Geometric Hermite curves with minimum strain energy ☆

Jun-Hai Yong, Fuhua (Frank) Cheng [*]

*Graphics & Geometric Modeling Lab., Department of Computer Science, University of Kentucky,
Lexington, KY 40506-0046, USA*

Abstract

The purpose of this paper is to provide yet another solution to a fundamental problem in computer aided geometric design, i.e., constructing a smooth curve satisfying given endpoint (position and tangent) conditions. A new class of curves, called *optimized geometric Hermite* (*OGH*) *curves*, is introduced. An OGH curve is defined by optimizing the magnitudes of the endpoint tangent vectors in the Hermite interpolation process so that the strain energy of the curve is a minimum. An OGH curve is not only mathematically smooth, i.e., with minimum strain energy, but also geometrically smooth, i.e., loop-, cusp- and fold-free if the geometric smoothness conditions and the tangent direction preserving conditions on the tangent angles are satisfied. If the given tangent vectors do not satisfy the tangent angle constraints, one can use a 2-segment or a 3-segment *composite optimized geometric Hermite* (*COH*) *curve* to meet the requirements. Two techniques for constructing 2-segment COH curves and five techniques for constructing 3-segment COH curves are presented. These techniques ensure automatic satisfaction of the tangent angle constraints for each OGH segment and, consequently, mathematical and geometric smoothness of each segment of the curve. The presented OGH and COH curves, combined with symmetry-based extension schemes, cover tangent angles of all possible cases. The new method has been compared with the high-accuracy Hermite interpolation method by de Boor et al. and the Pythagorean-hodograph (PH) curves by Farouki et al. While the other two methods both would generate unpleasant shapes in some cases, the new method generates satisfactory shapes in all the cases.

© 2003 Elsevier B.V. All rights reserved.

Keywords: Hermite; Geometric continuity; Strain energy; Smoothness

1. Introduction

Constructing a smooth curve with given endpoint conditions is a fundamental problem in computer aided geometric design (CAGD). The Hermite interpolation process is frequently used in the construction

---

[*] Corresponding author.
E-mail addresses: junhai@csr.uky.edu (J.-H. Yong), cheng@cs.uky.edu (F.F. Cheng).
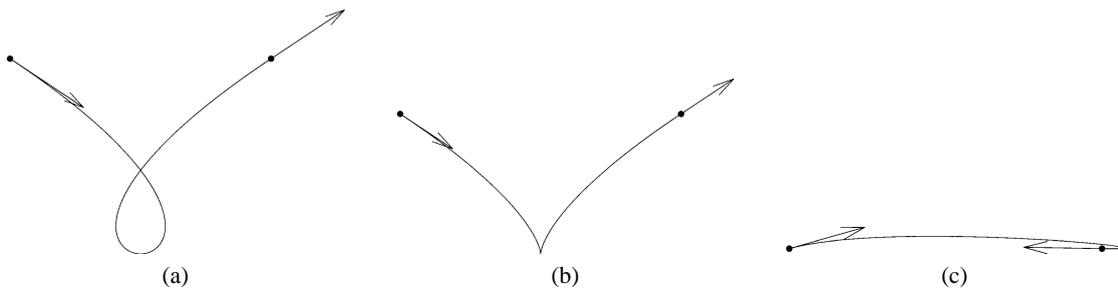
Fig. 1. Examples of a Hermite curve with (a) a loop, (b) a cusp, and (c) a fold.

of such a curve and the resulting cubic polynomial curve is called a Hermite curve. As pointed out in (Zhang et al., 2001), a cubic Hermite curve is mathematically smooth because it has the minimum strain energy among all C^1 cubic polynomial spline curves satisfying the same endpoint conditions. Hence, fairing a C^1 cubic spline curve with endpoint (position and tangent) constraints will eventually lead to a cubic Hermite curve. Unfortunately, a mathematically smooth Hermite curve might not be geometrically smooth. As shown in Fig. 1, cubic Hermite curves may have loops, cusps, or folds.

The standard Hermite technique has been extended in several directions. A recent focus is the so called *geometric Hermite curves* (de Boor et al., 1987; Höllig and Koch, 1995, 1996; Reif, 1999; Schaback, 1998; Wang and Cheng, 1997). A curve of this type allows flexibility on the magnitudes of the two given tangent vectors and, consequently, can satisfy additional requirements. Current research on geometric Hermite curves can be classified into two categories. In the first category, the research focuses on building a low degree geometric Hermite curve with high order geometric continuity and approximation accuracy. The first work in this category is presented by de Boor et al. (1987) in which an interpolation scheme using a cubic spline curve with G^2 continuity and sixth order approximation accuracy under a provided condition is introduced. Several results in this direction have been presented after that (Degen, 1993; Höllig and Koch, 1995, 1996). Höllig and Koch (1995) present a method for interpolating space curves by interpolating a third point in the parameter interval. The new interpolation scheme introduced by them produces a quadratic geometric Hermite interpolant with curvature continuity and fourth order approximation accuracy (Höllig and Koch, 1996). An analysis on the local existence of the quadratic geometric Hermite interpolant of the Höllig and Koch type is provided by Reif (1999). Schaback's work (1998) focuses on determining the minimal degree under a specific situation.

Research work in the second category focuses on producing a G^1 geometric Hermite curve with a pleasing shape (Meek and Walton, 1997a, 1997b). A curve with a pleasing shape should not contain undesired features such as loops, cusps or folds. The work of this paper falls into this category. A reason for one to get undesired shapes is unsuitable magnitudes of the given tangent vectors. Usually, the larger the magnitudes of the tangent vectors, the more likely the occurrence of a loop in the resulting curve. On the other hand, the smaller the magnitudes of the tangent vectors, the closer the resulting curve to the base line segment. Therefore, the problem is how to choose suitable magnitudes for the endpoint tangent vectors.

Another reason for one to get undesired shapes is unsuitable directions of the given tangent vectors. This can be easily verified by holding the (large enough) magnitudes of the endpoint tangent vectors fixed, while rotating the directions of the tangent vectors. Hence, another problem with Hermite interpolation is how to deal with the directions of the given tangent vectors.

Meek and Walton (1997a) use a Tschirnhausen cubic (or, simply, *T-cubic* (Farouki and Sakkalis, 1990)) curve to avoid unpleasant shapes. T-cubic curves have several very desirable properties and can be joined with circular arcs to form nice spirals. Actually, a T-cubic would also provide fourth order approximation accuracy if the curve segment is short enough (Meek and Walton, 1997a, 1997b). What we are interested here, however, are explicit ways to quantize the smoothness of a curve in the geometric Hermite interpolation process both mathematically and geometrically. As a curve with minimum strain energy is often considered as the "smoothest" curve mathematically (Horn, 1983) and strain energy minimization has become a popular method in curve/surface fairing (Zhang et al., 2001), it is desired to actually construct a curve with minimum strain energy in the designing process instead of obtaining it by approximate methods during the fairing process. Furthermore, it is also desired to know explicitly under what circumstance (or, circumstances) would one get a curve in the geometric Hermite interpolation process without the possibility of getting a loop, a cusp, or a fold. Meek and Walton achieve the goal of getting pleasing curves by implicitly putting restriction on the directions of the input tangent vectors (Meek and Walton, 1997a). In this paper, we provide a solution to these problems by presenting a new class of curves called *optimized geometric Hermite* (OGH) curves. A curve in this class is defined by optimizing the magnitudes of the endpoint tangent vectors in the Hermite interpolation process so that the strain energy of the curve is a minimum. An explicit formula for obtaining such a curve is presented. Circumstances (tangent direction preserving conditions and geometric smoothness conditions on the given tangent angles) under which an OGH curve would be geometrically smooth, i.e., loop-, cusp- and fold-free are also found. If the given tangent vectors do not satisfy these tangent angle constraints, one can use a 2-segment or 3-segment composite optimized geometric Hermite (COH) curve to meet the requirements. Two techniques for constructing 2-segment COH curves and five techniques for constructing 3-segment COH curves are constructed. These techniques guarantee automatic satisfaction of the tangent angle constraints for each segment and, consequently, the mathematical and geometric smoothness of each segment of the curve. The OGH curves and the (2-segment and 3-segment) COH curves, together with some symmetry-based extension schemes, can cover tangent angles of all different cases.

The methods developed in this paper find applications in several important areas, such as surface fairing (Chen, 1993; Chen et al., 1997; Zhang and Cheng, 1998). Local irregularities of a NURBS surface are detected by identifying the resulting abnormal portions of the highlight lines in the corresponding regions of the surface. Each abnormal portion of the highlight lines is then replaced with a smooth curve constructed using the above curve construction process. The surface is then deformed so that the new surface would have the modified highlight lines as the new highlight lines. The control points of the new surface are obtained by solving a system of linear equations. In (Chen, 1993; Chen et al., 1997), the abnormal highlight lines are manually moved to the desired positions. (Zhang and Cheng, 1998) uses Hermite curves in their fairing process, which may have loops, cusps, or folds as shown in Fig. 1. The performance of these methods would be improved by our new methods since the resulting curves are guaranteed to be mathematically and geometrically smooth.

The remaining part of the paper is arranged as follows. Related preliminary results, the definition of an optimized geometric Hermite (OGH) curve and constraints on the endpoint tangents to ensure tangent direction preserving property and geometric smoothness of the resulting OGH curve are given in Section 2. Definition of a composite optimized geometric Hermite (COH) curve and construction techniques of 2-segment COH curves and 3-segment COH curves are presented in Section 3. Extension of the one-segment, two-segment and three-segment COH curves to cover all possible tangent angles are

presented in Section 4. Concluding remarks and a comparison of the new method with the high-accuracy Hermite interpolation method by de Boor et al. (1987) and the Pythagorean-hodograph (PH) curves by Farouki et al. (Farouki and Sakkalis, 1990; Farouki and Neff, 1995) on some typical data sets are given in Section 5.

