $R(t)$

$$R''(t) = \frac{\{P''(t)W(t) - W''(t)P(t)\}(W(t)) - \{P'(t)W(t) - W'(t)P(t)\}(2W'(t))}{W^3(t)}$$

where

$$P''(t) = 6A\alpha(1-t) + B(-4+6t) + C(2-6t) + 6D\beta t,$$

$$W''(t) = 6\alpha(1-t) + (-4+6t) + (2-6t) + 6\beta t.$$

At $t = 0$,

$$R''(0) = \frac{[2(C-B)+4(B-A)](\alpha)-2(B-A)}{\alpha^2}. \tag{4.3}$$

At $t = 1$,

$$R''(1) = \frac{-2(C-B)\beta-4(D-C)\beta+2(D-C)}{\beta^2}. \tag{4.4}$$

Let $\kappa(t)$ be the curvature of the rational curve (3.3). By using (4.1) and (4.3) the values of curvature at A is

$$\kappa(0) = \frac{R'(0)\times R''(0)}{|R'(0)|^3} = \frac{2\alpha[(B-A)\times(C-B)]}{|B-A|^3}. \tag{4.5}$$

Similarly, by using (4.2) and (4.4) the value of curvature at D is

$$\kappa(1) = \frac{R'(1)\times R''(1)}{|R'(1)|^3} = \frac{2\beta[(C-B)\times(D-C)]}{|D-C|^3}. \tag{4.6}$$

We refer to (Goodman, 1988), suppose $I_i = (x_i, y_i)$, $i = 1,2,....,N$, $N \geq 3$ are the given data points in the plane. The curvature κ_i at each interpolation point I_i are assigned to be curvature of the circle passing through I_{i-1}, I_i, I_{i+1}, i.e.

$$\kappa_i = \frac{2(I_i - I_{i-1})\times(I_{i+1}-I_i)}{|I_i - I_{i-1}||I_{i+1}-I_i||I_{i+1}-I_{i-1}|}, \tag{4.7}$$

which allows for the possibility of reproducing circular arcs.

For an open curve, the curvature κ_1 at I_1 is defined as the curvature of the circular arc passing through I_1, I_2, I_3 except when I_1, I_2, I_3 are collinear, in which case κ_i=0. The curvature κ_N at I_N is similarly defined. For a closed curve, $I_{N+j} = I_j$ and $I_{1-j} = I_{N+1-j}$ for $j = 1, 2$.

4.2.2 Estimation of Tangents

From (Goodman, 1988), the tangent direction T_i at each interpolation point I_i is assigned to be

$$T_i = a_i(I_i - I_{i-1}) + b_i(I_{i+1} - I_i), \tag{4.8}$$

where

$$a_i = |\kappa_{i+1}||I_{i+1} - I_i|^2,$$

$$b_i = |\kappa_{i-1}||I_i - I_{i-1}|^2.$$

T_i is in the direction of $I_{i+1} - I_i$ if and only if I_i, I_{i+1}, I_{i+2} are collinear, and is in the direction of $I_i - I_{i-1}$ if and only if I_i, I_{i-1}, I_{i-2} are collinear. If $I_{i-2}, \ldots, I_{i+2}$ lie on a circular arc, then (4.8) gives a tangent which has the same direction as that of this circular arc.

For an open curve, if I_1, I_2, I_3 are collinear, then the tangent direction T_1 at I_1 is in the same direction as $I_2 - I_1$; otherwise T_1 is assigned to be the tangent to the circle which passes through I_1, I_2, I_3. The tangent direction T_N at I_N is similarly defined.

4.3 Determination of Control Points and the Weights

After the curvatures and tangent vectors have been assigned at each interpolation point as (4.7) and (4.8), data points are then joined by a parametric rational cubic in (3.3). In the case where three or more consecutive data points are collinear, straight line segments through these points are used instead of the rational cubic. Each curve segment is determined as follows.

For $i = 1,...,N-1$ (respectively, $i = 1,...,N$), the i th curve segment $R_i(t)$ between I_i and I_{i+1} of an open curve (respectively, a closed curve) is defined as below.

(a) If $\kappa_i \kappa_{i+1} = 0$, then

$$R_i(t) = (1-t)I_i + tI_{i+1} , \qquad 0 \le t \le 1,$$

which is a straight line segment.

