

HERMITE INTERPOLATION BY PYTHAGOREAN HODOGRAPH QUINTICS

R. T. FAROUKI AND C. A. NEFF

ABSTRACT. The *Pythagorean hodograph* (PH) curves are polynomial parametric curves $\{x(t), y(t)\}$ whose hodograph (derivative) components satisfy the Pythagorean condition $x'^2(t) + y'^2(t) \equiv \sigma^2(t)$ for some polynomial $\sigma(t)$. Thus, unlike polynomial curves in general, PH curves have arc lengths and offset curves that admit *exact* rational representations. The lowest-order PH curves that are sufficiently flexible for general interpolation/approximation problems are the quintics. While the PH quintics are capable of matching arbitrary first-order Hermite data, the solution procedure is not straightforward and furthermore does not yield a unique result—there are always *four* distinct interpolants (of which only one, in general, has acceptable “shape” characteristics). We show that formulating PH quintics as complex-valued functions of a real parameter leads to a compact Hermite interpolation algorithm and facilitates an identification of the “good” interpolant (in terms of minimizing the *absolute rotation number*). This algorithm establishes the PH quintics as a viable medium for the design or approximation of free-form curves, and allows a one-for-one substitution of PH quintics in lieu of the widely-used “ordinary” cubics.

1. INTRODUCTION

The hodograph of a differentiable parametric curve $\mathbf{r}(t) = \{x(t), y(t)\}$ is the locus described by its derivative, $\mathbf{r}'(t) = \{x'(t), y'(t)\}$. Whereas the *geometric* properties of hodographs yield valuable information [1] for curve design and analysis problems, we are concerned here with polynomial curves

$$(1) \quad x(t) = \sum_{k=0}^n a_k t^k, \quad y(t) = \sum_{k=0}^n b_k t^k$$

whose hodographs exhibit a special *algebraic* property [10]—namely, they satisfy the Pythagorean condition

$$(2) \quad x'^2(t) + y'^2(t) \equiv \sigma^2(t),$$

for some polynomial $\sigma(t)$. These *Pythagorean hodograph* (PH) curves have two computational advantages over polynomial parametric curves in general: (i) the arc length of any segment $t \in [t_0, t_1]$ may be determined exactly—i.e., without numerical quadrature—through arithmetic operations on the curve coefficients $\{a_k, b_k\}$ and the end-point parameter values t_0, t_1 ; and (ii) the *offset curves* at

Received by the editor October 26, 1993 and, in revised form, September 22, 1994.
1991 *Mathematics Subject Classification*. Primary 68U07; Secondary 53A04.

©1995 American Mathematical Society

each signed distance d , defined by

$$(3) \quad \mathbf{r}_d(t) = \mathbf{r}(t) + d\mathbf{n}(t),$$

where $\mathbf{n}(t)$ is the unit normal to $\mathbf{r}(t)$, admit exact *rational* parameterizations. (Offset curves arise in a variety of CAD/CAM applications; the fact that general polynomial curves do not have rational offsets [7] has prompted a profusion of more or less heuristic approximation schemes; see [18].)

It may be feared that condition (2) will transpire to be unduly restrictive in terms of the shape flexibility of curves that enjoy the above attributes; such curves would then be of little practical value. As a preliminary counter-argument to this concern, we cite the fact [6] that, although the set of regular PH curves forms a *proper subset* of the set of all regular polynomial curves, the former actually has the *same cardinality* as the latter.

In this paper we are more directly concerned with *geometric* arguments for the shape flexibility of PH curves. Specifically, we shall focus on establishing the suitability of the *regular PH quintics* for free-form design problems based on interpolating discrete point/tangent data. In the standard C^2 cubic spline [2]—and in more sophisticated “shape-preserving” interpolation schemes, e.g., [12, 13]—a key construction is the first-order Hermite interpolant $\mathbf{r}(t)$ to prescribed end points and “tangent vectors”

$$(4) \quad \mathbf{r}(0), \mathbf{r}(1) \quad \text{and} \quad \mathbf{r}'(0), \mathbf{r}'(1)$$

on the unit interval $t \in [0, 1]$. Expressing $\mathbf{r}(t)$ in the Bernstein-Bézier form

$$(5) \quad \mathbf{r}(t) = \sum_{k=0}^n \mathbf{p}_k \binom{n}{k} (1-t)^{n-k} t^k$$

on $t \in [0, 1]$ greatly facilitates this construction [4]: one merely assigns the first and last pairs of the “control points” $\{\mathbf{p}_k\}$ according to

$$(6) \quad \begin{aligned} \mathbf{p}_0 &= \mathbf{r}(0), & \mathbf{p}_1 &= \mathbf{r}(0) + \frac{1}{n}\mathbf{r}'(0), \\ \mathbf{p}_{n-1} &= \mathbf{r}(1) - \frac{1}{n}\mathbf{r}'(1), & \mathbf{p}_n &= \mathbf{r}(1). \end{aligned}$$

Clearly, this consumes all degrees of freedom when $n = 3$, establishing the existence and uniqueness of cubic interpolants to first-order Hermite data.

The simplest PH curves that are capable of inflecting and of matching arbitrary first-order Hermite data are the quintics. The construction of such interpolants is more involved than the simple assignment (6), however, and *does not yield a unique solution* for given Hermite data—there are always four distinct possibilities! While all four are rigorously consistent with the prescribed data, we observe empirically that only one of the interpolants has overall shape properties in agreement with those intuitively suggested by the point/tangent data; the others exhibit a contorted “looping” behavior.

The goal of this paper is to present a detailed account of the construction of PH quintic Hermite interpolants, including the formulation of quantitative criteria for automatically identifying the “good” solution among the various possibilities. Our analysis will ultimately be based on representing PH curves not as ordered pairs $\{x(t), y(t)\}$ of real polynomials, but rather as *complex-valued polynomials* $x(t) + iy(t)$ of a real parameter. This is not mathematical

sophistry: appealing to the algebra of complex numbers greatly simplifies the solution procedure and offers clearer insight into the geometrical properties of the resulting interpolants. Moreover, the complex formulation allows an exceedingly compact implementation of the PH quintic Hermite interpolation scheme in high-level computer languages that admit complex arithmetic.

The PH quintic Hermite interpolation scheme has immediate application in the smooth interpolation of ordered sequences of points by piecewise PH curves. There is, however, another practical implication of this scheme that is perhaps more pertinent to the design of *individual* segments. The PH quintics are the subset of all degree-5 Bézier curves of the form (5) whose six control points $\mathbf{p}_0, \dots, \mathbf{p}_5$ satisfy certain nonlinear constraints. Unfortunately, these constraints are not all amenable to simple geometrical interpretations [6, 10] and it is impractical to attempt to construct PH quintics by seeking to impose them upon general quintic control polygons.

The Hermite interpolation scheme, however, may be regarded as a means of automatically “filling in” the interior points $\mathbf{p}_2, \mathbf{p}_3$ so as to yield a PH curve when only the exterior control points $\mathbf{p}_0, \mathbf{p}_1$ and $\mathbf{p}_4, \mathbf{p}_5$ are specified. Thus, the scheme allows a one-for-one replacement of “ordinary” cubics by PH quintics that have similar shape properties (provided the Hermite data are consistent with a smooth arc without inflections or loops—in the latter cases the PH quintic surrogate is often a “fairer” curve, in the sense of having a more even curvature distribution, than its cubic counterpart).