2. Optimized geometric Hermite (OGH) curves

A cubic Hermite curve $\mathbf{Q}(t)$, $t \in [t_0, t_1]$ (where $t_0, t_1 \in \mathbb{R}$ and $t_0 < t_1$) , is a cubic polynomial curve satisfying the following endpoint location and tangent vector conditions:

$$\mathbf{Q}(t_0) = \mathbf{P}_0, \quad \mathbf{Q}(t_1) = \mathbf{P}_1, \quad \mathbf{Q}'(t_0) = \mathbf{V}_0, \quad \text{and} \quad \mathbf{Q}'(t_1) = \mathbf{V}_1,$$

where $\mathbf{P}_0$ and $\mathbf{P}_1$ are given 2D or 3D points, and $\mathbf{V}_0$ and $\mathbf{V}_1$ are given tangent vectors at $\mathbf{P}_0$ and $\mathbf{P}_1$ respectively. $\mathbf{Q}(t)$ can be expressed as follows:

$$\mathbf{Q}(t) = (2s + 1)(s - 1)^2 \mathbf{P}_0 + (-2s + 3)s^2 \mathbf{P}_1 + (1 - s)^2 s(t_1 - t_0)\mathbf{V}_0 + (s - 1)s^2(t_1 - t_0)\mathbf{V}_1, \quad (1)$$

where $s = \frac{t - t_0}{t_1 - t_0}$. The strain energy of a (piecewise) C^2-continuous curve $f(t)$ defined on $[t_0, t_1]$ is defined as follows:

$$\int_{t_0}^{t_1} \left[f''(t) \right]^2 \mathrm{d}t,$$

where $f''(t)$ is the second derivative of $f(t)$.

A Hermite curve is considered mathematically smooth because it has minimum strain energy among all C^1 cubic polynomial spline curves satisfying the same endpoint conditions. This follows from the following theorem in (Zhang et al., 2001).

Theorem 1. *If a cubic Hermite curve $\mathbf{Q}(t)$ and a C^1 cubic polynomial spline curve $\overline{\mathbf{Q}}(t)$ have the same parameter space, $[t_0, t_1]$, and the same endpoint conditions, i.e.,*

$$\overline{\mathbf{Q}}(t_0) = \mathbf{Q}(t_0) = \mathbf{P}_0, \quad \overline{\mathbf{Q}}(t_1) = \mathbf{Q}(t_1) = \mathbf{P}_1, \quad \overline{\mathbf{Q}}'(t_0) = \mathbf{Q}'(t_0) = \mathbf{V}_0, \quad \overline{\mathbf{Q}}'(t_1) = \mathbf{Q}'(t_1) = \mathbf{V}_1,$$

then

$$\int_{t_0}^{t_1} \left[\overline{\mathbf{Q}}''(t) \right]^2 \mathrm{d}t \geqslant \int_{t_0}^{t_1} \left[\mathbf{Q}''(t) \right]^2 \mathrm{d}t.$$

The above theorem can be proved by mathematical induction. Note that, while $\mathbf{Q}(t)$ in the above theorem is uniquely determined by the endpoint constrains, $\overline{\mathbf{Q}}(t)$ has much more degree of freedom than the cubic Hermite curve $\mathbf{Q}(t)$. This can be easily seen from the construction of $\overline{\mathbf{Q}}(t)$. Given $n + 1$ ($n \in \mathbb{Z}$) real numbers $\{s_i\}_{i=0}^n$, $n + 1$ points $\{\mathbf{R}_i\}_{i=0}^n$, and $n + 1$ vectors $\{\mathbf{T}_i\}_{i=0}^n$ with

$$t_0 = s_0 < s_1 < s_2 < \cdots < s_{n-1} < s_n = t_1. \quad (2)$$

$\mathbf{R}_0 = \mathbf{P}_0$, $\mathbf{R}_n = \mathbf{P}_1$, $\mathbf{T}_0 = \mathbf{V}_0$, and $\mathbf{T}_n = \mathbf{V}_1$, $\overline{\mathbf{Q}}(t)$ can be considered as a composition of cubic Hermit curve segments $\overline{\mathbf{Q}}_i(t)$, $t \in [s_{i-1}, s_i]$ ($i = 1, 2, \ldots, n$), satisfying the conditions

$$\overline{\mathbf{Q}}_i(s_{i-1}) = \mathbf{R}_{i-1}, \quad \overline{\mathbf{Q}}_i(s_i) = \mathbf{R}_i, \quad \overline{\mathbf{Q}}_i'(s_{i-1}) = \mathbf{T}_{i-1}, \quad \text{and} \quad \overline{\mathbf{Q}}_i'(s_i) = \mathbf{T}_i.$$

$\{s_i\}_{i=0}^{n}$ are usually called the knots of $\overline{\mathbf{Q}}(t)$. Equations of the segments of $\overline{\mathbf{Q}}(t)$ are similar to the one in Eq. (1). The expression of $\overline{\mathbf{Q}}(t)$ can be converted into a B-spline form by solving a system of linear equations. The degree of freedom of $\overline{\mathbf{Q}}(t)$ is easy to identify now: n can be any positive integer; the knot sequence $\{s_i\}_{i=0}^{n}$ only needs to satisfy the constrain (2); the joints $\{\mathbf{R}_i\}_{i=1}^{n-1}$ and the tangent vectors $\{\mathbf{T}_i\}_{i=1}^{n-1}$ at these joints can be arbitrarily chosen. Theorem 1 shows that fairing $\overline{\mathbf{Q}}(t)$ based on minimizing strain energy will eventually lead to a cubic Hermite curve, i.e., $\mathbf{Q}(t)$.

However, a cubic Hermite curve might not be geometrically smooth. A curve is said to be geometrically smooth if it is loop-, cusp- and fold-free. As shown in Fig. 1, Hermite curves could have loops, cusps, or folds. Hence, Hermite interpolation based curve construction techniques alone are not sufficient for generating curves that are both mathematically and geometrically smooth. We will present a solution to this problem by putting more flexibility on one aspect of the interpolation process, i.e., magnitudes of the endpoint tangent vectors. The intention is to construct a curve that is not only loop-, cusp- and fold-free, but also has minimum strain energy among all similar Hermite curves. We need a definition first.

Definition 1. Given two endpoints $\mathbf{P}_0$ and $\mathbf{P}_1$, and two endpoint tangent vectors $\mathbf{V}_0$ and $\mathbf{V}_1$, a cubic polynomial curve $\mathbf{Q}(t)$, $t \in [t_0, t_1]$, is called an *optimized geometric Hermite* (*OGH*) *curve* with respect to the given endpoint conditions $\{\mathbf{P}_0, \mathbf{P}_1, \mathbf{V}_0, \mathbf{V}_1\}$ if the curve has the smallest strain energy among all cubic Hermite curves $\overline{\mathbf{Q}}(t)$, $t \in [t_0, t_1]$, satisfying the following conditions:

$$\overline{\mathbf{Q}}(t_0) = \mathbf{P}_0, \quad \overline{\mathbf{Q}}(t_1) = \mathbf{P}_1, \quad \overline{\mathbf{Q}}'(t_0) = a_0 \mathbf{V}_0, \quad \overline{\mathbf{Q}}'(t_1) = a_1 \mathbf{V}_1, \tag{3}$$

where a_0 and a_1 are arbitrary real numbers.

Such an OGH curve always exists. Actually, one can find the values of a_0 and a_1 which define the OGH curve explicitly. This follows from the following theorem.

Theorem 2. *Given two endpoints* $\mathbf{P}_0$ *and* $\mathbf{P}_1$, *two endpoint tangent vectors* $\mathbf{V}_0$ *and* $\mathbf{V}_1$, *and a parameter space* $[t_0, t_1]$, *an OGH curve* $\mathbf{Q}(t)$, $t \in [t_0, t_1]$, *with respect to the endpoint conditions* $\{\mathbf{P}_0, \mathbf{P}_1, \mathbf{V}_0, \mathbf{V}_1\}$ *is obtained at* $a_0 = a_0^*$ *and* $a_1 = a_1^*$ *where*

$$\begin{cases} a_0^* = \dfrac{6[(\mathbf{P}_1 - \mathbf{P}_0) \cdot \mathbf{V}_0](\mathbf{V}_1^2) - 3[(\mathbf{P}_1 - \mathbf{P}_0) \cdot \mathbf{V}_1](\mathbf{V}_0 \cdot \mathbf{V}_1)}{[4\mathbf{V}_0^2(\mathbf{V}_1^2) - (\mathbf{V}_0 \cdot \mathbf{V}_1)^2](t_1 - t_0)}, \\ a_1^* = \dfrac{3[(\mathbf{P}_1 - \mathbf{P}_0) \cdot \mathbf{V}_0](\mathbf{V}_0 \cdot \mathbf{V}_1) - 6[(\mathbf{P}_1 - \mathbf{P}_0) \cdot \mathbf{V}_1](\mathbf{V}_0^2)}{[(\mathbf{V}_0 \cdot \mathbf{V}_1)^2 - 4\mathbf{V}_0^2(\mathbf{V}_1^2)](t_1 - t_0)}. \end{cases} \tag{4}$$

Proof. A cubic Hermite curve $\overline{\mathbf{Q}}(t)$, $t \in [t_0, t_1]$, satisfying the constraints in (3) can be expressed as

$$\overline{\mathbf{Q}}(t) = (2s + 1)(s - 1)^2 \mathbf{P}_0 + (-2s + 3)s^2 \mathbf{P}_1 + (1 - s)^2 s(t_1 - t_0) a_0 \mathbf{V}_0 + (s - 1)s^2(t_1 - t_0) a_1 \mathbf{V}_1,$$

where $s = \frac{t - t_0}{t_1 - t_0}$. With $\mathbf{P}_0, \mathbf{P}_1, \mathbf{V}_0, \mathbf{V}_1$ being known, the strain energy of $\overline{\mathbf{Q}}(t)$ can be represented as a function of a_0 and a_1 as follows:

$$f(a_0, a_1) = \frac{\mathbf{A}^2(t_1 - t_0)}{3} + (\mathbf{A} \cdot \mathbf{B})(t_1 - t_0) + \mathbf{B}^2(t_1 - t_0),$$

where

$$\mathbf{A} = \frac{12\mathbf{P}_0}{(t_1 - t_0)^2} - \frac{12\mathbf{P}_1}{(t_1 - t_0)^2} + \frac{6a_0\mathbf{V}_0}{t_1 - t_0} + \frac{6a_1\mathbf{V}_1}{t_1 - t_0}$$

and

$$\mathbf{B} = -\frac{6\mathbf{P}_0}{(t_1 - t_0)^2} + \frac{6\mathbf{P}_1}{(t_1 - t_0)^2} - \frac{4a_0\mathbf{V}_0}{t_1 - t_0} - \frac{2a_1\mathbf{V}_1}{t_1 - t_0}.$$

The optimization problem,

$$\min f(a_0, a_1),$$

is equivalent to two equations

$$\begin{cases} \frac{\partial f(a_0,a_1)}{\partial a_0} = 0, \\ \frac{\partial f(a_0,a_1)}{\partial a_1} = 0. \end{cases}$$

Solving these equations, we get

$$\begin{cases} \mathbf{B} \cdot \mathbf{V}_0 = 0, \\ \mathbf{A} \cdot \mathbf{V}_1 + \mathbf{B} \cdot \mathbf{V}_1 = 0. \end{cases}$$

