ELSEVIER
Journal of Computational and Applied Mathematics 134 (2001) 69–83
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

Curvature extrema of planar parametric polynomial cubic curves

D.J. Walton*, D.S. Meek
Department of Computer Science, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2

Received 2 April 1999; received in revised form 3 April 2000

Abstract

Parametric polynomial cubic curve segments are widely used in computer-aided design and computer-aided geometric design applications because their flexibility makes them suitable for use in the interactive design of curves and surfaces. Shape properties of these segments have been studied and results on the occurrence of cusps, and whether the segment is S- or C-shaped, are available in the literature. Their critical points, however, are not as well known. This paper presents results on the number and location of curvature extrema of these segments. All possible numbers of curvature extrema are found and it is demonstrated that all possible curvature extrema can be obtained numerically for any planar parametric cubic curve. These results are useful in the study of the fairness of curves designed with parametric polynomial cubic curve segments. © 2001 Elsevier Science B.V. All rights reserved.

MSC: 68U07; 65Y25

Keywords: Curvature extrema; Cubic Bézier curves; Fair curves

1. Introduction

Parametric cubic curve segments are widely used in computer-aided design (CAD) and computer-aided geometric design (CAGD) applications. They commonly occur as Bézier or B-spline representations [4,6]. They are popular mainly because

- they are polynomials of low degree, hence low overhead in implementation;
- they are flexible, hence suitable for use in interactive design of curves and surfaces.

* Corresponding author. Tel.: +1-204-474-8679; fax: +1-204-474-7609.
E-mail addresses: walton@cs.umanitoba.ca (D.J. Walton), dmeek@cc.umanitoba.ca (D.S. Meek).

PII: S0377-0427(00)00529-X

The discussion in this article is restricted to parametric cubics. A significant disadvantage of a parametric cubic curve is that its curvature is a complicated function of its parameter. It is thus not easy to use cubic curve segments in the design of curvature controlled curves. Control of curvature is important in the design of fair curves [3,11]. Fair curves are important in many CAD and CAGD applications [11]. They are also important in applications such as highway design, or the design of robot trajectories. For example, it is desirable that a transition curve between two circular arcs in the horizontal layout of highway design be free of curvature extrema, (except at its endpoints) [1].

The curvatures of some spirals, e.g., the Cornu spiral [8,9], and the logarithmic spiral [2], are simple functions of their parameters and are thus more easily controlled than the curvature of a parametric cubic curve. Unfortunately, such curves are not polynomial, which usually means more overhead on implementation. They are also not as flexible as cubic curve segments. Furthermore, many existing CAD software packages are based on NURBS (nonuniform rational B-splines), hence addition of nonpolynomial-based curve drawing features may not be feasible.

The lack of ease of control of the curvature of a parametric cubic curve segment in geometric design has been addressed by

- fairing of curves [3,11], and
- identifying which segments of polynomial curves have monotone curvature, and using such segments in the design of curves and surfaces [14–17].

According to Farin [3], "a curve is fair if its curvature plot consists of relatively few monotone pieces". Methods for curve fairing typically depend on an examination of the curvature plot. Techniques such as knot removal, adjustment of data points, degree elevation and reduction, are then applied to improve the curvature plot, i.e., reduce the number of monotone pieces. The process is usually iterative and may change the original curve.

An alternative to fairing of curves is to design with fair curves. In this approach, the designer works with curve segments of monotone curvature (i.e., spiral segments) and fits them together to form a curve whose curvature plot has relatively few monotone pieces. The advantage of this method is that the designer knows a priori that the segments of the curve have monotone curvature and can thus adjust them interactively to obtain a desired curve; a posteriori examination of curvature plots, or fairing, is not necessary and the curve need not be changed afterwards to make it fair [13,16]. The disadvantage of this approach is that such spiral segments are not as flexible as the usual NURBS curves, and thus not always as easy to manipulate.

In the design of fair curves, or in the fairing of curves, knowledge of the number and location of curvature extrema of the curves is useful. Earlier work focused on curvature bounds for planar cubic B-spline segments [12]. The purpose of this paper is to report on results of work done in determining the number and location of curvature extrema of planar parametric cubic curves. Care should thus be taken in transformations which are assumed for simplifying the analysis. Affine transformations, for example, do not preserve the number of curvature extrema. In this article, only translations, rigid rotations, uniform scaling, and linear re-parameterisations are used, to ensure that the number of curvature extrema and their relative positions are not inadvertently altered.

Wang [18] studied shape properties of parametric cubic curves, Patterson [10] studied parametric cubics as algebraic curves; this approach is convenient to determine some results for which an nalysis using the parametric representation is tedious. A cubic function curve may be defined as a curve given by $y = p(x)$, where $p(x)$ is a cubic polynomial in x. Neither of the two studies, [18,10], consider the nondegenerate cubic function $y = p(x)$ as a special case of a parametric cubic. This cubic always has a single inflection point and two curvature extrema.

The remainder of this paper is organised as follows. Section 2 gives a brief discussion of the notation and conventions used in this paper. Some theoretical background is presented in Section 3. Results for the cubic function as a special case are derived in Section 4. In Section 5 some known shape properties of cubics are reviewed. Results on curvature extrema of parametric cubics are presented in Section 6; these results are, to the authors' knowledge, new and not yet available in the literature. Illustrative examples are shown in Section 7. Concluding remarks are made in Section 8.