(b) If $\kappa_i \kappa_{i+1} \ne 0$, to determine the curve segment in (3.3) we need to determine the inner Bézier points B, C and the weights α, β. There are two cases that is $R_i(t)$ may be a convex segment or an inflection segment.

(i) If $\kappa_i \kappa_{i+1} > 0$ (convex segment), from (Goodman, 1988) we defined $|B-A|, |D-C|$ as

$$|B-A| = \frac{2|I_{i+1}-I_i||\sin b|}{2\lambda|\sin b| + (1-\lambda)|I_{i+1}-I_i||\kappa_{i+1}| + 2|\sin(a+b)|}, \qquad 0 < \lambda \le 1,$$

$$|D-C| = \frac{2|I_{i+1}-I_i||\sin a|}{2\mu|\sin a| + (1-\mu)|I_{i+1}-I_i||\kappa_i| + 2|\sin(a+b)|}, \qquad 0 < \mu \le 1,$$

where

$$\sin a = \frac{T_i \times (I_{i+1}-I_i)}{|T_i||I_{i+1}-I_i|},$$

$$\sin b = \frac{(I_{i+1}-I_i) \times T_{i+1}}{|I_{i+1}-I_i||T_{i+1}|},$$

and

$$\sin(a+b) = \frac{T_i \times T_{i+1}}{|I_{i+1}-I_i||T_{i+1}|}.$$

λ and μ are chosen to be $\lambda = \mu = 0.5$ through other values could also be used.

(ii) If $\kappa_i \kappa_{i+1} < 0$ (inflection segment), then the inner points B and C lie on opposite sides of the line joining A and D, and the curve segment will have a single point of inflection. We define

$$|B-A| = \gamma |I_{i+1} - I_i|,$$

$$|D-C| = \delta |I_{i+1} - I_i|, \quad 0 < \gamma, \delta < 0.5.$$

The bounds on γ and δ are to ensure that the curve segment does not exhibit any sharp turn. Here the values of γ and δ are chosen to be 0.25 each (Goodman, 1988).

The points B and C may be obtained by (4.1), (4.2) and (4.8),

$$B = A + |B-A| \frac{T_i}{|T_i|},$$

$$C = D + |D-C| \frac{T_{i+1}}{|T_{i+1}|},$$

and α, β are determined by (4.5), (4.6) and (4.7),

$$\alpha = \frac{\kappa_i |B-A|^3}{2[(B-A) \times (C-B)]},$$

$$\beta = \frac{\kappa_{i+1} |D-C|^3}{2[(C-B) \times (D-C)]}.$$

4.4 Curve Modification

After determine the inner control points B, C and the weights α, β then an initial interpolating curve is first generated by the rational cubic curve. Each curve segment of this default curve generated is then tested by the criteria as given in Corollary 3.1 to determine which segments have crossed over the given constrained line.

We shall refer these curve segments as "bad" segment. When the bad segments are identified, the signed distances of the points *A*, *B*, *C*, *D* from the line which are *a*, *b*, *c*, *d* satisfy strictly the inequalities (3.4) and (3.5). We would like to scale the weights α, β with factor λ and μ in such a way that

$$b^2c^2\lambda^2\mu^2 + 18\alpha\beta abcd\lambda\mu - 4(\alpha ac^3\lambda\mu^2 + \beta b^3 d\lambda\mu^2) - 27\alpha^2\beta^2a^2d^2 = 0,$$

and $0 < \lambda \leq 1$, $0 < \mu \leq 1$. This implies that the segment would just touch the given constraint line. There are three cases to be considered.

4.4.1 Case $b < 0$ and $c \geq 0$

In this case the inner Bézier point B lies on the other side of the given constraint line but C does not. This case is shown at below in Figure 4.1. So we shall choose $\mu = 1$, i.e. the weight β is kept fixed while only α is being scaled.