Substituting the values of $\mathbf{A}$ and $\mathbf{B}$ into the above two equations, we obtain

$$\begin{cases} \left[-\frac{6\mathbf{P}_0}{(t_1-t_0)^2} + \frac{6\mathbf{P}_1}{(t_1-t_0)^2} - \frac{4a_0\mathbf{V}_0}{t_1-t_0} - \frac{2a_1\mathbf{V}_1}{t_1-t_0} \right] \cdot \mathbf{V}_0 = 0, \\ \left[\frac{12\mathbf{P}_0}{(t_1-t_0)^2} - \frac{12\mathbf{P}_1}{(t_1-t_0)^2} + \frac{6a_0\mathbf{V}_0}{t_1-t_0} + \frac{6a_1\mathbf{V}_1}{t_1-t_0} - \frac{6\mathbf{P}_0}{(t_1-t_0)^2} + \frac{6\mathbf{P}_1}{(t_1-t_0)^2} - \frac{4a_0\mathbf{V}_0}{t_1-t_0} - \frac{2a_1\mathbf{V}_1}{t_1-t_0} \right] \cdot \mathbf{V}_1 = 0. \end{cases}$$

Simplifying the above equations, we get

$$\begin{cases} 2a_0\mathbf{V}_0^2(t_1 - t_0) + a_1(\mathbf{V}_0 \cdot \mathbf{V}_1)(t_1 - t_0) = 3(\mathbf{P}_1 - \mathbf{P}_0) \cdot \mathbf{V}_0, \\ a_0(\mathbf{V}_0 \cdot \mathbf{V}_1)(t_1 - t_0) + 2a_1\mathbf{V}_1^2(t_1 - t_0) = 3(\mathbf{P}_1 - \mathbf{P}_0) \cdot \mathbf{V}_1. \end{cases}$$

Theorem 2 is obtained by solving the above two linear equations.   □

a_0^* and a_1^* defined in Eqs. (4) are called the *optimized coefficients* of the tangent vectors of $\mathbf{Q}(t)$ at t_0 and t_1, respectively. Their values are not necessarily to be positive. A negative a_0^* means the tangent vector of the OGH curve $\mathbf{Q}(t)$ at t_0 is in the opposite direction of $\mathbf{V}_0$, and a negative a_1^* means the tangent vector of $\mathbf{Q}(t)$ at t_1 is in the opposite direction of $\mathbf{V}_1$. Either of these certainly is not desired. We would like a_0^* and a_1^* both to be positive so that the tangent vectors of $\mathbf{Q}(t)$ at t_0 and t_1 would have the same direction as that of $\mathbf{V}_0$ and $\mathbf{V}_1$, respectively. The conditions for a_0^* and a_1^* to be positive are given in Theorem 3. To make subsequent presentation simpler, we give a definition first.

Definition 2. An OGH curve is said to satisfy the *tangent direction preserving property* if both optimized coefficients of the tangent vectors of the curve are positive.

Theorem 3. $\mathbf{Q}(t)$, $t \in [t_0, t_1]$, *is an OGH curve with respect to the endpoint conditions* $\{\mathbf{P}_0, \mathbf{P}_1, \mathbf{V}_0, \mathbf{V}_1\}$, a_0^* *and* a_1^* *are the optimized coefficients of the tangent vectors of* $\mathbf{Q}(t)$ *at* t_0 *and* t_1, *respectively.* $\mathbf{Q}(t)$ *satisfies the tangent direction preserving property if and only if the following* tangent direction preserving conditions *are satisfied*:

$$3\cos\theta > \cos(\theta - 2\varphi) \quad and \quad 3\cos\varphi > \cos(\varphi - 2\theta), \tag{5}$$

where θ is the counterclockwise angle from the vector $\overrightarrow{\mathbf{P}_0\mathbf{P}_1}$ to $\mathbf{V}_0$, and φ is the counterclockwise angle from the vector $\overrightarrow{\mathbf{P}_0\mathbf{P}_1}$ to $\mathbf{V}_1$. θ and φ are both 2π-periodic (hence, a clockwise angle would be measured in negative degrees).
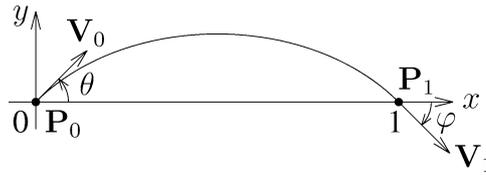
Fig. 2. Building the coordinate system.

Proof. It is sufficient to prove that $a_0^* > 0$ if and only if

$$3\cos\theta > \cos(\theta - 2\varphi),$$

and $a_1^* > 0$ if and only if

$$3\cos\varphi > \cos(\varphi - 2\theta).$$

Without loss of generality, we shall assume that $\mathbf{P}_0$ is at the origin and $\mathbf{P}_1$ is at $[1, 0]^{\mathrm{T}}$ (see Fig. 2). Suppose the magnitudes of $\mathbf{V}_0$ and $\mathbf{V}_1$ are r_0 and r_1, respectively, where $r_0, r_1 > 0$. Thus, coordinates of $\mathbf{V}_0$ and $\mathbf{V}_1$ are $[r_0 \cos\theta, r_0 \sin\theta]^{\mathrm{T}}$ and $[r_1 \cos\varphi, r_1 \sin\varphi]^{\mathrm{T}}$, respectively. Substitute these coordinates into Eqs. (4), we obtain

$$\begin{cases} a_0^* = \frac{6\cos(\theta) - 3\cos(\varphi)\cdot\cos(\theta - \varphi)}{[4 - \cos^2(\theta - \varphi)](t_1 - t_0)r_0}, \\ a_1^* = \frac{6\cos(\varphi) - 3\cos(\theta)\cdot\cos(\theta - \varphi)}{[4 - \cos^2(\theta - \varphi)](t_1 - t_0)r_1}. \end{cases}$$

Therefore, $a_0^* > 0$ if and only if

$$6\cos(\theta) - 3\cos(\varphi)\cdot\cos(\theta - \varphi) > 0,$$

and $a_1^* > 0$ if and only if

$$6\cos(\varphi) - 3\cos(\theta)\cdot\cos(\theta - \varphi) > 0.$$

Simplify the above two inequalities, we get the conclusion of Theorem 3. $\square$

     The regions of the $\theta\varphi$-plane where the tangent direction preserving conditions (5) are satisfied, i.e., where the corresponding a_0^* and a_1^* would be positive, are shown in Fig. 3.

     In some cases, one might want to hold the ratio of the tangent vector magnitudes unchanged, i.e., setting $a_0 = a_1$ in Definition 1, so that a fixed shape style can be maintained on the resulting curve. In such a case, $\mathbf{Q}(t)$ obtains the minimum energy when

$$a_0^* = a_1^* = a^* = \frac{3[(\mathbf{P}_1 - \mathbf{P}_0)\cdot(\mathbf{V}_0 + \mathbf{V}_1)]}{2(\mathbf{V}_0^2 + \mathbf{V}_0 \cdot \mathbf{V}_1 + \mathbf{V}_1^2)(t_1 - t_0)}.$$

In such a case, a^* is positive if and only if

$$r_0 \cos\theta + r_1 \cos\varphi > 0,$$

where $\theta \in [0, 2\pi)$ is the counterclockwise angle from the vector $\overrightarrow{\mathbf{P}_0\mathbf{P}_1}$ to $\mathbf{V}_0$, $\varphi \in [0, 2\pi)$ is the counterclockwise angle from the vector $\overrightarrow{\mathbf{P}_0\mathbf{P}_1}$ to $\mathbf{V}_1$, and $r_0(> 0)$ and $r_1(> 0)$ are the lengths of $\mathbf{V}_0$ and $\mathbf{V}_1$, respectively. The regions of the $\theta\varphi$-plane where a^* is positive when $r_0 = r_1$ are shown in Fig. 4.
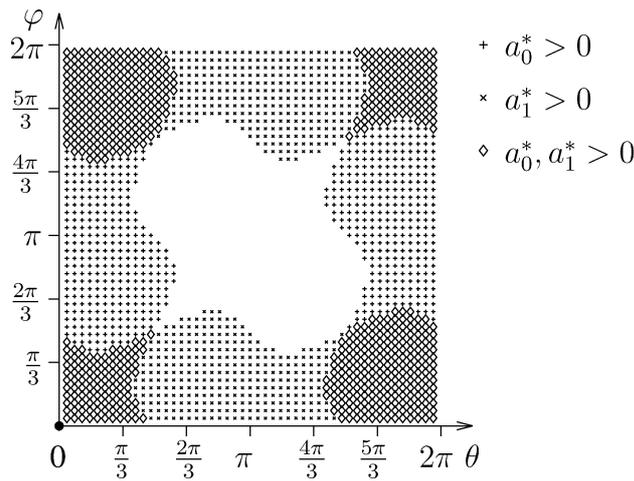
Fig. 3. Partition of the $\theta\varphi$-plane based on the sign of a_i^*, $i = 0, 1$.
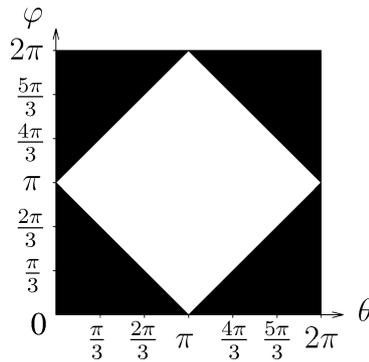

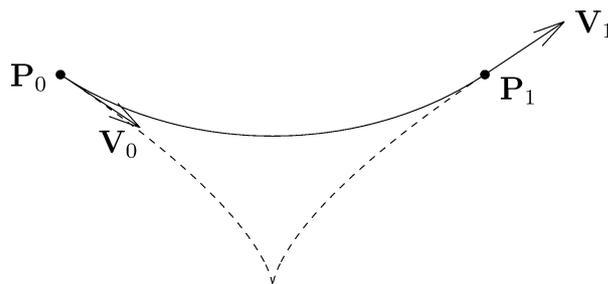Fig. 4. Regions where $a^* > 0$ when $r_0 = r_1$.

Fig. 5. Comparison of an OGH curve (solid) with a cubic Hermite curve (dashed).

Fig. 5 shows the comparison of an OGH curve (solid line) with an ordinary cubic Hermite curve (dashed line). In this example, the input parameters are:

$$t_0 = 0, \quad t_1 = 2; \quad \mathbf{P}_0 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \mathbf{P}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \quad \mathbf{V}_0 = \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \quad \mathbf{V}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

As show in this figure, the cubic Hermite curve in this case has a cusp, while the OGH curve has a quite good shape. The following theorem gives the conditions for an OGH curve to be loop-, cusp-, and fold-free.