2. Notation and conventions

The usual Cartesian coordinate system with x- and y-axis is presumed. Points, that correspond to ordered pairs in the Cartesian coordinate system, and vectors are shown as bold letters. Points and vectors may also be indicated using the ordered pair notation e.g. (x, y). In particular, the components of a vector $\boldsymbol{A}$ may be denoted as (A_x, A_y), or in the case of a subscripted vector $\boldsymbol{V}_0$, as $(V_{0,x}, V_{0,y})$. The dot product of two vectors, $\boldsymbol{A}$ and $\boldsymbol{B}$ is denoted as $\boldsymbol{A} \cdot \boldsymbol{B}$. The norm or length of a vector $\boldsymbol{A}$ is denoted as $\|\boldsymbol{A}\|$. The derivative of a function (scalar or vector valued), is denoted with a prime, e.g., $\boldsymbol{Q}'(t)$ or $x'(t)$.

A planar parametric curve is defined by the set of points $\boldsymbol{Q}(t) = (x(t), y(t))$ for t in some given interval of the real line. Both components of the zero vector are zeros. A cusp is formed at a point on a curve where the direction of the unit tangent vector changes through 180° [4, p. 132]. An inflection point is defined as a point at which the tangent to the curve intersects the curve [4, p. 381]; it occurs at a point where the curvature is zero, and changes sign in the neighbourhood of the point. It is known that the magnitude of the curvature of a parametric cubic tends to zero as the parameter tends to positive or negative infinity. While these may be considered as minima of the curvature magnitude, they are trivial cases of extrema and hence not counted. To avoid repetition, local extrema are simply referred to as extrema.

3. Theoretical background

The tangent vector of a plane parametric curve $\boldsymbol{Q}(t)$ is given by $\boldsymbol{Q}'(t)$. If $\boldsymbol{Q}'(t) \neq \boldsymbol{0}$, then the curvature of $\boldsymbol{Q}(t)$ is defined in [5, p. 28] as

$$\kappa(t) = \frac{x'(t)y''(t) - y'(t)x''(t)}{\|\boldsymbol{Q}'(t)\|^3}. \tag{3.1}$$

The curvature of a plane algebraic curve $y = p(x)$ is defined in [5, p. 28] as

$$\kappa(x) = \frac{p''(x)}{[1 + \{p'(x)\}^2]^{3/2}}. \tag{3.2}$$

The following facts are used several times throughout this paper.

- If a curve has a continuous tangent vector, then the tangent vector will be zero at the occurrence of a cusp. The curvature is consequently not defined at a cusp.
- A parametric curve intersects itself if there are distinct values of t, t_i and t_j, for which $\boldsymbol{Q}(t_i)=\boldsymbol{Q}(t_j)$.
- Local extreme values of a differentiable function $f(t)$ occur at points where $f'(t)$ changes sign [7, p. 290]. Local extrema thus occur at distinct zeros of $f'(t)$, or at those zeros of $f'(t)$ which have odd multiplicity.

A segment of parametric cubic curve and its derivative are, respectively, represented in Bézier form as [4]

$$\boldsymbol{Q}(t) = \boldsymbol{P}_0(1-t)^3 + 3\boldsymbol{P}_1(1-t)^2t + 3\boldsymbol{P}_2(1-t)t^2 + \boldsymbol{P}_3t^3, \quad 0 \leqslant t \leqslant 1, \tag{3.3}$$

and

$$\boldsymbol{Q}'(t) = 3\boldsymbol{W}_0(1-t)^2 + 6\boldsymbol{W}_1(1-t)t + 3\boldsymbol{W}_2t^2, \quad 0 \leqslant t \leqslant 1,$$

where $\boldsymbol{P}_i$, $i = 0,\ldots,3$, are the control points and

$$\boldsymbol{W}_i = \boldsymbol{P}_{i+1} - \boldsymbol{P}_i, \quad i = 0,1,2. \tag{3.4}$$

The whole curve, of which (3.3) is a segment, is obtained for $-\infty < t < \infty$. Subsequent discussion refers to this whole curve. The shape of the curve is not affected by translation, rotation, reflection about an arbitrary axis, or uniform scaling. Curvature is a curve intrinsic and a geometric invariant [4, p. 180]. Although uniform scaling changes the curvature, it preserves the number and relative positions of curvature extrema, cusps, inflection points, and self-intersections. Translation, rotation, uniform scaling, and reflection about an arbitrary axis will be introduced as needed to simplify algebraic expressions. These transformations do not affect the number and relative positions of curvature extrema, cusps, inflection points, and self-intersections.

Assume that $\boldsymbol{Q}(t)$ has been translated so that it passes through the origin at $t = 0$, i.e. $\boldsymbol{P}_0 = 0$, in which case (3.3) can be written as

$$\boldsymbol{Q}(t) = 3\boldsymbol{P}_1(1-t)^2t + 3\boldsymbol{P}_2(1-t)t^2 + \boldsymbol{P}_3t^3,$$

or in power basis form as

$$\boldsymbol{Q}(t) = 3\boldsymbol{P}_1t + 3(\boldsymbol{P}_2 - 2\boldsymbol{P}_1)t^2 + (\boldsymbol{P}_3 - 3\boldsymbol{P}_2 + 3\boldsymbol{P}_1)t^3.$$

An alternative representation using (3.4) is

$$\boldsymbol{Q}(t) = 3\boldsymbol{W}_0t + 3(\boldsymbol{W}_1 - \boldsymbol{W}_0)t^2 + (\boldsymbol{W}_0 - 2\boldsymbol{W}_1 + \boldsymbol{W}_2)t^3.$$

For a non-degenerate cubic, $\|\boldsymbol{W}_0 - 2\boldsymbol{W}_1 + \boldsymbol{W}_2\| \neq 0$. This condition is assumed in all subsequent discussion. It is also assumed that $\boldsymbol{Q}(t)$ is rotated so that

$$\boldsymbol{W}_0 - 2\boldsymbol{W}_1 + \boldsymbol{W}_2 = (0, \tfrac{1}{3}a), \quad a \neq 0.$$

Hence, $\boldsymbol{Q}(t)$ may be written as

$$\boldsymbol{Q}(t) = (x(t), y(t)) = 3\boldsymbol{W}_0t + 3(\boldsymbol{W}_1 - \boldsymbol{W}_0)t^2 + (0, \tfrac{1}{3}a)t^3. \tag{3.5}$$

Parametric cubics can now be classified into two broad categories, a special case for which $W_{1,x} - W_{0,x} = 0$, and a more general case for which $W_{1,x} - W_{0,x} \neq 0$.