The organization of this paper is as follows. Section 2 reviews the fundamental properties of PH curves and the motivation underlying their introduction. In §3 we formulate the PH quintic Hermite interpolation problem and derive its explicit real solutions. We argue in §4 that the “ordinary” cubic Hermite interpolant is not a viable comparison reference for identifying the “good” PH quintic interpolant among the four possibilities. The algebra of complex numbers proves to be a concise and elegant medium for analyzing PH curves: in §5 we show that it offers a compact alternative to the (real) algorithm of §3, while in §6 a simple means of selecting the best interpolant is derived from the complex formulation. Finally, §7 summarizes our results and makes some comments on their practical use.

2. PYTHAGOREAN HODOGRAPH CURVES

We commence by briefly recapitulating the definition and basic properties of PH curves (see [5, 6, 10] for further details).

Definition 1. Let $u(t)$, $v(t)$, and $w(t)$ be nonzero real polynomials such that $u(t)$ and $v(t)$ are relatively prime and not both constants. Then a parametric curve whose derivative is of the form

$$(7) \quad x'(t) = w(t)[u^2(t) - v^2(t)] \quad \text{and} \quad y'(t) = 2w(t)u(t)v(t)$$

is called a Pythagorean hodograph (PH) curve.

For any PH curve, the quantity $x'^2(t) + y'^2(t)$ evidently coincides with the *perfect square* of the polynomial

$$(8) \quad \sigma(t) = w(t)[u^2(t) + v^2(t)],$$

and hence the three polynomials $x'(t)$, $y'(t)$, $\sigma(t)$ form a *Pythagorean triple* satisfying equation (2). Moreover, by a theorem on Pythagorean triples in unique

factorization domains [16], any hodograph $\{x'(t), y'(t)\}$ that satisfies equation (2) for some polynomial $\sigma(t)$ must be of the form (7).

If $w(t)$ is nonnegative, we may identify (8) as the *parametric speed*

$$\sigma(t) = |\mathbf{r}'(t)| = \sqrt{x'^2(t) + y'^2(t)} = \frac{ds}{dt}$$

of the PH curve, i.e., the rate of change of the curve arc length s with respect to the parameter t . Thus, PH curves differ from polynomial curves in general, in that ds/dt is just a *polynomial* function of t .

(Since the elements of a Pythagorean triple (x', y', σ) are of indeterminate sign, we can always *assume* that $w(t)$ is nonnegative if it has no real roots of odd multiplicity; this affects only the *sense* of the curve parameterization. If it does have real odd-multiplicity roots, we must replace $w(t)$ in (8) by its absolute value if we wish to identify $\sigma(t)$ with ds/dt . At any rate, we are primarily concerned with the *regular* odd-degree curves obtained by taking $w(t) = 1$. Real roots of $w(t)$ give rise to irregular points—usually *cusps*—on the PH curve, where $x'(t) = y'(t) = 0$.)

Remark 1. If $w(t)$ is nonnegative, the cumulative arc length $s(t)$ of the PH curve, measured from $t = 0$, is given by

$$s(t) = \int_0^t w(\tau)[u^2(\tau) + v^2(\tau)] d\tau,$$

which is clearly also a *polynomial* function of t .

It follows that the *total* arc length S of any PH curve over a given domain $t \in [t_0, t_1]$ can be expressed rationally in terms of the curve coefficients and the initial and final parameter values t_0 and t_1 .

Remark 2. The differential characteristics of a PH curve—the unit tangent \mathbf{t} and normal \mathbf{n} , and the signed curvature κ —are all *rational* functions of the parameter t . In terms of the polynomials u, v, w and their derivatives, we have (assuming w is nonnegative):

$$(9) \quad \mathbf{t} = \frac{(u^2 - v^2, 2uv)}{u^2 + v^2}, \quad \mathbf{n} = \frac{(2uv, v^2 - u^2)}{u^2 + v^2}, \quad \kappa = \frac{2(uv' - u'v)}{w(u^2 + v^2)^2}.$$

Thus, the offset (3) to a PH curve clearly has a *rational* parameterization. (The above expression for κ follows by substituting from (7) into the usual form $\kappa(t) = |\mathbf{r}'(t)|^{-3}[\mathbf{r}'(t) \times \mathbf{r}''(t)] \cdot \mathbf{z}$ for the curvature of a plane parametric curve [15], where \mathbf{z} is a unit vector orthogonal to the plane of $\mathbf{r}(t)$.)

These attractive properties of PH curves may be contrasted with those of the broader set of polynomial curves in general, for which the parametric speed is the irreducible radical $\sqrt{x'^2(t) + y'^2(t)}$, and hence the differential characteristics and offset curves are irrational, and the determination of arc length requires numerical quadrature.

Discounting freedoms of rigid motion, scaling, and reparameterization, it transpires [10] that there is actually just *one* PH cubic. It corresponds to a classical curve, first studied in 1692 by Ehrenfried Walther von Tschirnhaus as the *caustic* for reflection of light rays by a parabola. The PH quartics are somewhat more flexible, but they always have a cusp (not necessarily within the domain of interest) and share with Tschirnhausen's cubic the constraint of being

convex, i.e., the curvature is of constant sign [10]. Also, even-degree curves are not suited to symmetric end-point Hermite interpolation.

We must therefore look to the quintics if we desire free-form capability while retaining the advantages of PH curves. There are actually two distinct PH quintic types: in (7) we may choose u and v linear and w quadratic, or we may take quadratics for u and v and set $w = 1$. The former always exhibit cusps when w has real roots, and cannot inflect [10]. We therefore confine our attention to the latter:

Proposition 1. *A Bézier curve of degree 5 is a (noncuspidal) PH curve if and only if its control points can be expressed in the form*

$$\begin{aligned} \mathbf{p}_1 &= \mathbf{p}_0 + \frac{1}{5}(u_0^2 - v_0^2, 2u_0v_0), \\ \mathbf{p}_2 &= \mathbf{p}_1 + \frac{1}{5}(u_0u_1 - v_0v_1, u_0v_1 + u_1v_0), \\ (10) \quad \mathbf{p}_3 &= \mathbf{p}_2 + \frac{2}{15}(u_1^2 - v_1^2, 2u_1v_1) + \frac{1}{15}(u_0u_2 - v_0v_2, u_0v_2 + u_2v_0), \\ \mathbf{p}_4 &= \mathbf{p}_3 + \frac{1}{5}(u_1u_2 - v_1v_2, u_1v_2 + u_2v_1), \\ \mathbf{p}_5 &= \mathbf{p}_4 + \frac{1}{5}(u_2^2 - v_2^2, 2u_2v_2) \end{aligned}$$

in terms of real values (u_0, u_1, u_2) and (v_0, v_1, v_2) , where \mathbf{p}_0 is arbitrary.

Proof. Substituting $w(t) = 1$ and the quadratic polynomials

$$(11) \quad \begin{aligned} u(t) &= u_0(1-t)^2 + u_12(1-t)t + u_2t^2, \\ v(t) &= v_0(1-t)^2 + v_12(1-t)t + v_2t^2 \end{aligned}$$

into (7), we use the arithmetic procedures for polynomials in Bernstein form [9] to express $x'(t)$ and $y'(t)$ in the quartic Bernstein basis on $t \in [0, 1]$. Integrating $x'(t)$ and $y'(t)$ and making use of the relation

$$(12) \quad \int \binom{n}{k} (1-t)^{n-k} t^k dt = \frac{1}{n+1} \sum_{j=k+1}^{n+1} \binom{n+1}{j} (1-t)^{n+1-j} t^j$$

for $k = 0, \dots, n$ [9] then gives expressions (10) for the control points $\{\mathbf{p}_k\}$ in the representation (5) of the resulting PH curve. The initial control point \mathbf{p}_0 , corresponding to the constants of integration, may be freely chosen. \square

Bear in mind that *neither* (u_0, u_1, u_2) nor (v_0, v_1, v_2) may be set to $(0, 0, 0)$ and they may not *both* be of the form (k, k, k) with $k \neq 0$, corresponding to the requirements that $u(t)$ and $v(t)$ be nonzero and not both constants. Furthermore, the inequality

$$(13) \quad (u_2v_0 - u_0v_2)^2 \neq 4(u_0v_1 - u_1v_0)(u_1v_2 - u_2v_1)$$

must hold to ensure that $u(t)$ and $v(t)$ are relatively prime.