Theorem 4. *Let* $\mathbf{Q}(t) = [x(t),\ y(t)]^{\mathrm{T}}$ *be an OGH curve with respect to the endpoint conditions* $\{\mathbf{P}_0,\ \mathbf{P}_1,\ \mathbf{V}_0,\ \mathbf{V}_1\}$, *with* $\mathbf{P}_0$ *being the origin of the coordinate system and the vector* $\overrightarrow{\mathbf{P}_0\mathbf{P}_1}$ *coinciding with the x-axis. Then* $x'(t) > 0,\ \forall t \in [t_0, t_1]$, *if*

$$\cos\theta > \frac{1}{3} \quad and \quad \cos\varphi > \frac{1}{3}, \tag{6}$$

where θ *and* φ *are defined in Theorem* 3.

Proof. Without loss of generality, we shall again assume that $\mathbf{P}_1$'s coordinates are $[1, 0]^{\mathrm{T}}$ (see Fig. 2). Besides, we shall also assume that t_0 and t_1 are 0 and 1, respectively, and $\mathbf{V}_0$ and $\mathbf{V}_1$ are both unit vectors. These assumptions simplify the proof of the theorem but do not change the sign of $x'(t)$, i.e., do not change the conditions for $x'(t)$ to be positive. Thus, coordinates of $\mathbf{P}_0$, $\mathbf{P}_1$, $\mathbf{V}_0$ and $\mathbf{V}_1$ are $[0, 0]^{\mathrm{T}}$, $[1, 0]^{\mathrm{T}}$, $[\cos\theta, \sin\theta]^{\mathrm{T}}$ and $[\cos\varphi, \sin\varphi]^{\mathrm{T}}$, respectively. These coordinates and the OGH curve are illustrated in Fig. 2 as well. Substitute the above values into Eq. (1), we have

$$x(t) = \big[-2 + a_0^* \cos(\theta) + a_1^* \cos(\varphi)\big]t^3 + \big[3 - 2a_0^* \cos(\theta) - a_1^* \cos(\varphi)\big]t^2 + a_0^* \cos(\theta)t,$$

where

$$\begin{cases} a_0^* = \frac{6\cos(\theta) - 3\cos(\varphi) \cdot \cos(\theta - \varphi)}{4 - \cos^2(\theta - \varphi)}, \\ a_1^* = \frac{6\cos(\varphi) - 3\cos(\theta) \cdot \cos(\theta - \varphi)}{4 - \cos^2(\theta - \varphi)}. \end{cases}$$

Therefore,

$$x'(t) = \big[-6 + 3a_0^* \cos(\theta) + 3a_1^* \cos(\varphi)\big]t^2 + \big[6 - 4a_0^* \cos(\theta) - 2a_1^* \cos(\varphi)\big]t + a_0^* \cos(\theta).$$

According to Theorem 3, we have

$$\begin{cases} a_0^* > 0, \quad \text{when } \cos(\theta) > 1/3, \\ a_1^* > 0, \quad \text{when } \cos(\varphi) > 1/3. \end{cases}$$

Hence, we obtain

$$\begin{cases} x'(0) = a_0^* \cos(\theta) > 0, \quad \text{when } \cos(\theta) > 1/3, \\ x'(1) = a_1^* \cos(\varphi) > 0, \quad \text{when } \cos(\varphi) > 1/3. \end{cases}$$

For simplicity, we shall use C to notate the coefficient of the second-degree item in $x'(t)$

$$C = -6 + 3a_0^* \cos(\theta) + 3a_1^* \cos(\varphi).$$

Substitute the values of a_0^* and a_1^* into C, we have

$$C = \frac{-24 + 6\cos^2(\theta - \varphi) + 18\cos^2(\theta) - 18\cos(\theta)\cos(\varphi)\cos(\theta - \varphi) + 18\cos^2(\varphi)}{4 - \cos^2(\theta - \varphi)}.$$

Through straightforward algebra, one can show that

$$C = \frac{-12 - 9[\cos(\theta + \varphi) - \cos(\theta - \varphi)]^2 + 3\cos^2(\theta - \varphi) + 9\cos^2(\theta + \varphi)}{2[4 - \cos^2(\theta - \varphi)]}.$$
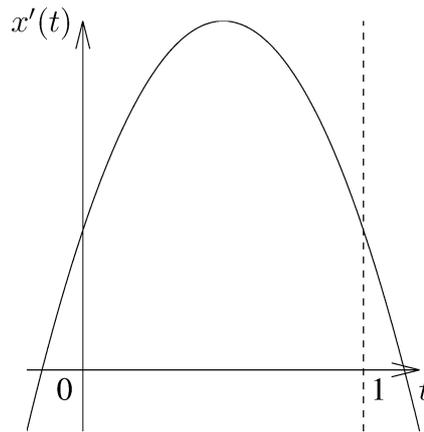
Fig. 6. $x'(t)$ is a concave down parabola when θ or $\varphi \notin \mathbf{S}$, and $\cos(\theta)$ and $\cos(\varphi)$ are both $> 1/3$.

Thus, we have

$$C \leqslant 0.$$

And the equality holds only when

$$\begin{cases} \cos(\theta + \varphi) = \cos(\theta - \varphi), \\ \cos^2(\theta - \varphi) = 1, \\ \cos^2(\theta + \varphi) = 1, \end{cases}$$

which is equivalent to

$$\theta, \varphi \in \mathbf{S}, \quad \text{where } \mathbf{S} = \{k\pi \mid k \in \mathbb{Z}\}.$$

When $\theta, \varphi \in \mathbf{S}$, we have

$$x'(t) = 1, \quad \forall t \in [0, 1].$$

If $\cos(\theta) > \frac{1}{3}$, $\cos(\varphi) > \frac{1}{3}$ and ($\theta \notin \mathbf{S}$ or $\varphi \notin \mathbf{S}$), as shown in Fig. 6, $x'(t)$ is a concave down parabola. Therefore, in this case, we have

$$x'(t) > 0, \quad \forall t \in [0, 1],$$

because its leading coefficient $C < 0$, $x'(0) > 0$, and $x'(1) > 0$. Summarizing the above two cases in the proof, we obtain the conclusion of Theorem 4. □

Since $x'(t) > 0, \forall t \in [t_0, t_1]$, $\mathbf{Q}(t)$ does not have any degenerate points. Thus, $\mathbf{Q}(t)$ does not have any cusps. The reason that $\mathbf{Q}(t)$ is loop- and fold-free follows from the fact that $x(t)$ is an increasing function in this case. The conditions in (6) will be called the *geometric smoothness conditions* since they guarantee the geometric smoothness of the resulting curve. Note that the tangent direction preserving conditions of Theorem 3 are satisfied if the geometric smoothness conditions of Theorem 4 are satisfied.

It should be pointed out that an OGH curve is in general different from a circular arc in that a circular arc cannot be represented in parametric polynomial form. Furthermore, a circular arc (Yong et al., 1999) requires that $\varphi = 2k\pi - \theta$, where $k \in \mathbb{Z}$. This is not required for an OGH curve.

3. Composite optimized geometric Hermite (COH) curves

As shown in Theorem 3, an OGH curve satisfies the tangent direction preserving property if and only if the tangent direction preserving conditions in (5) are satisfied. If the given tangent vectors do not satisfy the conditions in (5), one needs to use a composite OGH curve, instead of a single OGH curve, to ensure the satisfaction of the tangent direction preserving property and the fairness of the resulting curve. The definition of a composite OGH curve is given below.

Definition 3. A piecewise cubic polynomial curve is called a *composite optimized geometric Hermite* (*COH*) *curve* if the curve is G^1 continuous and each segment of the curve is an OGH curve.

The reason to study and use COH curves is obvious. Consider the example given in Fig. 7. The angles of the given tangent vectors in this case are $\theta = \pi/6$ and $\varphi = 2\pi/3$, respectively (Fig. 7(a)). So $3\cos\varphi = -3/2 < \cos(\varphi - 2\theta) = 1/2$. Hence, the tangent direction preserving condition for $t = t_1$ in Theorem 3 is not satisfied and, consequently, the direction of the tangent vector $\mathbf{V}_1$ at $t = t_1$ is not preserved by the OGH curve (Fig. 7(b)). This is certainly not desired. A standard cubic Hermite curve (defined by Eq. (1)), on the other hand, would always retain the directions of the given tangent vectors. However, as shown in Fig. 8, the shape of a standard cubic Hermite curve (solid line) is not good if the given tangent angles or magnitudes of the given tangent vectors are not in suitable conditions. The remedy here is to use COH curves. As shown in Fig. 8, a COH curve (dashed line) would not only satisfy the tangent direction preserving property, but also carry a shape much better than that of the corresponding standard Hermite curve (solid line). The COH curves in Fig. 8 are the same as the ones in Figs. 9(a), 10(a) and 10(d), respectively. In general, the number of OGH segments in a COH curve should be as small as possible, so long as all possible endpoint tangent vectors can be covered. Each COH curve should satisfy the conditions of Theorems 3 and 4 so the directions of the endpoint tangent vectors would be retained and each segment of the COH curve would be loop-, cusp- and fold-free. Our investigation shows that it
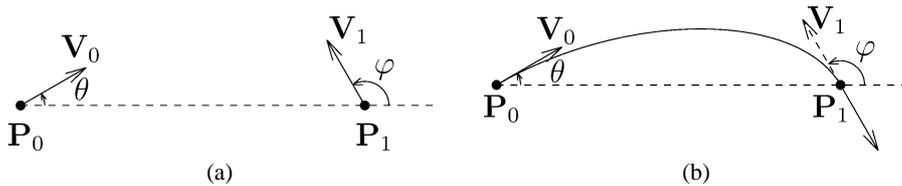

Fig. 7. Direction of a tangent vector is reversed when conditions in Theorem 2 are not satisfied. (a) Given endpoint conditions. (b) OGH curve.
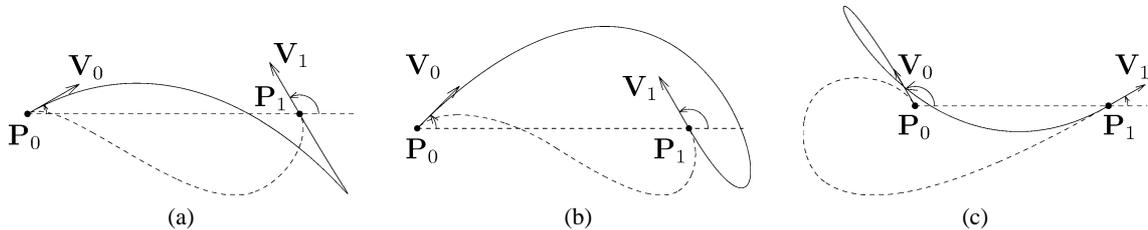