4. The special case $W_{1,x} - W_{0,X} = 0$

This is the case of a cubic function curve. This special case can be treated without categorising it further. If $W_{0,x} = 0$ then $\boldsymbol{Q}(t)$ degenerates to the straight line $x = 0$, otherwise assume that $W_{0,x} > 0$; if not, a reflection across the y-axis will ensure this. Now, $\boldsymbol{Q}(t)$ can be written in the algebraic form

$$y = \tfrac{1}{3}\alpha_3 x^3 + \alpha_2 x^2 + \alpha_1 x, \tag{4.1}$$

where

$$x = 3W_{0,x}t \quad \text{and} \quad \alpha_3 \neq 0.$$

It is assumed that $\alpha_3 > 0$. If not, a reflection across the x-axis will ensure this. Uniform scaling by $1/\sqrt{\alpha_3}$ puts (4.1) in the form

$$y = \tfrac{1}{3}x^3 + \frac{\alpha_2}{\sqrt{\alpha_3}}x^2 + \alpha_1 x.$$

The curve can now be translated by $\alpha_2/\sqrt{\alpha_3}$ along the x-axis to yield

$$y = \tfrac{1}{3}x^3 + \beta_1 x + \beta_0, \tag{4.2}$$

where

$$\beta_1 = \alpha_1 - \frac{\alpha_2^2}{\alpha_3}$$

and

$$\beta_0 = \frac{\alpha_2}{\sqrt{\alpha_3}}\left(\frac{2}{3}\frac{\alpha_2^2}{\alpha_3} - \alpha_1\right).$$

The curvature of the resulting curve is obtained from (3.2) as

$$\kappa(x) = \frac{2x}{(x^4 + 2\beta_1 x^2 + \beta_1^2 + 1)^{3/2}}.$$

The curve (3.2) thus has a single inflection point at $x = 0$. The derivative of its curvature is

$$\kappa'(x) = -\frac{2f(x)}{(x^4 + 2\beta_1 x^2 + \beta_1^2 + 1)^{5/2}},$$

where

$$f(x) = 5x^4 + 4\beta_1 x^2 - \beta_1^2 - 1.$$

Now $f(x)$, as a quadratic in x^2, has a negative and a positive zero. Only the positive zero gives real values for x, so the curve has exactly two curvature extrema, at

$$x = \pm\sqrt{\tfrac{1}{5}(-2\beta_1 + \sqrt{9\beta_1^2 + 5})}.$$

This special case of the parametric cubic never has a cusp or a loop because $\kappa(x)$ is always defined, and there is exactly one y for each x.

5. The case $W_{1,x} - W_{0,x} \neq 0$

It is assumed that $W_{1,x} - W_{0,x} > 0$; if not, a reflection across the y-axis will ensure this. Let $\mu = 6(W_{1,x} - W_{0,x})$. Uniform scaling of $\boldsymbol{Q}(t)$ in (3.5) by a^2/μ^3 followed by some algebraic manipulation and re-arrangement produce

$$x = \tfrac{1}{2}\left(\frac{at}{\mu}\right)^2 + \lambda\left(\frac{at}{\mu}\right)$$

and

$$y = \tfrac{1}{3}\left(\frac{at}{\mu}\right)^3 + \gamma_2\left(\frac{at}{\mu}\right)^2 + \gamma_1\left(\frac{at}{\mu}\right),$$

where

$$\lambda = \frac{3aW_{0,x}}{\mu^2}.$$

By completing the square in the x-component, it can also be written as

$$x = \tfrac{1}{2}\left(\frac{at}{\mu} + \lambda\right)^2 - \tfrac{1}{2}\lambda^2.$$

To simplify subsequent algebraic expressions, reparameterise using

$$\sigma = \frac{at}{\mu} + \lambda$$

to obtain

$$x = \tfrac{1}{2}\sigma^2 - \tfrac{1}{2}\lambda^2$$

and

$$y = \tfrac{1}{3}(\sigma - \lambda)^3 + \gamma_2(\sigma - \lambda)^2 + \gamma_1(\sigma - \lambda).$$

Translation by $(\frac{1}{2}\lambda^2, \frac{1}{3}\lambda^3 - \gamma_2\lambda^2 + \gamma_1\lambda)$ puts the curve in the form

$$\boldsymbol{Q}(\sigma) = (\tfrac{1}{2}\sigma^2, \tfrac{1}{3}\sigma^3 + \tfrac{1}{2}b\sigma^2 + d\sigma) = (x(\sigma), y(\sigma)), \tag{5.1}$$

where

$$b = 2(\gamma_2 - \lambda)$$

and

$$d = \lambda^2 - 2\gamma_2\lambda + \gamma_1.$$

Observe the following from (5.1):

- $x(\sigma) \geqslant 0$, so the curve is entirely in the right half-plane.
- if $b = 0$ then $x(\sigma) = x(-\sigma)$ and $y(\sigma) = -y(-\sigma)$, i.e., the curve is symmetric about the x-axis when $b = 0$.
- for $b < 0$, $x(\sigma) = x(-\sigma)$ and $y(\sigma) = -\{\frac{1}{3}(-\sigma)^3 + \frac{1}{2}(-b)(-\sigma)^2 + d(-\sigma)\}$, i.e., the curve for $b < 0$ is a reflection about the x-axis of the curve obtained when b is replaced by $-b$ and σ replaced by $-\sigma$.