Lemma 1. *The PH quintic defined by (10) has either two real inflections or none, according to whether the quantity*

$$(14) \quad \Delta = (u_2v_0 - u_0v_2)^2 - 4(u_0v_1 - u_1v_0)(u_1v_2 - u_2v_1)$$

is positive or negative.

Proof. Substituting from (11) into the numerator $k(t) = u(t)v'(t) - u'(t)v(t)$ of expression (9) for the curvature, $k(t)$ is seen to be quadratic with Bernstein

coefficients $k_0 = 2(u_0v_1 - u_1v_0)$, $k_1 = -(u_2v_0 - u_0v_2)$, $k_2 = 2(u_1v_2 - u_2v_1)$. The corresponding discriminant $k_1^2 - k_0k_2$ is proportional to (14), and we note that $\Delta \neq 0$ by virtue of the constraint (13). Thus, if $\Delta > 0$, there are two real t values for which κ vanishes, while if $\Delta < 0$ there are none. \square

3. THE HERMITE INTERPOLATION PROBLEM

Discounting the arbitrary initial point \mathbf{p}_0 , we see from expressions (10) that the ten coordinates of the remaining control points $\mathbf{p}_1, \dots, \mathbf{p}_5$ depend upon the six quantities u_0, u_1, u_2 and v_0, v_1, v_2 . Thus, there exist four "constraint relations" among the control points, whose satisfaction distinguishes the PH quintics from quintic Bézier curves in general. These constraints have been explicitly derived elsewhere [6], based on the complex formulation used in §5 below. Unfortunately, unlike the case of PH cubics [10], they do not all admit fruitful interpretations in terms of the control-polygon geometry.

We are concerned here with the construction of PH quintic arcs on $t \in [0, 1]$ that have prescribed end points $\mathbf{r}(0), \mathbf{r}(1)$ and derivatives $\mathbf{r}'(0), \mathbf{r}'(1)$ at those end points. As previously noted in §1, this amounts to specifying the first and last two control points $\mathbf{p}_0, \mathbf{p}_1$ and $\mathbf{p}_4, \mathbf{p}_5$ —our goal is thus to compute coordinates for the middle points $\mathbf{p}_2, \mathbf{p}_3$ in such a manner that the "constraint relations" are automatically satisfied.

We shall rely exclusively on the Bernstein-Bézier form (5), which affords a numerically stable representation of finite arcs [8]. Note that the hodograph of (5) may itself be expressed as a Bézier curve of degree $n - 1$,

$$\mathbf{r}'(t) = n \sum_{k=0}^{n-1} \Delta \mathbf{p}_k \binom{n-1}{k} (1-t)^{n-1-k} t^k,$$

where the "control-polygon legs" $\Delta \mathbf{p}_k$ for $k = 0, \dots, n - 1$ are the forward differences of the control points:

$$(15) \quad \Delta \mathbf{p}_k = (\Delta x_k, \Delta y_k) = \mathbf{p}_{k+1} - \mathbf{p}_k.$$

In terms of the representation (5), we may phrase the PH quintic Hermite interpolation problem as follows:

Problem 1. Given arbitrary points $\mathbf{p}_0 \neq \mathbf{p}_1$ and $\mathbf{p}_4 \neq \mathbf{p}_5$, can two additional points \mathbf{p}_2 and \mathbf{p}_3 be found such that all six are expressible in the form (10) for real values of (u_0, u_1, u_2) and (v_0, v_1, v_2) ?

The answer is in the affirmative—indeed, there is a multiplicity of choices for \mathbf{p}_2 and \mathbf{p}_3 . In order to demonstrate this, we need the following result:

Lemma 2. For all real values a and b , the real solutions to the equations

$$(16) \quad u^2 - v^2 = a \quad \text{and} \quad 2uv = b$$

may be expressed in the form

$$(17) \quad (u, v) = \pm \left(\sqrt{\frac{1}{2}(c+a)}, \text{sign}(b)\sqrt{\frac{1}{2}(c-a)} \right),$$

where $c = \sqrt{a^2 + b^2}$ and we take $\text{sign}(b) = \pm 1$ when $b = 0$.

Proof. Suppose first that $a \neq 0$ and $b \neq 0$. Then $u \neq 0$, and from the second equation we may substitute $v = b/2u$ into the first, giving the biquadratic

$$(18) \quad u^4 - au^2 - \frac{1}{4}b^2 = 0$$

for u . Setting $c = \sqrt{a^2 + b^2}$, we then have $u^2 = \frac{1}{2}(a \pm c)$, and since $a - c < 0$ and $a + c > 0$, the real solutions to (18) are seen to be

$$u = \pm \sqrt{\frac{1}{2}(c + a)}.$$

We obtain the corresponding values of v by substituting into $v = b/2u$:

$$v = \pm \frac{b}{2\sqrt{\frac{1}{2}(c + a)}} = \pm \text{sign}(b) \sqrt{\frac{1}{2}(c - a)}.$$

Thus, when a and b are both nonzero, the real solutions agree with (17). We now show that (17) also gives the real solutions in cases where a or b vanishes. First, if $a = b = 0$, the only solution is $(u, v) = (0, 0)$, which clearly agrees with (17). If $a = 0 \neq b$, the first equation gives $v = \pm u$, and on substituting into the second, we obtain $\pm 2u^2 = b$, which has real solutions $u = \pm \sqrt{\frac{1}{2}|b|}$. Since $b \neq 0$, corresponding v values are obtained from $v = b/2u$, giving

$$(u, v) = \pm \left(\sqrt{\frac{1}{2}|b|}, \text{sign}(b) \sqrt{\frac{1}{2}|b|} \right)$$

consistent with (17). Finally, if $a \neq 0 = b$, either u or v must vanish and the solutions to (16) are thus of the form

$$(u, v) = \begin{cases} \pm(\sqrt{a}, 0) & \text{when } a > 0, \\ \pm(0, \sqrt{-a}) & \text{when } a < 0. \end{cases}$$

With the convention that $\text{sign}(b) = \pm 1$ when $b = 0$, the above may also be regarded as a special case of equation (17). \square

Geometrically, the solutions to equations (16) are the intersections of two rectangular hyperbolae in the (u, v) plane, both centered on the origin but with asymptotes at 45° to each other; when $a = 0$ or $b = 0$ these hyperbolae degenerate into the pairs of lines defined by their asymptotes.