Fig. 8. Standard cubic Hermite curves (solid) compared with COH curves (dashed). (a) Example 1. (b) Example 2. (c) Example 3.
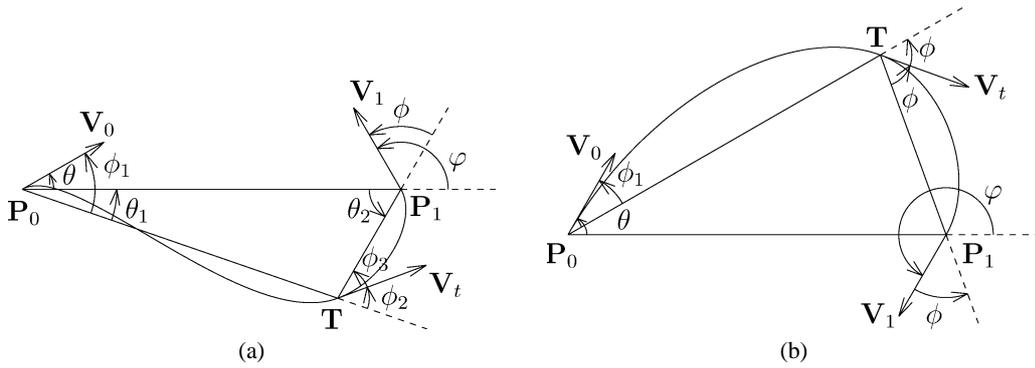
Fig. 9. Two methods, $\mathbf{M}_1$ and $\mathbf{M}_2$, for generating two-segment COH curves. (a) Method $\mathbf{M}_1$. (b) Method $\mathbf{M}_2$.

is sufficient to study two-segment and three-segment COH curves only because these curves are enough to cover all possible cases of the tangent vectors (see Section 4). It is necessary to consider three-segment COH curves because two-segment COH curves alone can not cover all the cases. For example, there is no two-segment COH curve that can satisfy the tangent vectors given in Fig. 10(e) and the conditions of Theorem 3 at the same time.

Several techniques can be used to construct two-segment and three-segment COH curves. In the following, two methods for generating two-segment COH curves are presented. These methods are illustrated in Fig. 9. As will be proved in Theorem 5, these methods guarantee that each segment of the constructed COH curve satisfies the conditions of Theorems 3 and 4 and, therefore, is not only loop-, cusp- and fold-free, but also retains the directions of the given tangent vectors. In these methods, the two OGH segments, their joint and the tangent vector at the joint are denoted $\mathbf{Q}_0(t)$, $\mathbf{Q}_1(t)$, $\mathbf{T}$ and $\mathbf{V}_t$, respectively.

Method $\mathbf{M}_1$. If the tangent angles θ and φ satisfy the condition

$$(\theta, \varphi) \in [0, \pi/6] \times [\pi/3, 2\pi/3]$$

(see Fig. 9(a)) where θ and φ are defined in Theorem 3, then the joint of the two OGH segments, $\mathbf{T}$, is determined by setting the counterclockwise angle from the vector $\overrightarrow{\mathbf{TP}_1}$ to $\mathbf{V}_1$, ϕ, to be $\frac{1}{2}\varphi$ and the length of $\overrightarrow{\mathbf{TP}_1}$ to be $\frac{1}{3}$ of that of $\overrightarrow{\mathbf{P}_0\mathbf{P}_1}$, and $\mathbf{V}_t$ is a vector bisecting the counterclockwise angle from the vector $\overrightarrow{\mathbf{P}_0\mathbf{T}}$ to the vector $\overrightarrow{\mathbf{TP}_1}$.

Method $\mathbf{M}_2$. If the tangent angles θ and φ satisfy the condition

$$(\theta, \varphi) \in [0, \pi/3] \times [\pi, 5\pi/3] \quad \text{or} \quad [\pi/3, 2\pi/3] \times [4\pi/3, 5\pi/3]$$

(see Fig. 9(b)), then the joint of the OGH segments, $\mathbf{T}$, is determined by setting the counterclockwise angle from the vector $\overrightarrow{\mathbf{P}_0\mathbf{T}}$ to $\mathbf{V}_0$ as follows:

$$\phi_1 = \begin{cases} \theta - \frac{\pi}{18}, & \text{if } \theta < \frac{\pi}{9}, \\ \frac{1}{2}\theta, & \text{otherwise}, \end{cases}$$

and the counterclockwise angle from $\mathbf{V}_1$ to the vector $\overrightarrow{\mathbf{TP}_1}$ as

$$\phi = \frac{2\pi - \varphi + \phi_1}{3}.$$

And $\mathbf{V}_t$ is determined by setting the counterclockwise angles from the vector $\overrightarrow{\mathbf{TP_1}}$ to $\mathbf{V}_t$, and from $\mathbf{V}_t$ to the vector $\overrightarrow{\mathbf{P_0T}}$ both equal to ϕ.

Theorem 5. *Given two endpoints $\mathbf{P}_0$ and $\mathbf{P}_1$ and two endpoint tangent vectors $\mathbf{V}_0$ and $\mathbf{V}_1$, if $\mathbf{Q}_0(t)$ is constructed with respect to the endpoint conditions $\{\mathbf{P}_0, \mathbf{T}, \mathbf{V}_0, \mathbf{V}_t\}$ and $\mathbf{Q}_1(t)$ is constructed with respect to the endpoint conditions $\{\mathbf{T}, \mathbf{P}_1, \mathbf{V}_t, \mathbf{V}_1\}$ where $\mathbf{T}$ and $\mathbf{V}_t$ are determined by method $\mathbf{M}_1$ or $\mathbf{M}_2$ then $\mathbf{Q}_0(t)$ and $\mathbf{Q}_1(t)$ both satisfy the conditions of Theorems 3 and 4.*

Proof. We prove the case for $\mathbf{M}_1$ only. The proof of the case for $\mathbf{M}_2$ is similar.

Because $\phi = \varphi/2$ and $\varphi \in [\pi/3, 2\pi/3]$, we have $\phi \in [\pi/6, \pi/3]$. So, obviously,

$$\cos\phi > \frac{1}{3}.$$

In Fig. 9(a), let θ_1 be the counterclockwise angle from the vector $\overrightarrow{\mathbf{P_0T}}$ to the vector $\overrightarrow{\mathbf{P_0P_1}}$, θ_2 be the counterclockwise angle from the vector $\overrightarrow{\mathbf{P_1P_0}}$ to the vector $\overrightarrow{\mathbf{P_1T}}$, ϕ_1 be the counterclockwise angle from the vector $\overrightarrow{\mathbf{P_0T}}$ to the vector $\mathbf{V}_0$, ϕ_2 be the counterclockwise angle from the vector $\overrightarrow{\mathbf{P_0T}}$ to $\mathbf{V}_t$, and ϕ_3 be the counterclockwise angle from $\mathbf{V}_t$ to the vector $\overrightarrow{\mathbf{TP_1}}$. We have

$$\phi_2 + \phi_3 = \theta_1 + \theta_2$$

and

$$\theta_2 = \frac{\varphi}{2} \in [\pi/6, \pi/3].$$

Since the length of the line segment $\overline{\mathbf{TP_1}}$ is $\frac{1}{3}$ of $\overline{\mathbf{P_0P_1}}$, it follows that

$$\sin\theta_1 < \frac{1}{3}.$$

Therefore,

$$\theta_1 \in [0, \pi/6).$$

Consequently,

$$\phi_1 = \theta + \theta_1 \in [0, \pi/3)$$

and

$$\phi_2 = \phi_3 = \frac{\theta_1 + \theta_2}{2} \in [\pi/12, \pi/4).$$

And then we have

$$\cos\phi_1 > \frac{1}{3}, \quad \cos\phi_2 > \frac{1}{3}, \quad \text{and} \quad \cos\phi_3 > \frac{1}{3}.$$

Hence, both OGH segments satisfy the conditions of Theorem 4. Theorem 3 now follows from Theorem 4. $\quad\square$

Five methods for generating three-segment COH curves are presented below. These methods are illustrated in Fig. 10. These methods also guarantee that each segment of the constructed COH curve satisfies the conditions of Theorems 3 and 4 and, therefore, is not only loop-, cusp- and fold-free, but
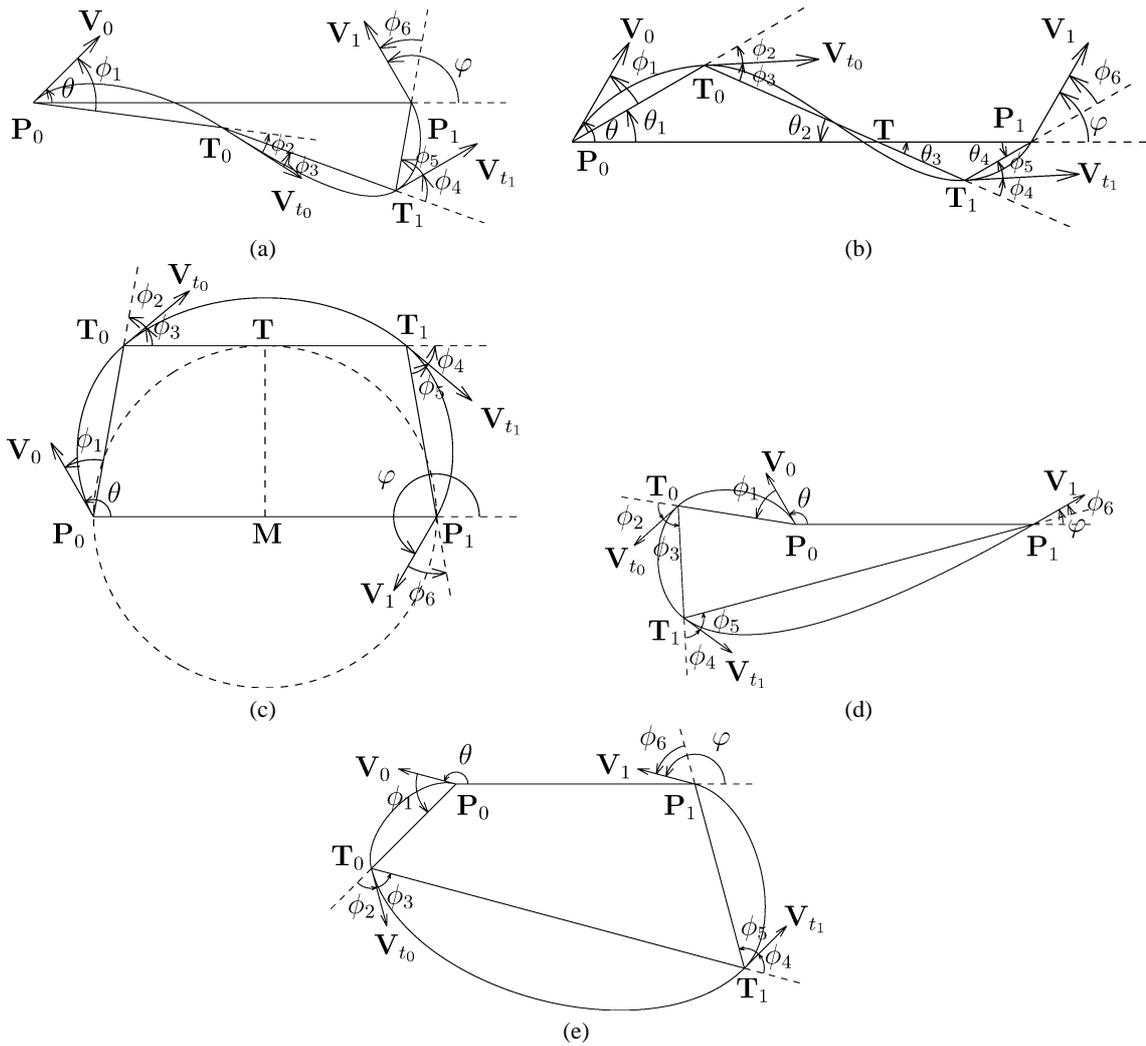
Fig. 10. Methods for generating three-segment COH curves. (a) Method $\mathbf{M}_3$. (b) Method $\mathbf{M}_4$. (c) Method $\mathbf{M}_5$. (d) Method $\mathbf{M}_6$. (e) Method $\mathbf{M}_7$.