Based on these observations, for the remainder of the paper it is assumed that $b \geqslant 0$. From (5.1), the tangent vector and second derivative of $\boldsymbol{Q}(\sigma)$ is

$$\boldsymbol{Q}'(\sigma) = (x'(\sigma), y'(\sigma)) = (\sigma, \sigma^2 + b\sigma + d) \tag{5.2}$$

and

$$\boldsymbol{Q}''(\sigma) = (x''(\sigma), y''(\sigma)) = (1, 2\sigma + b) \tag{5.3}$$

It follows from (3.1), (5.2), and (5.3) that the curvature of $\boldsymbol{Q}(t)$ is

$$\kappa(\sigma) = \frac{\sigma^2 - d}{\{\sigma^4 + 2b\sigma^3 + (1 + b^2 + 2d)\sigma^2 + 2bd\sigma + d^2\}^{3/2}}, \tag{5.4}$$

where $\sigma \neq 0$ for $d = 0$.

The derivative of the curvature is obtained from (5.4) which, after some manipulation, can be expressed as

$$\kappa'(\sigma) = \begin{cases} \dfrac{-g(\sigma)}{\sigma|\sigma|(\sigma^2 + 2b\sigma + 1 + b^2)^{5/2}} & \text{for } d = 0 \text{ and } \sigma \neq 0 \\ \dfrac{-f(\sigma)}{\{\sigma^4 + 2b\sigma^3 + (1 + b^2 + 2d)\sigma^2 + 2bd\sigma + d^2\}^{5/2}} & \text{for } d \neq 0, \end{cases}$$

where

$$g(\sigma) = 4\sigma^2 + 5b\sigma + 1 + b^2, \quad d = 0, \tag{5.5}$$

and

$$f(\sigma) = 4\sigma^5 + 5b\sigma^4 + (1 + b^2 - 4d)\sigma^3 - 10bd\sigma^2 - (8d^2 + 3b^2d + 3d)\sigma - 3bd^2, \quad d \neq 0. \tag{5.6}$$

Simple conditions for the occurrence of a cusp, inflection points, or a self-intersection, of a parametric cubic can now be stated.

(A) If a cusp occurs, then $\boldsymbol{Q}'(t) = 0$. Form (5.2) this implies that $d = 0$ and $\sigma = 0$.

(B) It follows immediately from (5.4) that inflection points can only occur for $d > 0$; they occur as a pair at $\sigma = \pm\sqrt{d}$.

(C) Self-intersections (if any) occur at distinct values of σ, σ_1 and σ_2, that satisfy $\boldsymbol{Q}(\sigma_1) = \boldsymbol{Q}(\sigma_2)$. Using (5.1), the following two equations in σ_1 and σ_2 are obtained:

$$\tfrac{1}{2}\sigma_1^2 = \tfrac{1}{2}\sigma_2^2 \tag{5.7}$$

and

$$\tfrac{1}{3}\sigma_1^3 + \tfrac{1}{2}b\sigma_1^2 + d\sigma_1 = \tfrac{1}{3}\sigma_2^3 + \tfrac{1}{2}b\sigma_2^2 + d\sigma_2. \tag{5.8}$$

From (5.7), $(\sigma_1 - \sigma_2)(\sigma_1 + \sigma_2) = 0$, but σ_1 and σ_2 being distinct implies that $\sigma_1 - \sigma_2 \neq 0$, and $\sigma_1 \neq 0$, $\sigma_2 \neq 0$. Hence, $\sigma_2 = -\sigma_1 \neq 0$. Substitution of this into (5.8) yields $\sigma_1^2 + 3d = 0$, which has two real solutions if $d < 0$, namely

$$\sigma_1 = \pm\sqrt{-3d}. \tag{5.9}$$

The following theorem provides an alternative derivation of results presented by Wang [18]. It also includes the special case of the cubic function curve which is not covered in [18]. Parametric cubics of this case (i.e., satisfying $W_{1,x} - W_{0,x} \neq 0$) can be classified into three sub-categories according as $d = 0$, $d < 0$, and $d > 0$. It is shown in Theorem 1 that these sub-categories correspond to the cases of the cubic having a cusp, a self-intersection (loop or crunode), or two distinct inflection points (acnode) as described in Wang [18] and Patterson [10].

Theorem 1. *Any non-degenerate parametric cubic whose standard from is given in* (5.1) *falls into one of the following four classes*:

(i) *It has a single inflection point, no cusps, and no points of self-intersection.*
(ii) *It has a single cusp, no points of self-intersection and no inflection points.*
(iii) *It has a point of self-intersection at* $\sigma = \pm\sqrt{-3d}$, *no inflection points, and no cusps.*
(iv) *It has two inflection points at* $\sigma = \pm\sqrt{d}$, *no cusps and no points of self-intersection.*

Proof.