Proposition 2. *Problem 1 always has real solutions, given by the following values for the coefficients (u_0, u_1, u_2) and (v_0, v_1, v_2) :*

$$(19) \quad \begin{aligned} (u_0, v_0) &= \pm \sqrt{\frac{5}{2}} \left(\sqrt{|\Delta \mathbf{p}_0| + \Delta x_0}, \text{sign}(\Delta y_0) \sqrt{|\Delta \mathbf{p}_0| - \Delta x_0} \right), \\ (u_2, v_2) &= \pm \sqrt{\frac{5}{2}} \left(\sqrt{|\Delta \mathbf{p}_4| + \Delta x_4}, \text{sign}(\Delta y_4) \sqrt{|\Delta \mathbf{p}_4| - \Delta x_4} \right), \\ (u_1, v_1) &= -\frac{3}{4}(u_0 + u_2, v_0 + v_2) \pm \sqrt{\frac{1}{2}}(\sqrt{c + a}, \text{sign}(b)\sqrt{c - a}), \end{aligned}$$

where the quantities a , b , and $c = \sqrt{a^2 + b^2}$ are defined by

$$(20) \quad \begin{aligned} a &= \frac{9}{16}(u_0^2 - v_0^2 + u_2^2 - v_2^2) + \frac{5}{8}(u_0u_2 - v_0v_2) + \frac{15}{2}(x_4 - x_1), \\ b &= \frac{9}{8}(u_0v_0 + u_2v_2) + \frac{5}{8}(u_0v_2 + u_2v_0) + \frac{15}{2}(y_4 - y_1). \end{aligned}$$

Proof. Since $\mathbf{p}_0, \mathbf{p}_1$ and $\mathbf{p}_4, \mathbf{p}_5$ are specified, Lemma 2 can be immediately applied to the x - and y -components of the first and last equations in (10) with $(a, b) = (5\Delta x_0, 5\Delta y_0)$ and $(5\Delta x_4, 5\Delta y_4)$, respectively, to obtain the quoted expressions (19) for (u_0, v_0) and (u_2, v_2) , where $|\Delta \mathbf{p}_k| = \sqrt{(\Delta x_k)^2 + (\Delta y_k)^2}$. Writing $\mathbf{p}_4 - \mathbf{p}_1 = (\mathbf{p}_4 - \mathbf{p}_3) + (\mathbf{p}_3 - \mathbf{p}_2) + (\mathbf{p}_2 - \mathbf{p}_1)$ and substituting from (10) for each of the three terms on the right, we see that the quantities (u_1, v_1) satisfy the simultaneous equations

$$(21) \quad \begin{aligned} \frac{15}{2}(x_4 - x_1) &= u_1^2 - v_1^2 + \frac{3}{2}(u_0 + u_2)u_1 - \frac{3}{2}(v_0 + v_2)v_1 + \frac{1}{2}(u_0u_2 - v_0v_2), \\ \frac{15}{2}(y_4 - y_1) &= 2u_1v_1 + \frac{3}{2}(v_0 + v_2)u_1 + \frac{3}{2}(u_0 + u_2)v_1 + \frac{1}{2}(u_0v_2 + u_2v_0), \end{aligned}$$

and if we make the change of variables

$$\tilde{u}_1 = u_1 + \frac{3}{4}(u_0 + u_2) \quad \text{and} \quad \tilde{v}_1 = v_1 + \frac{3}{4}(v_0 + v_2),$$

we find that the transformed equations for \tilde{u}_1 and \tilde{v}_1 assume the form (16) with right-hand sides a and b given in terms of the known values of (u_0, v_0) and (u_2, v_2) by expressions (20). Hence, with $c = \sqrt{a^2 + b^2}$, the solutions to equations (21) are of the form given in the last expression of (19). \square

Remark 3. Three independent sign ambiguities arise in expressions (19). Thus, it might appear at first sight that there are *eight* PH quintics matching prescribed end points and tangents. However, if (u_0, u_1, u_2) and (v_0, v_1, v_2) are the coefficients obtained for a specific choice of signs, a careful inspection of (19) reveals that the converse choice simply yields $(-u_0, -u_1, -u_2)$ and $(-v_0, -v_1, -v_2)$. Since expressions (10) involve only homogeneous quadratic forms in the coefficients of $u(t)$ and $v(t)$, the control points they define are invariant under a sign reversal of those coefficients. Thus, there are actually just *four* distinct Hermite interpolants—they may be generated without replication by arbitrarily fixing the sign in any one of the three expressions (19) and exercising the sign freedoms of the other two.

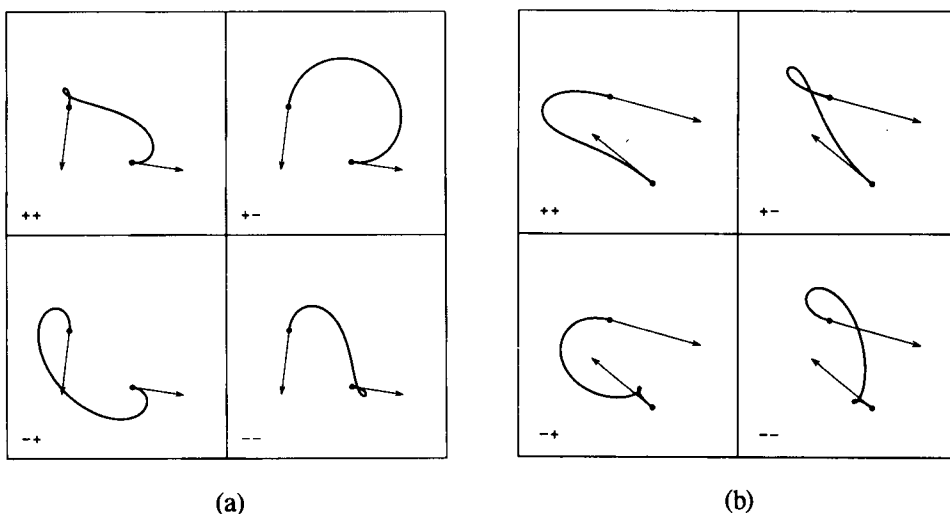


FIGURE 1. Examples of the four different PH quintic interpolants to given end points and end derivatives

Figure 1 shows examples of the four PH quintics that match given first-order Hermite data (we take a fixed sign in the (u_1, v_1) expression in (19) and obtain the four PH quintics from the sign choices $++$, $+-$, $-+$, $--$ in the expressions for (u_0, v_0) and (u_2, v_2)). While the four interpolants are all consistent with the Hermite data, they are otherwise markedly dissimilar. In view of this, we now draw attention to a surprising feature of these solutions:

Corollary 1. *The four PH quintics matching prescribed first-order Hermite data can be grouped into two pairs such that the members of each pair have identical total arc lengths, given by*

$$(22) \quad S = \frac{5}{8}(|\Delta \mathbf{p}_0| + |\Delta \mathbf{p}_4|) - \frac{1}{12}(u_0 u_2 + v_0 v_2) + \frac{2}{15}c.$$

Proof. Substitute from (19) into the expressions given in [5] for the Bernstein coefficients of the parametric speed $\sigma(t)$ and the total arc length S in terms of u_0, u_1, u_2 and v_0, v_1, v_2 to obtain expression (22). It can then be verified from (22) that the value of S is unaltered upon simultaneously reversing the choice of signs for (u_0, v_0) and (u_2, v_2) in equations (19). \square

4. COMPARISON WITH CUBICS

Among the four PH interpolants in the cases shown in Figure 1, one choice of signs seems to give a solution with more “reasonable” shape characteristics than the others. However, the appropriate choice varies with the specified Hermite data in a manner that is not easy to fathom a priori. Thus, we need to formulate a quantitative criterion for identifying the “good” solution.