also retains the directions of the given tangent vectors (see Theorem 6 below). In these methods, the joints of the three OGH segments $\mathbf{Q}_0(t)$, $\mathbf{Q}_1(t)$ and $\mathbf{Q}_2(t)$ are denoted $\mathbf{T}_0$ and $\mathbf{T}_1$, and the tangent vectors at these points are denoted $\mathbf{V}_{t_0}$ and $\mathbf{V}_{t_1}$, respectively. The signed (slope) angles of vectors $\overrightarrow{\mathbf{P}_0\mathbf{T}_0}$, $\mathbf{V}_{t_0}$, $\overrightarrow{\mathbf{T}_0\mathbf{T}_1}$, $\mathbf{V}_{t_1}$ and $\overrightarrow{\mathbf{T}_1\mathbf{P}_1}$ with respect to $\overrightarrow{\mathbf{P}_0\mathbf{P}_1}$ are denoted α_1, α_2, α_3, α_4, and α_5, respectively. The counterclockwise angles at the endpoints of these three OGH segments with respect to their base lines are denoted ϕ_1, ϕ_2, ϕ_3, ϕ_4, ϕ_5 and ϕ_6, respectively (see Fig. 10).

Method $\mathbf{M}_3$. If the tangent angles θ and φ satisfy the condition

$$(\theta, \varphi) \in [0, \pi/3] \times [\pi/3, \pi]$$

(see Fig. 10(a)), then $\mathbf{T}_0$, $\mathbf{T}_1$, $\mathbf{V}_{t_0}$, and $\mathbf{V}_{t_1}$, are determined by requiring

$$\phi_6 = \phi = \frac{\varphi}{3},$$

$$\alpha_1 = \begin{cases} \frac{\theta - \phi}{2} - \frac{\pi}{18}, & \text{if } \frac{\theta - \phi}{2} - \frac{\pi}{18} \geqslant 0, \\ \frac{\theta - \phi}{2} + \frac{35\pi}{18}, & \text{otherwise}, \end{cases}$$

$$\alpha_3 = \frac{17\pi}{9},$$

$$\alpha_4 = \frac{\alpha_3 + \alpha_5}{2} - \pi,$$

$$\alpha_2 = \alpha_1 - 2t,$$

with

$$t = \begin{cases} |\alpha_3 - \alpha_1|, & \text{if } \frac{\pi}{18} < |\alpha_3 - \alpha_1| < \pi, \\ 2\pi - |\alpha_3 - \alpha_1|, & \text{if } \pi < |\alpha_3 - \alpha_1| < \frac{35\pi}{18}, \\ \frac{\pi}{18}, & \text{otherwise}, \end{cases}$$

and $\mathbf{T}_0$ is on the perpendicular bisector of the line segment $\overline{\mathbf{P}_0\mathbf{P}_1}$.

Method M$_4$. If the tangent angles θ and φ satisfy the condition

$$(\theta, \varphi) \in [\pi/3, 2\pi/3] \times [0, 2\pi/3]$$

(see Fig. 10(b)), then the joints, $\mathbf{T}_0$ and $\mathbf{T}_1$, are determined by requiring

$$\alpha_1 = \frac{\theta}{2}, \quad \alpha_5 = \frac{\varphi}{2}, \quad \frac{\|\overline{\mathbf{P}_0\mathbf{T}_0}\|}{\|\overline{\mathbf{P}_0\mathbf{P}_1}\|} = \frac{1}{3}, \quad \frac{\|\overline{\mathbf{T}_1\mathbf{P}_1}\|}{\|\overline{\mathbf{P}_0\mathbf{P}_1}\|} = \frac{1}{6},$$

the tangent vector $\mathbf{V}_{t_0}$ bisects the angle from the vector $\overrightarrow{\mathbf{T}_0\mathbf{T}_1}$ to the vector $\overrightarrow{\mathbf{P}_0\mathbf{T}_0}$, and the tangent vector $\mathbf{V}_{t_1}$ bisects the angle from the vector $\overrightarrow{\mathbf{T}_0\mathbf{T}_1}$ to the vector $\overrightarrow{\mathbf{T}_1\mathbf{P}_1}$.

Method M$_5$. If the tangent angles θ and φ satisfy the condition

$$(\theta, \varphi) \in [\pi/3, 2\pi/3] \times [\pi, 4\pi/3] \quad \text{or} \quad [2\pi/3, \pi] \times [\pi, 5\pi/3]$$

(see Fig. 10(c)), then the joints and their tangent vectors, $\{\mathbf{T}_0, \mathbf{T}_1, \mathbf{V}_{t_0}, \mathbf{V}_{t_1}\}$, are determined by requiring

$$\phi_1 = \phi_2 = \phi_3 = \phi_4 = \phi_5 = \phi_6,$$

and the line $\overline{\mathbf{T}_0\mathbf{T}_1}$ tangent to the circle whose center is at $\mathbf{M}$ (midpoint of the line segment $\overline{\mathbf{P}_0\mathbf{P}_1}$) and radius is $\frac{1}{2}\|\overline{\mathbf{P}_0\mathbf{P}_1}\|$ at $\mathbf{T}$.

Method M$_6$. If the tangent angles θ and φ satisfy the condition

$$(\theta, \varphi) \in [2\pi/3, \pi] \times [\pi/6, 2\pi/3]$$

(see Fig. 10(d)), then the joints and their tangent vectors, $\{\mathbf{T}_0, \mathbf{T}_1, \mathbf{V}_{t_0}, \mathbf{V}_{t_1}\}$, are determined by requiring

$$\alpha_5 = \frac{\varphi}{2}, \quad \phi_1 = \phi_2 = \phi_3 = \phi_4 = \phi_5 = \frac{2\pi + \alpha_5 - \theta}{5}$$

and

$$\|\overline{\mathbf{P}_0\mathbf{T}_0}\| = \frac{1}{2}\|\overline{\mathbf{P}_0\mathbf{P}_1}\|.$$

Method M_7. If the tangent angles θ and φ satisfy the condition

$$(\theta, \varphi) \in [2\pi/3, \pi] \times [2\pi/3, \pi]$$

(see Fig. 10(e)), then the joints and their tangent vectors, $\{\mathbf{T}_0, \mathbf{T}_1, \mathbf{V}_{t_0}, \mathbf{V}_{t_1}\}$, are determined by requiring

$$\phi_1 = \phi_2 = \phi_3 = \phi_4 = \phi_5 = \phi_6 = \phi = \frac{2\pi + \varphi - \theta}{6}$$

and

$$\|\overline{\mathbf{P}_0\mathbf{T}_0}\| = \frac{1}{2}\|\overline{\mathbf{P}_0\mathbf{P}_1}\|.$$

Theorem 6. *Given two endpoints $\mathbf{P}_0$ and $\mathbf{P}_1$ and two endpoint tangent vectors $\mathbf{V}_0$ and $\mathbf{V}_1$, if $\mathbf{Q}_0(t)$ is constructed with respect to $\{\mathbf{P}_0, \mathbf{T}_0, \mathbf{V}_0, \mathbf{V}_{t_0}\}$, $\mathbf{Q}_1(t)$ is constructed with respect to $\{\mathbf{T}_0, \mathbf{T}_1, \mathbf{V}_{t_0}, \mathbf{V}_{t_1}\}$, and $\mathbf{Q}_2(t)$ is constructed with respect to $\{\mathbf{T}_1, \mathbf{V}_1, \mathbf{V}_{t_1}, \mathbf{V}_1\}$, where the joints and their tangent vectors, $\{\mathbf{T}_0, \mathbf{T}_1, \mathbf{V}_{t_0}, \mathbf{V}_{t_1}\}$, are determined by one of the above methods (M_3–M_7) then $\mathbf{Q}_0(t)$, $\mathbf{Q}_1(t)$ and $\mathbf{Q}_2(t)$ all satisfy the conditions of Theorems 3 and 4.*

Proof. We prove the case for method M_4 only. The other cases can be proved similarly.