(i) This is the special case, $W_{1,x} - W_{0,x} = 0$, of Section 4, the cubic function curve.
(ii) This is the case, $W_{1,x} - W_{0,x} \neq 0$ with $d = 0$. From condition (A) above, a cusp could occur at $\sigma = 0$. At $\sigma = -\varepsilon$ the unit tangent vector is parallel to $(-1, -b + \varepsilon)$ and in the limits as $\varepsilon \to 0$, it is parallel to $(-1, -b)$. On the other hand, at $\sigma = \varepsilon$, the unit tangent vector is parallel to $(1, b + \varepsilon)$ and in the limit as $\varepsilon \to 0$, it is parallel to $(1, b)$. Hence, the unit tangent vector changes orientation by 180° when σ goes through zero. Thus, there is a cusp at $\sigma = 0$. Inflection points do not occur for $d = 0$ by (B) above. Points of self-intersection do not occur for $d = 0$ by (C) above.
(iii) This is the case, $W_{1,x} - W_{0,x} \neq 0$ with $d < 0$. Self-intersection occurs for values of σ that satisfy (5.9); since there are two such distinct parametric values, there is a single point of self-intersection, forming a single loop. The curve has a crunode according to Patterson [10]. Cusps do not occur for $d \neq 0$ by (A). Inflection points do not occur for $d < 0$ by (B).
(iv) This is the case, $W_{1,x} - W_{0,x} \neq 0$ with $d > 0$. From (B), the curve has two inflection points at $\sigma = \pm\sqrt{d}$. Self-intersection does not occur for $d > 0$ by (C). A cusp does not occur by (A). □

The various shapes that a parametric cubic curve segment may assume are illustrated in Fig. 1. A dot (small black-filled circle) marks the beginning of each curve. One of the curvature extrema in (g) looks like a cusp in the figure. The reason for this effect is the very high curvature at that point. Control polygons are shown in dashed lines. The various curves are labelled as follows:

(a) A curve with a single inflection point.
(b) Part of a self-intersecting curve with a single curvature extremum.
(c) An affine transformation of (b) obtained by a scaling of 2 in the x-direction and 1 in the y-direction; it has three curvature extrema.
(d) A self-intersecting curve with one curvature extremum in the loop, and two curvature extrema outside the loop.

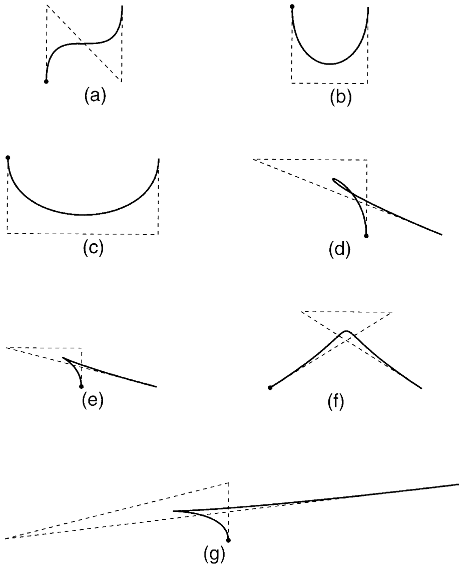

Fig. 1. Various shapes of a parametric cubic curve.

(e) A curve with a cusp and two curvature extrema.
(f) A curve with two inflections and three curvature extrema.
(g) A curve with two inflections and five curvature extrema.

The control points used to generate the curves, as well as the parametric values at which a cusp, self-intersections, inflections, and curvature extrema (if any) occur, are given in Section 7.

6. Curvature extrema for the case $W_{1,x} - W_{0,x} \neq 0$

Curvature extrema occur at places where the derivative of the curvature changes sign. The derivative of the curvature changes sign at the real zeros of odd multiplicity of $g(\sigma)$ as defined by (5.5) for $d = 0$, or $f(\sigma)$ as defined by (5.6) for $d \neq 0$. Differentiation of $f(\sigma)$ yields

$$f'(\sigma) = 20\sigma^4 + 20b\sigma^3 + 3(1 + b^2 - 4d)\sigma^2 - 20bd\sigma - (8d^2 + 3b^2d + 3d).$$

Note that $f'(\sigma)$ may be factored as

$$f'(\sigma) = (\sigma^2 - d)(20\sigma^2 + 20b\sigma + 3 + 3b^2 + 8d). \tag{6.1}$$

The zeros of the second quadratic factor of $f'(\sigma)$ are

$$\xi_1 = \frac{-5b - \sqrt{5(2b^2 - 8d - 3)}}{10} \tag{6.2}$$

and

$$\xi_2 = \frac{-5b + \sqrt{5(2b^2 - 8d - 3)}}{10}. \tag{6.3}$$

Observe that for $2b^2 - 8d - 3 \leqslant 0$, $f'(\sigma)$ does not change sign if $d < 0$, and changes sign at only two places if $d > 0$ ($at \pm \sqrt{d}$). For $2b^2 - 8d - 3 > 0$, the zeros ξ_1 and ξ_2 are real and distinct and $\xi_1 < 0$.

If the curve is symmetric, $b = 0$, then it follows from (5.6) that the numerator of the derivative of the curvature is

$$f(\sigma) = \sigma\{4\sigma^4 + (1 - 4d)\sigma^2 - (8d^2 + 3d)\}. \tag{6.4}$$

So, for $d \neq 0$, $f(\sigma)$ has a distinct zero at $\sigma = 0$. The zeros of the second factor in (6.4) as a quadratic in σ^2 are

$$\zeta_1 = \frac{4d - 1 - \sqrt{144d^2 + 40d + 1}}{8} \tag{6.5}$$

and

$$\zeta_2 = \frac{4d - 1 + \sqrt{144d^2 + 40d + 1}}{8}. \tag{6.6}$$

Observe that

$$144d^2 + 40d + 1 > 0 \quad \text{for } d < -\tfrac{1}{4} \text{ or } d > -\tfrac{1}{36} \tag{6.7}$$

and that

$$144d^2 + 40d + 1 > (4d - 1)^2$$

is equivalent to

$$8d^2 + 3d > 0.$$

So

$$144d^2 + 40d + 1 > (4d - 1)^2 \quad \text{for } d < -\tfrac{3}{8} \text{ or } d > 0. \tag{6.8}$$

Now if ζ_1 is real, then if $d < 0$, then $\zeta_1 < 0$; also, if $d > 0$, then from (6.8), $\zeta_1 < 0$. On the other hand, from (6.7) and (6.8), ζ_2 is real and positive for $d < -\frac{3}{8}$ or $d > 0$ in which case the second factor of $f(\sigma)$ in (6.4) has two real zeros at $\pm\sqrt{\zeta_2}$.