We begin by comparing shape properties of the four PH quintic Hermite interpolants with that of the *unique* “ordinary” cubic Hermite interpolant. In taking this approach, we are motivated simply by the current widespread use of cubic Hermite interpolation; it is not our intent to impute “ideal” shape characteristics to the latter. Indeed, we shall ultimately be obliged to *reject* comparison with cubics as a basis for identifying the “good” PH quintics, but in view of the inevitability of comparisons between cubic and PH quintic Hermite interpolation schemes, the following discussion seems warranted.

In Figure 2 (next page) we show the *Gauss maps* [17] for the curves in Figure 1, which describe the rotation of the tangent \mathbf{t} as the curves are traversed. Since the initial and final tangent directions are fixed by the Hermite data, the Gauss maps of the various PH quintic interpolants clearly differ from each other, and from that of the cubic, only in the manner in which they “wind” between these limits (inflections correspond to sudden reversals of the Gauss map). We may characterize the *net* amount of winding of the Gauss map as follows:

Definition 2. The rotation number \mathcal{R} of a plane C^2 curve $\mathbf{r}(t) = \{x(t), y(t)\}$ defined on $t \in [0, 1]$ is given by

$$(23) \quad \mathcal{R} = \frac{1}{2\pi} \int_0^1 \kappa(t) |\mathbf{r}'(t)| dt.$$

Rotation numbers are a classical means [20] of characterizing the global shape properties of curves. They indicate the fraction of a full revolution that the curve tangent \mathbf{t} , or normal \mathbf{n} , executes upon traversing a curve (this follows from the familiar interpretation, $\kappa = d\theta/ds$, of the curvature as the derivative

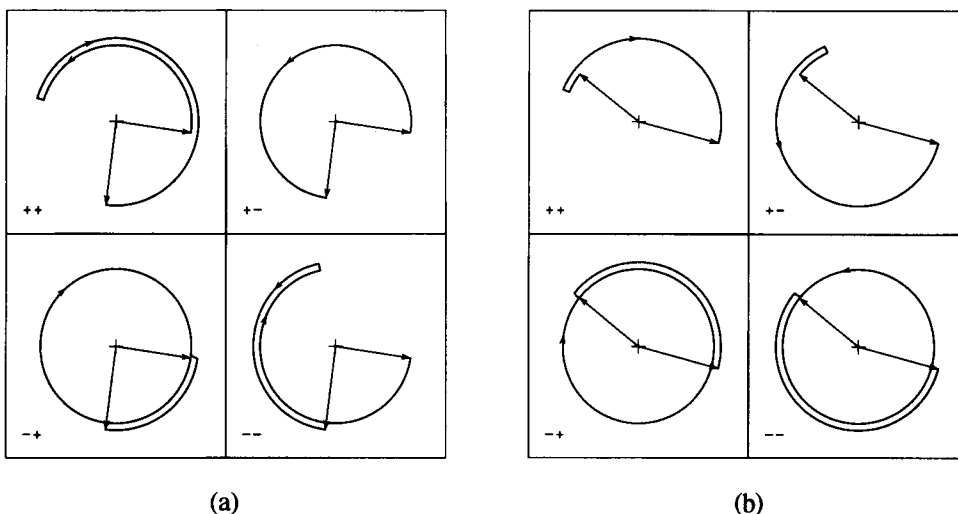


FIGURE 2. Gauss maps of the PH quintic Hermite interpolants shown in Figure 1

with respect to arc length s of the angle θ that \mathbf{t} or \mathbf{n} makes with a fixed direction). For simple closed curves, the “theorem of turning tangents” [3] states that $\mathcal{R} = \pm 1$, depending on the curve orientation. Here we are concerned with open curves, for which \mathcal{R} is generally not an integral value.

Definition 3. The Cauchy index $I_a^b \Phi(t)$ of a rational function $\Phi(t)$ on an interval $t \in [a, b]$ is equal to the number of poles at which $\Phi(t)$ jumps from $-\infty$ to $+\infty$ minus the number at which it jumps from $+\infty$ to $-\infty$, as t increases from a to b .

Note that poles at the interval end points a and b do not contribute to $I_a^b \Phi(t)$. Also, if $\Phi(t) = q(t)/p(t)$, where $\gcd(p, q) = 1$, only odd-multiplicity roots of the denominator $p(t)$ contribute to the Cauchy index [14].

Lemma 3. Let $p(t)$ and $q(t)$ be polynomials, not both constants, such that $\gcd(p, q) = 1$. Then the value of the definite integral

$$J = \int_a^b \frac{p(t)q'(t) - p'(t)q(t)}{p^2(t) + q^2(t)} dt$$

is given in terms of the rational function $\Phi(t) = q(t)/p(t)$ by

$$J = \tan^{-1} \Phi(b) - \tan^{-1} \Phi(a) - \pi I_a^b \Phi(t).$$

Proof. We set $\Phi = q/p$, so that $d\Phi/dt = (pq' - p'q)/p^2$, and hence we have

$$(24) \quad \int \frac{pq' - p'q}{p^2 + q^2} dt = \int \frac{d\Phi}{1 + \Phi^2} = \tan^{-1} \Phi,$$

where we take $\tan^{-1} \Phi \in [-\pi/2, +\pi/2]$. Note, however, that the change of variables $t \rightarrow \Phi = q/p$ does not define a bijective map from $t \in [a, b]$ to $\Phi \in [\Phi(a), \Phi(b)]$ if $p(t)$ has roots between a and b . In such a case, the range of integration must be split at the roots $t_1, \dots, t_N \in (a, b)$ of $p(t)$, and (24)

can then be applied to each subinterval delineated by those roots:

$$(25) \quad J = \int_{\Phi(a)}^{\Phi_1^-} \frac{d\Phi}{1 + \Phi^2} + \sum_{k=1}^{N-1} \int_{\Phi_k^+}^{\Phi_{k+1}^-} \frac{d\Phi}{1 + \Phi^2} + \int_{\Phi_N^+}^{\Phi(b)} \frac{d\Phi}{1 + \Phi^2},$$

where Φ_k^- and Φ_k^+ are the left- and right-hand limits (either $-\infty$ or $+\infty$) of $\Phi(t)$ at the pole t_k . Making use of (24) and writing $\tan^{-1}\Phi_k^\pm = \text{sign}(\Phi_k^\pm)\pi/2$, we see that (25) can be expressed as

$$(26) \quad J = \tan^{-1}\Phi(b) - \tan^{-1}\Phi(a) + \frac{\pi}{2} \sum_{k=1}^N [\text{sign}(\Phi_k^-) - \text{sign}(\Phi_k^+)].$$

Now on traversing each pole t_k , the value of $\text{sign}(\Phi_k^-) - \text{sign}(\Phi_k^+)$ is -2 or $+2$ if $\Phi(t)$ jumps from $-\infty$ to $+\infty$ or from $+\infty$ to $-\infty$, respectively, but is otherwise zero. The sum in (26) is thus simply $-2I_a^b\Phi(t)$. \square

In general, $I_a^b\Phi(t)$ is computed by invoking the Euclidean algorithm to construct a Sturm sequence for $p(t)$ and $q(t)$, and then taking the difference between the number of sign variations of this sequence evaluated at $t = a$ and $t = b$ [11]. To compare the rotation numbers of PH quintics and "ordinary" cubics, however, a simpler approach will suffice.