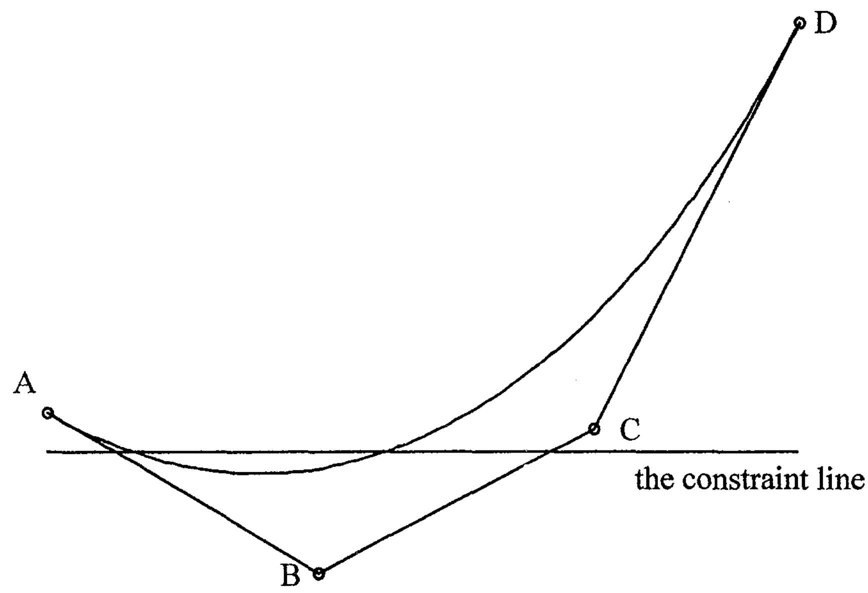

Figure 4.1 Initial curve crossed the constraint line for case $b < 0$, $c \geq 0$.

We then require $f(\lambda)=0$ where

$$f(\lambda)=(b^2c^2-4\beta b^3 d)\lambda^2+(18\alpha\beta abcd-4\alpha ac^3)\lambda-27\alpha^2\beta^2a^2d^2. \quad (4.9)$$

Observe that when $\lambda=0$,

$$f(0)=-27\alpha^2\beta^2a^2d^2<0.$$

If $f(1)<0$, then it means $R(t)$ does not crosses over the line. According to Corollary 3.1, if $R(t)$ crosses over the given constrained line, then $f(1)>0$. By the Intermediate Value Theorem, there exist $\lambda\in(0,1)$ for which $f(\lambda)=0$. From (4.9), we observe that when $f(\lambda)\geq 0$ then

$$[(b^2c^2-4\beta b^3 d)\lambda+(18\alpha\beta abcd-4\alpha ac^3)]>27\alpha^2\beta^2a^2d^2>0. \quad (4.10)$$

If we consider the derivative of $f(\lambda)$,

$$f'(\lambda)=2(b^2c^2-4\beta b^3 d)\lambda+18\alpha\beta abcd-4\alpha ac^3$$

and compare to (4.10) then we get $f'(\lambda)>0$. Hence f is increasing whenever $f(\lambda)\geq 0$ on [0, 1]. This implies that $f(\lambda)$ has a unique positive root λ in (0,1).

4.4.2 Case $b\geq 0$ and $c<0$

This case is similar to Section 4.4.1 and it is shown in Figure 4.2. We shall choose $\lambda=1$ and require $g(\mu)=0$ where

$$g(\mu)=(b^2c^2-4\alpha ac^3)\mu^2+(18\alpha\beta abcd-4\beta b^3 d)\mu-27\alpha^2\beta^2a^2d^2.$$

As Section 4.4.1, $g(0)<0$, $g(1)>0$ and g is increasing whenever $g(\mu)\geq 0$ on [0, 1]. Thus μ to be the unique root of g in (0, 1).