Fig. 2 illustrates curvature plots showing extrema corresponding to the various parametric cubic curve segments in Fig. 1. Curves (d), (e) and (g) have very large variations in curvature; the magnitudes of their curvatures were truncated. In Examples (a), (f) and (g), zero curvature on curvature plots is shown as a dashed line. Results on curvature extrema for the three sub-categories of the case $W_{1,x} - W_{0,x} \neq 0$ are now presented.

Theorem 2. *If the parametric cubic* (5.1) *has a cusp then it has zero or two curvature extrema*; *in the latter case*, *both extrema are on the same side of the cusp.*

Proof. This is the case $d = 0$. Observe that the quadratic $g(\sigma)$ as given by (5.5) does not change sign for $b \leqslant \frac{4}{3}$, in which case the curvature as given by (5.4) is nondecreasing as σ goes from

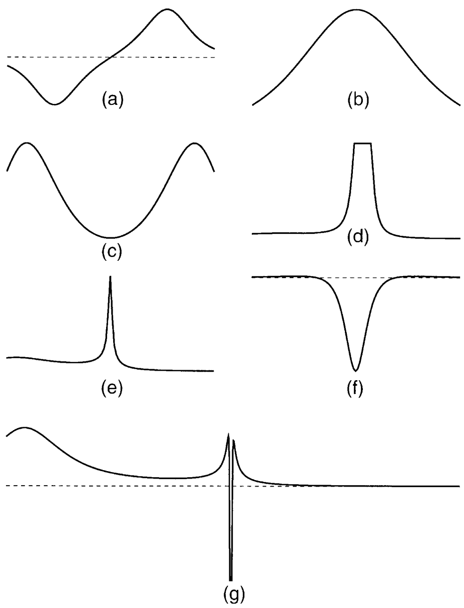

Fig. 2. Curvature plots for corresponding shapes in Fig. 1.

almost $-\infty$ to almost 0, and nondecreasing as σ goes from almost 0 to almost ∞. Hence there are no extrema in the intervals $(-\infty, 0)$ and $(0, \infty)$ for $b \leqslant \frac{4}{3}$. For $b > \frac{4}{3}$, $g(\sigma)$ changes sign twice at $\sigma = (-5b \pm \sqrt{9b^2 - 16})/8$, so in this case the curve with a cusp has two extrema (at these two values of σ). Note that both extrema occur at negative values of σ, i.e. on the same side of the cusp which occurs at $\sigma = 0$. □

Corollary. *A symmetric parametric cubic with a cusp has no curvature extrema.*

Proof. For $b = 0$ and $\sigma \neq 0$, $g(\sigma)$ in the proof of Theorem 2 is nonzero. □

Theorem 3. *A self-intersecting parametric cubic (i.e., with a loop) has one or three curvature extrema. If it has a single curvature extremum, the extremum occurs in the interval* $(0, \sqrt{-3d})$, *i.e., in the loop. If it has three curvature extrema, they occur in the intervals* $(-\infty, \xi_1)$, $(\xi_1, \min(0, \xi_2))$, *and* $(\max(0, \xi_2), \sqrt{-3d})$ *where* ξ_1 *and* ξ_2 *are given by* (6.2) *and* (6.3), *i.e., at least one of the extrema is in the loop.*

Proof. This is the case $d < 0$, so it follows from (6.1) that $f'(\sigma)$ changes sign zero or two times. Hence $f(\sigma)$, defined by (5.6), changes sign at most three times. Since $f(\sigma)$ is of odd degree, it changes sign an odd number of times, so its curvature has one or three extreme values.

It follows from Theorem 1 that the curve intersects itself at $\sigma = \pm\sqrt{-3d}$. Now $f(0) < 0$ and $f(\sqrt{-3d}) = 72bd^2 + (40d^2 - 6b^2d - 6d)\sqrt{-3d} > 0$, so at least one of the extrema is in the interval $(0, \sqrt{-3d})$, i.e., in the loop.

Consider first the case $2b^2 - 8d - 3 \leqslant 0$. Recall (from the observations at the beginning of this section) that for this case $f'(\sigma)$ does not change sign, so $f(\sigma)$ will change sign once and there will be a single curvature extremum.

Consider now the case $2b^2 - 8d - 3 > 0$. Note that, for $d < 0$, and $b \geqslant 0$,

$$15b^2 - 260d + 100\sqrt{-3d} + 15 > 0$$

Therefore,

$$-5b + \sqrt{5(2b^2 - 8d - 3)} < 10\sqrt{-3d}.$$

So, from (6.2) and (6.3), $\xi_1 < \xi_2 < \sqrt{-3d}$. The zeros of $f(\sigma)$ are separated by its extreme points which occur at the distinct zeros of $f'(\sigma)$ namely ξ_1 and ξ_2 as given by (6.2) and (6.3). So if $f(\sigma)$ is zero three times, then the zeros of $f(\sigma)$ are in $(-\infty, \xi_1)$, $(\xi_1, \min(0, \xi_2))$, and $(\max(0, \xi_2), \sqrt{-3d})$. □

Corollary. *A symmetric self-intersecting parametric cubic has one or three curvature extrema, all of which occur in the loop. One extremum is at* $\sigma = 0$. *If there are three, the remaining two are at* $\pm\sqrt{\zeta_2}$ *where* ζ_2 *is given by* (6.6).