Corollary 2. *The rotation number of the PH curve defined on $t \in [0, 1]$ by (7) with $w(t) = 1$ may be expressed in terms of $\Psi(t) = v(t)/u(t)$ as*

$$(27) \quad \mathcal{R} = \frac{\tan^{-1}\Psi(1) - \tan^{-1}\Psi(0)}{\pi} - I_0^1\Psi(t).$$

Similarly, for the general polynomial curve (1) on $t \in [0, 1]$ we have, in terms of $\Theta(t) = y'(t)/x'(t)$,

$$(28) \quad \mathcal{R} = \frac{1}{2} \left[\frac{\tan^{-1}\Theta(1) - \tan^{-1}\Theta(0)}{\pi} - I_0^1\Theta(t) \right].$$

Proof. The above expressions for \mathcal{R} follow directly from Lemma 3 and the formulae $\kappa = 2(uv' - u'v)/(u^2 + v^2)^2$ and $|\mathbf{r}'| = u^2 + v^2$, or $\kappa = (x'y'' - x''y')/(x'^2 + y'^2)^{3/2}$ and $|\mathbf{r}'| = (x'^2 + y'^2)^{1/2}$, as appropriate. \square

For cubics and PH quintics, $\Theta(t)$ and $\Psi(t)$ are evidently just *quadratic* rational functions. Now if $\Phi(t) = q(t)/p(t)$, where $p(t)$ and $q(t)$ are relatively prime quadratics, $I_a^b\Phi(t)$ can be computed by inspecting the sign of $q(t)/p'(t)$ at the *simple* roots of $p(t)$ between a and b . For a quadratic $p(t)$, elementary arguments based on the "variation-diminishing property" [19] show that if \mathcal{V} is the number of sign variations in the Bernstein coefficients b_0, b_1, b_2 of $p(t)$ on $t \in [a, b]$ and $\Delta = b_1^2 - b_2b_0$ is its discriminant, then $p(t)$ has *one* or *two* simple roots between a and b if $\Delta > 0$ and $\mathcal{V} = 1$ or 2 ; otherwise, it has none.

The rational functions $\Psi(t) = v(t)/u(t)$ and $\Theta(t) = y'(t)/x'(t)$ occurring in (27) and (28) are closely related to the angle $\theta(t)$ that the curve tangent makes with the positive x -axis ($-\pi < \theta \leq +\pi$). For the PH curve, we have

$$\theta(t) = 2 \tan^{-1}\Psi(t),$$

while for the general polynomial curve,

$$\theta(t) = \begin{cases} \tan^{-1}\Theta(t) & \text{if } x'(t) \geq 0, \\ \tan^{-1}\Theta(t) + \text{sign}(y'(t))\pi & \text{if } x'(t) < 0. \end{cases}$$

Lemma 4. *The rotation number of an ordinary cubic satisfies $-1 < \mathcal{R} < +1$, while that of a PH quintic satisfies $-2 < \mathcal{R} < +2$.*

Proof. We show that, for relatively prime quadratics $p(t)$ and $q(t)$,

$$-2\pi < \int_a^b \frac{p(t)q'(t) - p'(t)q(t)}{p^2(t) + q^2(t)} dt < +2\pi$$

—the lemma then follows from the fact that $\kappa|\mathbf{r}'| = (x'y'' - x''y')/(x'^2 + y'^2)$ for the cubic, while $\kappa|\mathbf{r}'| = 2(uv' - u'v)/(u^2 + v^2)$ for the PH quintic. For brevity, we argue geometrically. The locus $x = p(t)$, $y = q(t)$ for $t \in [a, b]$ is a parabolic arc, and the radius vector from the origin to a point of that arc makes an angle θ with the positive x -direction, where $\tan \theta = q(t)/p(t)$. The integrand $(pq' - p'q)/(p^2 + q^2)$ is thus simply $d\theta/dt$, and it is evident geometrically that, regardless of the location or shape of the arc, the radius vector cannot execute more than a full revolution in tracing it out. \square

For given Hermite data, the rotation numbers of the four PH quintics differ from that of the cubic only by integral amounts, and one might hope that there is always a *unique* PH quintic whose rotation number coincides precisely with that of the cubic—this quintic being the “good” interpolant. However, some simple examples suffice to show that this conjecture is false:

Example 1. *There may not be a unique interpolant whose rotation number agrees with that of the cubic.* Consider the Hermite data

$$\begin{aligned} \mathbf{r}(0) &= (0, 5), & \mathbf{r}'(0) &= (25, -15), \\ \mathbf{r}(1) &= (-3, -4), & \mathbf{r}'(1) &= (25, -15), \end{aligned}$$

for which we have control points $\mathbf{p}_0 = (0, 5)$, $\mathbf{p}_1 = (5, 2)$ and $\mathbf{p}_4 = (-8, -1)$, $\mathbf{p}_5 = (-3, -4)$. Here, $\mathcal{R} = 0$ for the cubic, but there are *two* PH quintics for which \mathcal{R} vanishes, corresponding to the sign choices $++$ and $--$. As seen in Figure 3a, the latter choice clearly gives an undesirable interpolant. \square

Example 2. *Interpolants whose rotation numbers agree with that of the cubic do not necessarily have the best “shape” characteristics.* For the data

$$\begin{aligned} \mathbf{r}(0) &= (-6, -1), & \mathbf{r}'(0) &= (30, 25), \\ \mathbf{r}(1) &= (1, 0), & \mathbf{r}'(1) &= (25, -30), \end{aligned}$$

i.e., control points $\mathbf{p}_0 = (-6, -1)$, $\mathbf{p}_1 = (0, 4)$ and $\mathbf{p}_4 = (-4, 6)$, $\mathbf{p}_5 = (1, 0)$, we have $\mathcal{R} = 3/4$ for the cubic, while for the four PH quintics the sign choices $++$, $+-$, $-+$, $--$ give $\mathcal{R} = -1/4, 3/4, -5/4, 7/4$. In Figure 3b we see that the $+-$ solution, having the same \mathcal{R} value as the cubic, is indeed most similar in shape to the latter, but the $++$ solution (which inflects but does not intersect itself) is arguably a preferable interpolant. \square

It might be thought that, instead of using just the rotation number \mathcal{R} , a detailed comparison of the Gauss maps for the four PH quintics with that of

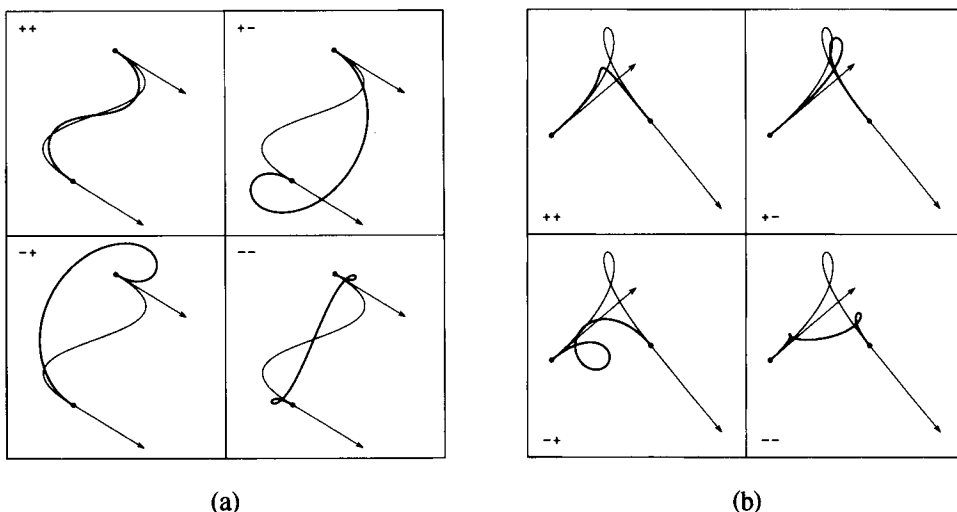


FIGURE 3. Comparison of PH quintic Hermite interpolants (bold curves) with the unique cubic interpolant (light curve) for: (a) the data of Example 1; and (b) the data of Example 2

the cubic might serve to identify the “good” PH quintic. One might hope, for example, for a unique PH quintic whose Gauss map is *isomorphic* to that of the cubic, i.e., there is a one-to-one correspondence between reversals (inflections) and a common sense of rotation between them. However, it is again possible to construct examples in which *none* of the PH quintics has the same number of inflections as the cubic, or in which the aesthetically pleasing PH quintic differs from the cubic in its number of inflections.