In Fig. 10(b), let $\mathbf{T}$ be the intersection point of the line segment $\overline{\mathbf{T}_0\mathbf{T}_1}$ and the line segment $\overline{\mathbf{P}_0\mathbf{P}_1}$, θ_1 be the counterclockwise angle from the vector $\overrightarrow{\mathbf{P}_0\mathbf{P}_1}$ to the vector $\overrightarrow{\mathbf{P}_0\mathbf{T}_0}$, θ_2 be the counterclockwise angle from the vector $\overrightarrow{\mathbf{TT}_0}$ to the vector $\overrightarrow{\mathbf{TP}_0}$, θ_3 be the counterclockwise angle from the vector $\overrightarrow{\mathbf{TT}_1}$ to the vector $\overrightarrow{\mathbf{TP}_1}$, and θ_4 be the counterclockwise angle from the vector $\overrightarrow{\mathbf{P}_1\mathbf{P}_0}$ to the vector $\overrightarrow{\mathbf{P}_1\mathbf{T}_1}$. We have

$$\phi_2 + \phi_3 = \theta_1 + \theta_2 \quad \text{and} \quad \phi_4 + \phi_5 = \theta_3 + \theta_4.$$

Because $\alpha_1 = \theta/2$ and $\alpha_5 = \varphi/2$, one sees that

$$\phi_1 = \theta_1 \in l[\pi/6, \pi/3] \quad \text{and} \quad \phi_6 = \theta_4 \in [0, \pi/3].$$

On the other hand, since

$$\|\overline{\mathbf{P}_0\mathbf{T}_0}\| + \|\overline{\mathbf{T}_0\mathbf{T}}\| + \|\overline{\mathbf{TT}_1}\| + \|\overline{\mathbf{T}_1\mathbf{P}_1}\| > \|\overline{\mathbf{P}_0\mathbf{T}}\| + \|\overline{\mathbf{TP}_1}\| = \|\overline{\mathbf{P}_0\mathbf{P}_1}\|$$

and, by hypothesis,

$$\|\overline{\mathbf{P}_0\mathbf{T}_0}\| = \frac{1}{3}\|\overline{\mathbf{P}_0\mathbf{P}_1}\| \quad \text{and} \quad \|\overline{\mathbf{T}_1\mathbf{P}_1}\| = \frac{1}{6}\|\overline{\mathbf{P}_0\mathbf{P}_1}\|,$$

one must have

$$\|\overline{\mathbf{T}_0\mathbf{T}}\| > \|\overline{\mathbf{P}_0\mathbf{T}_0}\| \quad \text{or} \quad \|\overline{\mathbf{TT}_1}\| > \|\overline{\mathbf{T}_1\mathbf{P}_1}\|$$

or, equivalently,

$$\theta_2 < \theta_1 \quad \text{or} \quad \theta_3 < \theta_4.$$

Therefore, we see that

$$\theta_2 = \theta_3 \in [0, \pi/3].$$

Consequently,

$$\phi_2 = \phi_3 = \frac{\theta_1 + \theta_2}{2} \in [0, \pi/3]$$

and

$$\phi_4 = \phi_5 = \frac{\theta_3 + \theta_4}{2} \in [0, \pi/3].$$

With

$$\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6 \in [0, \pi/3],$$

we have

$$\cos \phi_1, \cos \phi_2, \cos \phi_3, \cos \phi_4, \cos \phi_5, \cos \phi_6 \in (1/3, 1].$$

Thus, all OGH segments generated by $\mathbf{M}_4$ satisfy the conditions of Theorem 4 and, consequently, the conditions of Theorem 3 as well. □

4. Symmetry-based extension schemes

Let $\mathbf{M}_0$ denote the method that generates one-segment COH curves (i.e., OGH curves) satisfying the conditions of Theorem 3. Thus, totally, we have eight methods for generating COH curves. These methods are not enough to cover all the possible cases of the tangent angles since they do not cover the entire $\theta\varphi$-space, $[0, 2\pi) \times [0, 2\pi)$. Instead of introducing more methods to generate two-segment or three-segment COH curves, in this section, we will present symmetry-based techniques to extend the coverage of the above methods so that all cases can be considered. Three schemes will be used.

As shown in Fig. 11(a), the first scheme creates a new curve $\mathbf{Q}^{\mathrm{T}}(t)$ symmetric to the original curve $\mathbf{Q}(t)$ with respect to the base line of $\mathbf{Q}(t)$. The second scheme constructs a new curve $\mathbf{Q}^{\mathrm{R}}(t)$ that reverses the original curve $\mathbf{Q}(t)$ (Fig. 11(b)). The third one is a mixture of the first two schemes, i.e., perform the first scheme then the second scheme, or the second scheme then the first scheme ($\mathbf{Q}^{\mathrm{TR}}(t)$ or $\mathbf{Q}^{\mathrm{RT}}(t)$, Fig. 11(c)). The result of the third scheme is order independent, i.e., the result will be the same no matter which scheme is applied first (i.e., $\mathbf{Q}^{\mathrm{RT}}(t) = \mathbf{Q}^{\mathrm{TR}}(t)$).

For a given method $\mathbf{M}_i$, we use $\mathbf{M}_i^{\mathrm{T}}$ to refer to the combination of $\mathbf{M}_i$ and the first scheme, $\mathbf{M}_i^{\mathrm{R}}$ to refer to the combination of $\mathbf{M}_i$ and the second scheme, and $\mathbf{M}_i^{\mathrm{RT}}$ to refer to the combination of $\mathbf{M}_i$ and the third scheme. Define the *applicable region* of a method as the tensor product of the applicable region of the start point tangent angle and the applicable region of the end point tangent angle that would ensure the satisfaction of Theorems 3 and 4 of the created curve. For example, the applicable region of method $\mathbf{M}_1$ is $[0, \pi/6] \times [\pi/3, 2\pi/3]$ (see Method $\mathbf{M}_1$). Then we have the following theorem.

Theorem 7. *The applicable regions of the methods* $\mathbf{M}_i$, $\mathbf{M}_i^{\mathrm{T}}$, $\mathbf{M}_i^{\mathrm{R}}$, *and* $\mathbf{M}_i^{\mathrm{RT}}$, $i = 0, 1, \ldots, 7$, *cover the entire* $[0, 2\pi) \times [0, 2\pi)$ *space.*
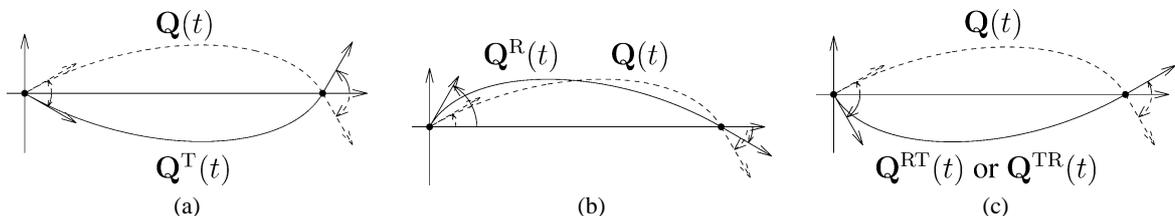

Fig. 11. Symmetry-based schemes. (a) First scheme. (b) Second scheme. (c) Third scheme.

Proof. For a given curve $\mathbf{Q}(t)$, let θ and φ be the counterclockwise angles of the start point tangent vector and the end point tangent vector with respect to the vector from the start point to the end point. Similarly, we have θ^T and φ^T for the curve $\mathbf{Q}^\mathrm{T}(t)$, θ^R and φ^R for the curve $\mathbf{Q}^\mathrm{R}(t)$, and θ^RT and φ^RT for the curve $\mathbf{Q}^\mathrm{RT}(t)$. These angles satisfy the following formulae:

$$\begin{cases} \theta^\mathrm{T} = 2\pi - \theta, \\ \varphi^\mathrm{T} = 2\pi - \varphi, \\ \theta^\mathrm{R} = 2\pi - \varphi, \\ \varphi^\mathrm{R} = 2\pi - \theta, \end{cases} \quad \text{and} \quad \begin{cases} \theta^\mathrm{RT} = \varphi, \\ \varphi^\mathrm{RT} = \theta. \end{cases}$$

Note that if $[\theta_0, \theta_1] \times [\varphi_0, \varphi_1]$ is the applicable region of method $\mathbf{M}_i$, then

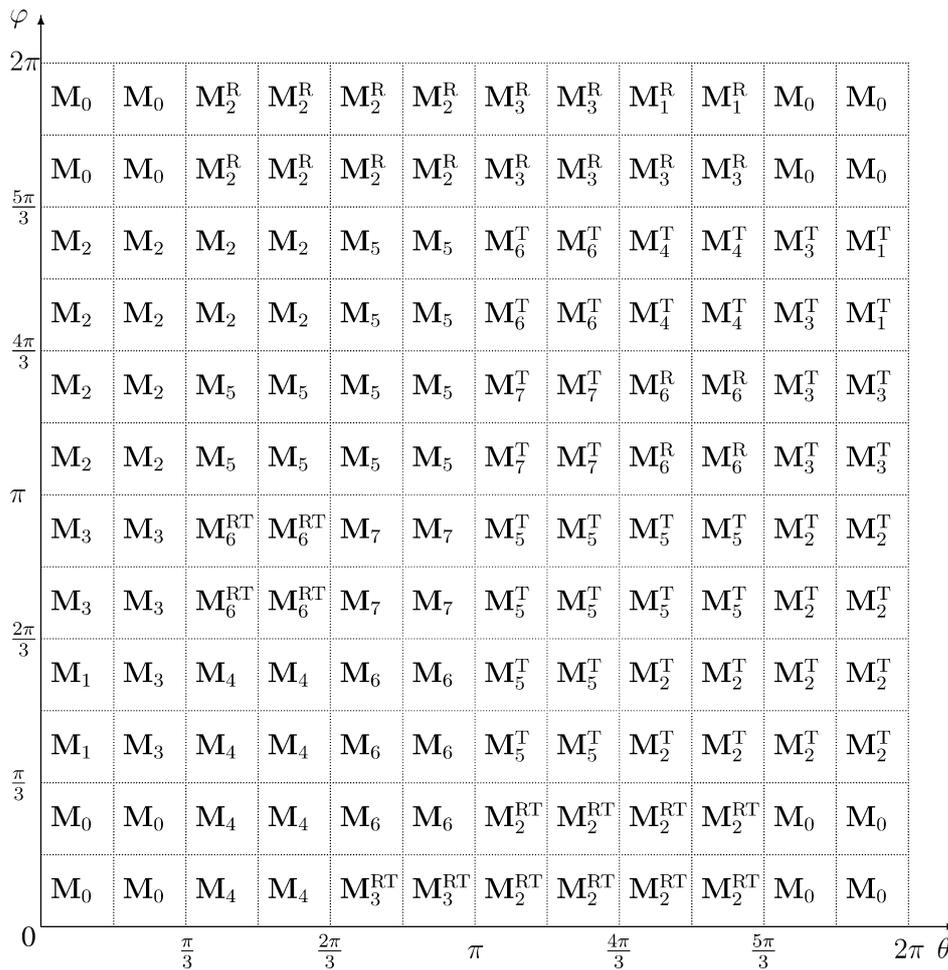$$[2\pi - \theta_1, 2\pi - \theta_0] \times [2\pi - \varphi_1, 2\pi - \varphi_0]$$

Fig. 12. Applicable chart of the methods $\mathbf{M}_i$, $\mathbf{M}_i^\mathrm{T}$, $\mathbf{M}_i^\mathrm{R}$, and $\mathbf{M}_i^\mathrm{RT}$, $i = 0, 1, \ldots, 7$.

is the applicable region of $\mathbf{M}_i^{\mathrm{T}}$;

$$[2\pi - \varphi_1, 2\pi - \varphi_0] \times [2\pi - \theta_1, 2\pi - \theta_0]$$

is the applicable region of $\mathbf{M}^{\mathrm{R}}$ and

$$[\varphi_0, \varphi_1] \times [\theta_0, \theta_1]$$

is the applicable region of $\mathbf{M}_i^{\mathrm{RT}}$.