Proof. It is clear from (6.4) that an extremum occurs at $\sigma = 0$. It also follows from (6.5) to (6.8) that, for $d < -\frac{3}{8}$, two more extrema occur at $\pm\sqrt{\zeta_2}$. Using some algebraic manipulation it can be verified that $\zeta_2 < -3d$, so if there are three extrema, they all occur in the loop. □

Theorem 4. *A parametric cubic with two inflection points has three or five curvature extrema. If it has three curvature extrema, they occur in the intervals* $(-\infty, -\sqrt{d})$, $(-\sqrt{d}, \sqrt{d})$ *and* $(\sqrt{d}, \infty)$; *if there are five curvature extrema, then they are in the intervals* $(-\infty, \xi_1)$, (ξ_1, ξ_2), $(\xi_2, -\sqrt{d})$, $(-\sqrt{d}, \sqrt{d})$ *and* $(\sqrt{d}, \infty)$. *In both cases there is exactly one extremum between the two inflection points.*

Proof. This is the case $d > 0$. It follows from (5.6) that

$$\lim_{\sigma \to -\infty} f(\sigma) < 0, \tag{6.9}$$

$$\begin{aligned} f(-\sqrt{d}) &= 2d\sqrt{d}(1 + 4d + b^2) - 8bd^2 \\ &= 2d\sqrt{d}[(b - 2\sqrt{d})^2 + 1] > 0, \end{aligned} \tag{6.10}$$

$$f(\sqrt{d}) = -2d\sqrt{d}(1 + 4d + b^2) - 8bd^2 < 0, \tag{6.11}$$

and

$$\lim_{\sigma \to \infty} f(\sigma) > 0. \tag{6.12}$$

Thus, $f(\sigma)$ changes sign at least three times. From (6.10) and (6.11) it follows that one of the sign changes occurs between the two inflection points. If $2b^2 - 8d - 3 \leqslant 0$, the second factor in (6.1) does not change sign, so $f'(\sigma)$ changes sign twice. In this case, $f(\sigma)$ changes sign exactly three times. If $2b^2 - 8d - 3 > 0$, the second factor in (6.1) has two distinct real zeros ξ_1 and ξ_2 given in (6.2) and (6.3). Furthermore, $2b^2 - 8d - 3 > 0$ implies that $2b^2 > 8d + 3 > 8d$, or $b/2 > \sqrt{d}$. Thus from (6.2)

$$\xi_1 < -\frac{b}{2} < -\sqrt{d}. \tag{6.13}$$

Two sub-cases are (i) $\xi_2 \geqslant -\sqrt{d}$ and (ii) $\xi_2 < -\sqrt{d}$.

(i) $\xi_2 \geqslant -\sqrt{d}$:

For this sub-case $f'(\sigma)$ changes sign at most twice in the interval $(-\sqrt{d}, \infty)$, so $f(\sigma)$ changes sign at most three times in $(-\sqrt{d}, \infty)$. However, from (6.10) and (6.12) it follows that $f(\sigma)$ changes sign an even number of times in $(-\sqrt{d}, \infty)$, so it changes sign only twice, once in each of $(-\sqrt{d}, \sqrt{d})$ and $(\sqrt{d}, \infty)$, to be consistent with (6.11). So, noting (6.9), it follows that $f(\sigma)$ changes sign exactly three times, once in each of $(-\infty, \sqrt{d})$, $(-\sqrt{d}, \sqrt{d})$ and $(\sqrt{d}, \infty)$.

(ii) $\xi_2 < -\sqrt{d}$:

This sub-case has two further sub-cases, $f(\xi_1)f(\xi_2) \geqslant 0$, and $f(\xi_1)f(\xi_2) < 0$. The former, in conjunction with (6.9) to (6.12), implies that $f(\sigma)$ changes sign exactly three times, with one sign change in each of the intervals $(-\infty, -\sqrt{d}), (-\sqrt{d}, \sqrt{d})$ and $(\sqrt{d}, \infty)$. In the latter, if $f(\xi_1) < 0$, then $f(\sigma)$ changes sign exactly three times, as in the case of the former, while $f(\xi_1) > 0$ in conjunction with (6.9)–(6.12), implies that $f(\sigma)$ changes sign exactly five times, with one sign change in each of the intervals $(-\infty, \xi_1)$, (ξ_1, ξ_2), $(\xi_2 - \sqrt{d})$, $(-\sqrt{d}, \sqrt{d})$ and $(\sqrt{d}, \infty)$. □

Corollary. *A symmetric parametric cubic with two inflection points has exactly three curvature extrema. They occur at* $\sigma = 0$ *and* $\pm\sqrt{\zeta_2}$, *where* ζ_2 *is given by* (6.6).

Note: Since $f(0) < 0$ for b, $d > 0$ in Theorem 4, the interval $(-\sqrt{d}, \sqrt{d})$ can be shortened to $(-\sqrt{d}, 0)$.

7. Examples

The following examples were used to generate the curves of Fig. 1. They illustrate some practical applications of the theorems. The theoretical results are for the parametric cubic, $-\infty < t < \infty$. The examples are cubic Bézier curves $\boldsymbol{Q}(t)$, $0 \leqslant t \leqslant 1$.