5. SOLUTION USING COMPLEX REPRESENTATIONS

Evidently, comparing PH quintic Hermite interpolants with their ordinary cubic counterparts is not a fruitful approach; the four PH quintics must be judged against “absolute” criteria in order to identify the one with the most desirable shape properties. To facilitate the formulation and analysis of such criteria, we now adopt a radically different approach: instead of regarding plane curves $\mathbf{r}(t)$ as pairs $\{x(t), y(t)\}$ of real polynomials, we express them as *complex-valued* polynomials $x(t) + iy(t)$ of a real parameter.

Remark 4. Henceforth we use bold characters (e.g., \mathbf{z}) for complex numbers and italics for explicitly real quantities. Of course, the former always admit vectorial interpretations, but if they appear in products, quotients, radicals, etc., $(\mathbf{z}_1\mathbf{z}_2, \mathbf{z}_1/\mathbf{z}_2, \sqrt{\mathbf{z}}, \dots)$, the algebra of complex numbers must be invoked.

Lemma 5. *In the complex representation, the regular PH curves correspond to those curves whose hodographs are perfect squares of complex polynomials having relatively prime real and imaginary parts.*

Proof. The square of the complex polynomial $\mathbf{p}(t) = u(t) + iv(t)$ is $\mathbf{p}^2(t) = u^2(t) - v^2(t) + i2u(t)v(t)$, whose real and imaginary parts are seen to be of the Pythagorean form (7) with $w(t) = 1$. Conversely, any complex hodograph having real and imaginary parts of the form (7) with $w(t) = 1$ is simply the

square of the complex polynomial $\mathbf{p}(t) = u(t) + iv(t)$. Moreover, we note that $\gcd(u, v) = 1$ if and only if $\gcd(u^2 - v^2, 2uv) = 1$. \square

Thus, the hodograph of any regular PH quintic may be expressed as

$$(29) \quad \mathbf{r}'(t) = \mathbf{k}[(t - \mathbf{a})(t - \mathbf{b})]^2$$

for certain complex numbers \mathbf{a} , \mathbf{b} , and \mathbf{k} . To guarantee that $|\mathbf{r}'(t)| \neq 0$ for all real t , the numbers \mathbf{a} and \mathbf{b} should have nonzero imaginary parts. Further, \mathbf{a} and \mathbf{b} should not be conjugates, since the argument of $\mathbf{r}'(t)$ would then be independent of t , corresponding to a degenerate linear locus.

Remark 5. The correspondence between expression (29) and equations (7) with $w(t) = 1$ amounts to identifying the polynomials $u(t)$ and $v(t)$ with the real and imaginary parts of $\sqrt{\mathbf{k}}(t - \mathbf{a})(t - \mathbf{b})$. In particular, the Bernstein coefficients u_0, u_1, u_2 and v_0, v_1, v_2 correspond to the real and imaginary parts of the complex numbers

$$\sqrt{\mathbf{k}}\mathbf{a}\mathbf{b}, \quad -\frac{1}{2}\sqrt{\mathbf{k}}[\mathbf{a}(1 - \mathbf{b}) + (1 - \mathbf{a})\mathbf{b}], \quad \sqrt{\mathbf{k}}(1 - \mathbf{a})(1 - \mathbf{b}).$$

Definition 4. Arbitrary complex Hermite data $\mathbf{r}(0)$, $\mathbf{r}'(0)$ and $\mathbf{r}(1)$, $\mathbf{r}'(1)$ is reduced to the *standard form*

$$\mathbf{r}(0) = 0, \quad \mathbf{r}'(0) = \mathbf{d}_0 \quad \text{and} \quad \mathbf{r}(1) = 1, \quad \mathbf{r}'(1) = \mathbf{d}_1$$

by (i) subtracting $\mathbf{r}(0)$ from the end points; and (ii) dividing the end points and end derivatives by $\mathbf{r}(1) - \mathbf{r}(0)$.

Thus, the shape of the Hermite interpolants depends only on the relative magnitudes and orientations of the end derivatives \mathbf{d}_0 and \mathbf{d}_1 . Since it is a trivial matter to transform to and from the original coordinate system by the appropriate complex arithmetic operations, we shall consider henceforth only the interpolation of Hermite data in standard form.

By differentiation, it is readily verified that the PH quintic corresponding to the hodograph (29) may be expressed as

$$(30) \quad \mathbf{r}(t) = \frac{\mathbf{k}}{30}[(t - \mathbf{a})^5 - 5(t - \mathbf{a})^4(t - \mathbf{b}) + 10(t - \mathbf{a})^3(t - \mathbf{b})^2] + \mathbf{c},$$

where, in order to satisfy $\mathbf{r}(0) = 0$, the constant of integration \mathbf{c} is given by

$$(31) \quad \mathbf{c} = \frac{\mathbf{k}}{30}(\mathbf{a}^5 - 5\mathbf{a}^4\mathbf{b} + 10\mathbf{a}^3\mathbf{b}^2).$$

Thus, in standard form the problem of Hermite interpolation by PH quintics amounts to computing complex constants \mathbf{a} , \mathbf{b} , and \mathbf{k} such that the curve defined by (30) and (31) satisfies $\mathbf{r}'(0) = \mathbf{d}_0$, $\mathbf{r}'(1) = \mathbf{d}_1$, and $\mathbf{r}(1) = 1$, and then deciding which of the resulting solutions gives the "best" curve.

Proposition 3. Let ρ be either of the two complex numbers defined by

$$(32) \quad \rho^2 = \frac{\mathbf{d}_0}{\mathbf{d}_1},$$

and hence let α be either of the two solutions to the quadratic equation

$$(33) \quad \alpha^2 - 3(1 + \rho)\alpha + 6\rho^2 + 2\rho + 6 - \frac{30}{\mathbf{d}_1} = 0.$$

Then if μ_1 and μ_2 are the two roots of

$$(34) \quad \mu^2 - \alpha\mu + \rho = 0,$$

the values of \mathbf{a} and \mathbf{b} in (30) are given by

$$(35) \quad \mathbf{a} = \frac{\mu_1}{\mu_1 + 1} \quad \text{and} \quad \mathbf{b} = \frac{\mu_2}{\mu_2 + 1}.$$

The corresponding value of \mathbf{k} is then

$$(36) \quad \mathbf{k} = \frac{\mathbf{d}_0}{\mathbf{a}^2 \mathbf{b}^2} = \frac{\mathbf{d}_1}{(1 - \mathbf{a})^2 (1 - \mathbf{b})^2},$$

and \mathbf{c} is given in terms of \mathbf{a} , \mathbf{b} , and \mathbf{k} by expression (31).