Using these formulae and relationships, it is easy to see that the applicable regions of the methods $\mathbf{M}_i$, $\mathbf{M}_i^{\mathrm{T}}$, $\mathbf{M}_i^{\mathrm{R}}$, and $\mathbf{M}_i^{\mathrm{RT}}$, $i = 0, 1, \ldots, 7$, satisfy the chart shown in Fig. 12. In this chart, if a region has more than one available method, only one is shown to avoid clustering. □

5. Discussion and conclusions

A new class of curves called *optimized geometric Hermite* (OGH) curves is presented. An OGH curve is mathematically and geometrically smooth, i.e., loop-, cusp- and fold-free and with minimum strain energy, if the geometric smoothness conditions and the tangent direction preserving conditions on the tangent angles are satisfied. If the given tangent vectors do not satisfy these constraints, one can use a *two-segment* or *three-segment composite optimized geometric Hermite* (COH) curve instead. The construction techniques of the *two-segment* and *three-segment composite optimized geometric Hermite* (COH) curves guarantee that each segment of the curve automatically satisfies the tangent angle constraints and consequently is both mathematically and geometrically smooth. Symmetry-based schemes have also been given to extended the coverage of the presented methods so that all cases of the tangent angles can be considered.

In general, a COH curve has only G^1 continuity. However, in some cases a COH curve could be G^2 continuous. Fig. 13 gives such an example, where the two given end points are $\mathbf{P}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\mathbf{P}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$; and the tangent vectors at those two end points are $\mathbf{V}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\mathbf{V}_1 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$. As show in Fig. 13(b), two segments in a COH curve share the same curvature $k = -2.667$ at the joint $\mathbf{T}$. This result is very close to the high accuracy geometric Hermite interpolant (de Boor et al., 1987) whose curvatures at the end points are both -1.5 (Fig. 13(a)). As a comparison, Fig. 13(c) shows several Pythagorean hodograph quintics (Farouki and Neff, 1995) with the same end-point conditions. The main concerns of these three papers ((de Boor et al., 1987; Farouki and Neff, 1995) and this paper) are quite different. (de Boor et
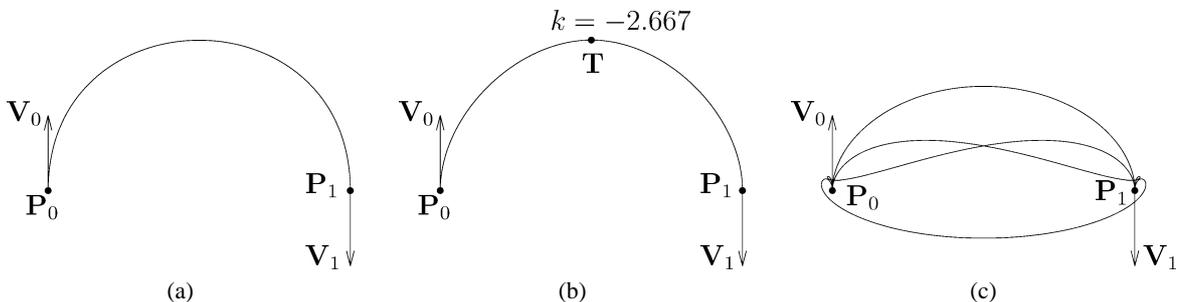

Fig. 13. Example 4: interpolated by (a) a high accuracy geometric Hermite interpolant (de Boor et al., 1987), (b) a G^2 COH curve with the method $\mathbf{M}_2$, and (c) Pythagorean hodograph quintics (totally 4 curves) (Farouki and Neff, 1995).
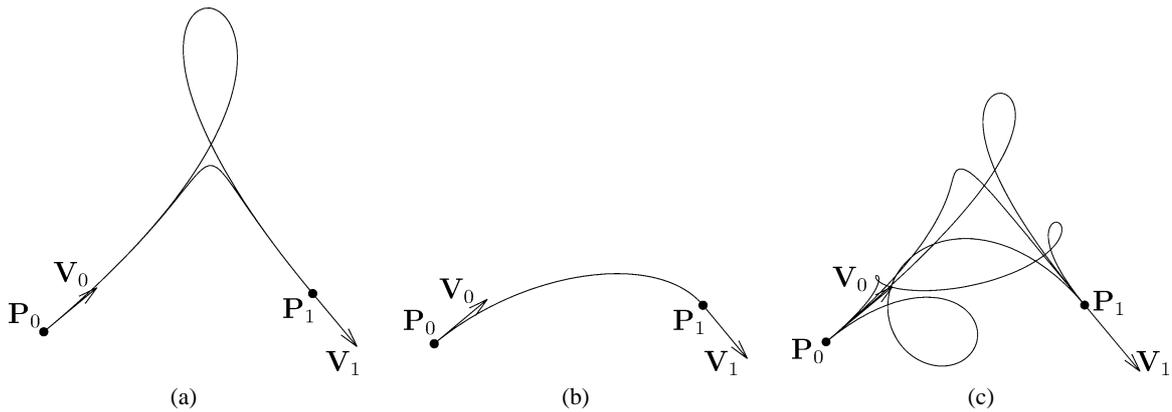
Fig. 14. Example 5: different contributions among (a) high accuracy geometric Hermite interpolants (totally 2 curves) (de Boor et al., 1987), (b) an OGH curve, and (c) Pythagorean hodograph quintics (totally 4 curves) (Farouki and Neff, 1995).

al., 1987) puts its focus on G^2 continuity and sixth order approximation accuracy. It does not deal with loops, cusps and folds. As shown in Fig. 14(a), a high accuracy geometric Hermite interpolant (de Boor et al., 1987) may have a loop, which is avoided by this paper. (Farouki and Neff, 1995) pays most of its attention on Pythagorean hodograph (PH) condition. (Farouki and Neff, 1995) points out that the simplest PH curves with first-order Hermite conditions are quintics. Therefore, the degree required by (Farouki and Neff, 1995) is larger than that in (de Boor et al., 1987) and this paper. This paper tries to produce cubic curves with pleasing shapes under all kinds of given conditions, using strain energy minimization technique, composition methods and symmetry skills. Thus, as shown in Fig. 14, the shape of the OGH curve is more pleasing and more natural than the shape of other interpolants. The original data of Fig. 14 is from (Farouki and Neff, 1995). The curves in Fig. 14(c) are the same as the curves provided in Fig. 3 of (Farouki and Neff, 1995).

The presented methods can be used in applications such as *shape design* and *curve/surface fairing* in geometric modeling. For example, in fairing the abnormal regions of a NURBS surface (Zhang and Cheng, 1998) uses Hermite curves to replace abnormal portions of the highlight lines in those regions and deforms the surface so that the new surface would have the modified highlight lines as the new highlight lines. Traditional Hermite interpolation technique is used to construct those Hermite curves. As has been shown in previous sections, such curves could contain undesired features such as loops, cusps or even folds. Obviously, the deformed surface would not have the best possible shape if undesired features are contained in the constructed Hermite curves. One should use optimized geometric Hermite (OGH) curves in the above highlight line modification process instead. This would not only avoid singular cases such as cusps or loops in the modified highlight lines, but also maintain low strain energy of the resultant curves. It might be possible to use COH curves in curve interpolation problem. This will be a work to look at in the future.

Acknowledgements

References

Chen, Y. 1993. Highlight lines for surface quality control and shape manipulation. PhD Thesis, Department of Naval Architecture and Marine Engineering, The University of Michigan.

Chen, Y., Beier, K.-P., Papageorgiou, D., 1997. Direct highlight line modification on NURBS surfaces. Computer Aided Geometric Design 14 (6), 583–601.

de Boor, C., Höllig, K., Sabin, M., 1987. High accuracy geometric Hermite interpolation. Computer Aided Geometric Design 4, 269–278.

Degen, W.L.F., 1993. High accurate rational approximation of parametric curves. Computer Aided Geometric Design 10, 293–313.

Farouki, R.T., Neff, C.A., 1995. Hermite interpolation by Pythagorean hodograph quintics. Math. Comput. 64 (212), 1589–1609.

Farouki, R.T., Sakkalis, T., 1990. Pythagorean hodographs. IBM J. Res. Develop. 34, 736–752.

Höllig, K., Koch, J., 1995. Geometric Hermite interpolation. Computer Aided Geometric Design 12 (6), 567–580.

Höllig, K., Koch, J., 1996. Geometric Hermite interpolation with maximal order and smoothness. Computer Aided Geometric Design 13 (8), 681–695.

Horn, B.K.P., 1983. The curve of least energy. ACM Trans. Math. Software 9 (4), 441–460.

Meek, D.S., Walton, D.J., 1997a. Geometric Hermite interpolation with Tschirnhausen cubics. J. Comput. Appl. Math. 81 (2), 299–309.

Meek, D.S., Walton, D.J., 1997b. Hermite interpolation with Tschirnhausen cubic spirals. Computer Aided Geometric Design 14 (7), 619–635.

Reif, U., 1999. On the local existence of the quadratic geometric Hermite interpolant. Computer Aided Geometric Design 16 (3), 217–221.

Schaback, R., 1998. Optimal geometric Hermite interpolation of curves. In: Dahlen, M., Lyche, T., Schumaker, L.L. (Eds.), Mathematical Methods for Curves and Surface II, pp. 1–12.

Wang, X., Cheng, F., 1997. Surface design based on Hermite spline interpolation with tension control and optimal twist vectors. Neural Parallel & Scientific Computations 5 (1&2), 37–54.

Yong, J.-H., Hu, S.-M., Sun, J.-G., 1999. A note on approximation of discrete data by G^1 arc splines. Computer-Aided Design 31 (14), 911–915.

Zhang, C., Cheng, F., 1998. Removing local irregularities of NURBS surfaces by modifying highlight lines. Computer-Aided Design 30 (12), 923–930.

Zhang, C., Zhang, P., Cheng, F., 2001. Fairing spline curves and surfaces by minimizing energy. Computer-Aided Design 33 (13), 913–923.