The control points for the curves were determined by experimentation, but guided by the results of the theorems. For some examples (the symmetric ones) an initial symmetric control polygon, with a shape corresponding to the desired curve, was used as a basic curve. Location of curvature extrema were then calculated, with subsequent adjustment of the control polygon until a suitable curve was obtained. Finding some of the other curves (e.g., Examples (d) and (g)) was more challenging. For example, in (d) there is a trade-off in the size of the loop versus the distinction of the curvature extrema. From the numerical results below it is clear that (d) has three extrema, but two of them are so close and vary so little that they are not distinguishable in the figures. The numerical results (location of curvature extrema, etc.) were obtained using a combination of bisection and Newton's method on intervals determined by applying the theorems.

Example (*a*)

Control points: (0,0) (0,1) (1,0) (1,1)

Inflection at: $t = 0.5$

Curvature extrema at: $t = 0.228, 0.772$
Example (*b*)
Control points: (0,1) (0,0) (1,0) (1,1)
Self-intersection at: $t = -0.366, 1.366$
Curvature extremum at: $t = 0.5$
Example (*c*)
Control points: (0,1) (0,0) (2,0) (2,1)
Self-intersection at: $t = -0.366, 1.366$
Curvature extrema at: $t = 0.095, 0.500, 0.905$
Example (*d*)
Control points: (1.5,0) (1.5,1) (0,1) (2.5,0.2)
Self-intersection at: $t = 0.249, 0.757$
Curvature extrema at: $t = 0.145, 0.209, 0.532$
Example (*e*)
Control points: (1,0) (1,0.5) (0,0.5) (2,0)
cusp at: $t = 0.5$
Curvature extrema at: $t = 0.048, 0.327$
Example (*f*)
Control points: (0,0) (1.6,1) (0.4,1) (2,0)
Inflections at: $t = 0.311, 0.689$
Curvature extrema at: $t = 0.219, 0.500, 0.781$
Example (*g*)
Control points: (0.987,0) (0.987,0.25) (0,0) (2,0.25)
Inflections at: $t = 0.494, 0.500$
Curvature extrema at: $t = 0.040, 0.370, 0.491, 0.498, 0.502$

8. Conclusion

In conclusion, the foregoing results for nondegenerate cubic curves are now summarised. The cubic function curve, $y = \beta_3 x^3 + \beta_2 x^2 + \beta_1 x + \beta_0$, $\beta_3 \neq 0$, is a special case of a parametric cubic. For this special case, the curve always has one inflection point, and the curvature always has two-curvature extrema positioned symmetrically about the inflection point. This case is not mentioned in [18,10].

Except for the above special case, a parametric cubic can be classified as having a cusp, a self-intersection, or two inflection points. Only one of the above three features occur in any non-degenerate parametric cubic. A cubic with a cusp has zero or two curvature extrema. A self-intersecting cubic has one or three curvature extrema; at least one of the extrema occurs in the loop part of the curve. A cubic with two inflections has three or five curvature extrema; there is exactly one extremum between the two inflection points. Search intervals, each containing a single extremum have been established in all of the above cases.

Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada.

References

[1] K.G. Baass, The use of clothoid templates in highway design, Transport. Forum 1 (1984) 47–52.
[2] C. Baumgarten, G. Farin, Approximation of logarithmic spirals, Comput. Aided Geom. Des. 14 (1997) 515–532.
[3] G. Farin, Degree reduction fairing of cubic B-spline curves, in: R.E. Barnhill (Ed.), Geometry Processing for Design and Manufacturing, SIAM, Philadelphia, 1992, pp. 87–99.
[4] G. Farin, Curves and Surfaces for Computer Aided Geometric Design: A Practical Guide, 4th Edition, Academic Press, San Diego, 1997.
[5] I.D. Faux, M.J. Pratt, Computational Geometry for Design and Manufacture, Ellis Horwood, Chichester, UK, 1979.
[6] J. Hoschek, D. Lasser, Fundamentals of Computer Aided Geometric Design (Translated from German by L.L. Schumaker). A K Peters, Wellesley, MA, 1993.
[7] J.H. Mathews, Numerical Methods for Computer Science, Engineering, and Mathematics, Prentice-Hall, Englewood Cliffs, NJ, 1987.
[8] D.S. Meek, D.J. Walton, The use of Cornu spirals in drawing planar curves of controlled curvature, J. Comput. Appl. Math. 25 (1989) 69–78.
[9] D.S. Meek, D.J. Walton, Clothoid spline transition spirals, Math. Comput. 59 (1992) 117–133.
[10] R.R. Patterson, Parametric cubics as algebraic curves, Comput. Aided Geom. Des. 5 (1988) 139–159.
[11] N.A. Sapidis (Ed.), Designing Fair Curves and Surfaces: Shape Quality in Geometric Modeling and Computer-Aided Design, SIAM, Philadelphia, 1994.
[12] D.J. Walton, D.S. Meek, Curvature bounds for planar B-spline curve segments, Comput.-Aided Des. 20 (3) (1988) 146–150.
[13] D.J. Walton, D.S. Meek, Clothoidal splines, Comput. Graphics 14 (1) (1990) 95–100.
[14] D.J. Walton, D.S. Meek, A planar cubic Bézier spiral, J. Comput. Appl. Math. 72 (1996) 85–100.
[15] D.J. Walton, D.S. Meek, A Pythagorean hodograph quintic spiral, Comput.-Aided Des. 28 (1996) 943–950.
[16] D.J. Walton, D.S. Meek, G^2 curves composed of planar cubic and Pythagorean hodograph quintic spirals, Comput. Aided Geom. Des. 15 (1998) 547–566.
[17] D.J. Walton, D.S. Meek, Planar G^2 curve design with spiral segments, Comput.-Aided Des. 30 (7) (1998) 529–538.
[18] C.Y. Wang, Shape classification of the parametric cubic curve and parametric B-spline cubic curve, Comput.-Aided Des. 13 (1981) 199–206.