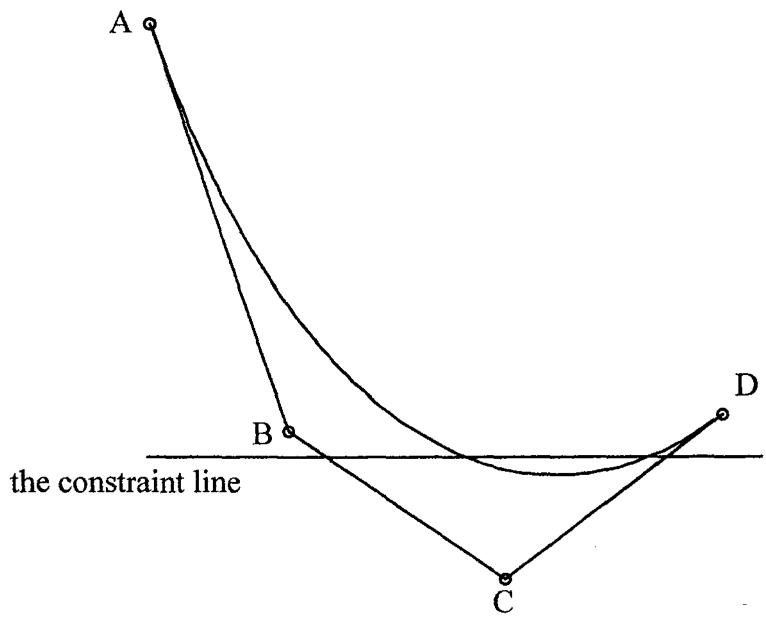

Figure 4.2 Initial curve crossed the constraint line for case $b \geq 0$, $c < 0$.

4.4.3 Case $b < 0$ and $c < 0$

In this case (see Figure 4.3), we shall scale both of the weights α and β by the same scaling factor. Let $\mu = \lambda$, we would require $h(\lambda) = 0$ where

$$h(\lambda) = b^2c^2\lambda^4 - 4(\alpha ac^3 + \beta b^3 d)\lambda^3 + 18\alpha\beta abcd\lambda^2 - 27\alpha^2\beta^2a^2d^2.$$

When $\lambda = 0$,

$$h(0) = -27\alpha^2\beta^2a^2d^2 < 0,$$

and $\lambda = 1$, $R(t)$ crosses over the line then $h(1) > 0$. Differentiate the $h(\lambda)$, we have

$$h'(\lambda) = 4b^2c^2\lambda^3 - 12(\alpha ac^3 + \beta b^3 d)\lambda^2 + 36\alpha\beta abcd\lambda.$$

Observe that $h'(\lambda) > 0$, if $\lambda > 0$. Thus, h is increasing for $\lambda > 0$. By Intermediate Value Theorem, there exist a unique root $\lambda \in (0,1)$ of $h(\lambda) = 0$. We choose λ to be the unique root of $h(\lambda) = 0$ in (0, 1). This value can be obtained by solving the equation numerically.

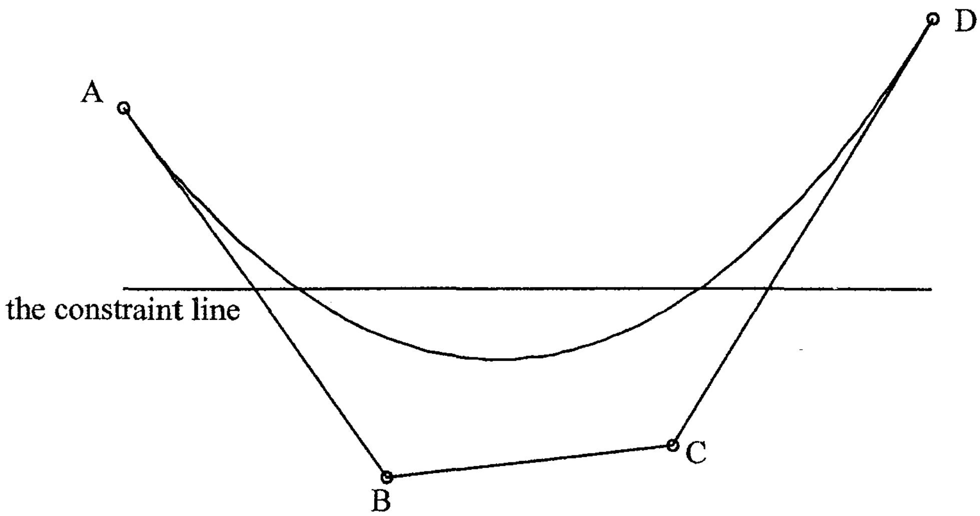

Figure 4.3 Initial curve crosses the constraint line for case $b < 0$ and $c < 0$.

4.5 Graphical Examples

We shall illustrate our discussion with five examples. The first three examples are showing the cases in Sections 4.4.1, 4.4.2 and 4.4.3 respectively. The fourth example is a closed curve with the arbitrary line as the constraint line, while last example shows a closed curve with several constraint lines. In these examples, the data points are marked by "o" and the two inner Bézier points of each curve segments are marked by "*". The initial curve is drawn in dotted line form and the modified curve is drawn in full line form. The corresponding data points of each example and the related scaling factors are given respectively in the following tables.