Proof. From (29) we see that

$$(37) \quad \mathbf{k} \mathbf{a}^2 \mathbf{b}^2 = \mathbf{d}_0 \quad \text{and} \quad \mathbf{k} (1 - \mathbf{a})^2 (1 - \mathbf{b})^2 = \mathbf{d}_1,$$

while the condition $\mathbf{r}(1) = 1$ may be cast in the form

$$(38) \quad \begin{aligned} & \mathbf{k} \{ [6(1 - \mathbf{a})^2 - 3(1 - \mathbf{a})\mathbf{a} + \mathbf{a}^2](1 - \mathbf{b})^2 \\ & + [-3(1 - \mathbf{a})^2 + 4(1 - \mathbf{a})\mathbf{a} - 3\mathbf{a}^2](1 - \mathbf{b})\mathbf{b} \\ & + [(1 - \mathbf{a})^2 - 3(1 - \mathbf{a})\mathbf{a} + 6\mathbf{a}^2]\mathbf{b}^2 \} = 30. \end{aligned}$$

We begin by using expressions (37) to eliminate \mathbf{b} and \mathbf{k} from equation (38); this is accomplished by substituting

$$\mathbf{b}^2 = \frac{\mathbf{d}_0}{\mathbf{k} \mathbf{a}^2}, \quad (1 - \mathbf{b})\mathbf{b} = \frac{\sqrt{\mathbf{d}_0 \mathbf{d}_1}}{\mathbf{k}(1 - \mathbf{a})\mathbf{a}}, \quad (1 - \mathbf{b})^2 = \frac{\mathbf{d}_1}{\mathbf{k}(1 - \mathbf{a})^2}.$$

The resulting equation depends only on \mathbf{d}_0 , \mathbf{d}_1 and the ratio $\mu = \mathbf{a}/(1 - \mathbf{a})$:

$$\mathbf{d}_1(\mu^2 - 3\mu + 6) - \sqrt{\mathbf{d}_0 \mathbf{d}_1}(3\mu - 4 + 3\mu^{-1}) + \mathbf{d}_0(6 - 3\mu^{-1} + \mu^{-2}) = 30.$$

With $\rho^2 = \mathbf{d}_0/\mathbf{d}_1$, the above may be further simplified to give the quadratic equation (33) by setting $\alpha = \mu + \rho\mu^{-1}$. Once a root α of (33) has been computed, the corresponding values of μ may be obtained by inverting the relation $\alpha = \mu + \rho\mu^{-1}$, i.e., by solving equation (34). Note here that, since equations (37) and (38) are symmetric in \mathbf{a} and \mathbf{b} , the two roots μ_1 and μ_2 of (34) must identify values of μ appropriate to *corresponding values* of \mathbf{a} and \mathbf{b} . The latter are obtained by inverting the relations $\mu_1 = \mathbf{a}/(1 - \mathbf{a})$ and $\mu_2 = \mathbf{b}/(1 - \mathbf{b})$, i.e., by expressions (35). Finally, knowing a pair of corresponding values for \mathbf{a} and \mathbf{b} , the appropriate value (36) for \mathbf{k} is obtained by substituting $t = 0$ or $t = 1$ into (29). \square

Not surprisingly, the procedure described in Proposition 3 generates four PH quintics, corresponding to distinct pairs (\mathbf{a}, \mathbf{b}) of complex numbers, for any given complex values \mathbf{d}_0 and \mathbf{d}_1 . (Four α -values may occur in equation (34), since these values are roots of the quadratic (33) in which the coefficient ρ assumes the two values defined by (32).) Compared to the real-arithmetical approach in Proposition 2, the complex formulation yields a concise solution. Moreover, we will show that the location of \mathbf{a} and \mathbf{b} relative to the interval $[0, 1]$ gives insight into the shape properties of the four possible interpolants, and offers a fairly simple basis for identifying the "good" solution.

Remark 6. Expressions (29) and (30) can be cast in Bernstein-Bézier form to give complex control points for the hodograph and curve. Setting $t - \mathbf{a} =$

$-\mathbf{a}(1-t) + (1-a)t$ in (29), and likewise for $t - \mathbf{b}$, and expanding gives

$$\begin{aligned} \mathbf{h}_0 &= \mathbf{k}a^2b^2, \\ \mathbf{h}_1 &= -\frac{1}{2}\mathbf{k}ab(\mathbf{a} + \mathbf{b} - 2ab), \\ \mathbf{h}_2 &= \frac{1}{6}\mathbf{k}[(\mathbf{a} + \mathbf{b} - 2ab)^2 + 2\mathbf{a}(1-a)\mathbf{b}(1-b)], \\ \mathbf{h}_3 &= -\frac{1}{2}\mathbf{k}(1-a)(1-b)(\mathbf{a} + \mathbf{b} - 2ab), \\ \mathbf{h}_4 &= \mathbf{k}(1-a)^2(1-b)^2, \end{aligned}$$

for the hodograph control points, and in terms of the above the control points of the PH quintic are [9]:

$$\mathbf{p}_k = \frac{1}{5} \sum_{j=0}^{k-1} \mathbf{h}_j \quad \text{for } k = 1, \dots, 5.$$

Here we take $\mathbf{p}_0 = 0$ as the constant of integration; note also that $\mathbf{p}_5 = 1$ when \mathbf{a} , \mathbf{b} , and \mathbf{k} are computed according to Proposition 3.

A somewhat simpler construction of the Hermite interpolants results if we commence with the form

$$\mathbf{r}'(t) = [\delta_0(1-t)^2 + \beta(1-t)t + \delta_1t^2]^2$$

in lieu of expression (29). Here, δ_0 and δ_1 are square roots of \mathbf{d}_0 and \mathbf{d}_1 , while β is a root of the equation

$$\beta^2 + 3\gamma\beta + \gamma^2 + 5(\mathbf{d}_0 + \mathbf{d}_1) - 30 = 0,$$

where $\gamma = \delta_0 + \delta_1$. Although simpler, this approach does not directly yield information on the location of \mathbf{a} and \mathbf{b} in the complex plane that we need below to differentiate among the solutions. (It is, however, better suited to formulating a priori constraints on the Hermite data guaranteeing that the "good" interpolant has certain desirable properties—this is a substantive problem in its own right, which we hope to address in a future study.)

6. THE ABSOLUTE ROTATION NUMBER

We now consider the identification of the "best" of the four interpolants that Proposition 3 yields for given Hermite data. For each solution $(\mathbf{a}, \mathbf{b}, \mathbf{k})$ the hodograph (29) indicates that the curve tangent at point t makes an angle

$$(39) \quad \theta(t) = \arg(\mathbf{k}) + 2\arg(\mathbf{a} - t) + 2\arg(\mathbf{b} - t)$$

with the positive real axis (where $-\pi < \arg(\mathbf{z}) \leq \pi$). We will differentiate among the four solutions according to the variation of $\theta(t)$ over $t \in [0, 1]$.

As noted in §4, the four PH quintics must be judged according to criteria that do not depend upon comparison with the variation of $\theta(t)$ for the *unique* cubic interpolant. Furthermore, it is clear from examples such as those shown in Figures 1 and 3 that minimizing the *net* variation of $\theta(t)$, as expressed by the rotation number (23), is not a satisfactory criterion—cases in which large amounts of positive and negative rotation nearly cancel each other may be as undesirable as those exhibiting too much rotation of a fixed sense.

Consequently, we are motivated to introduce the *absolute rotation number*

$$(40) \quad \mathcal{R}_{\text{abs}} = \frac{1}{2\pi} \int_0^1 |\kappa(t)| |\mathbf{r}'(t)| dt$$

as the quantity whose minimization identifies the “best” interpolant. \mathcal{R}_{abs} measures the *total* amount—i.e., regardless of sense—of “winding” of the curve tangent. If κ is of constant sign over $t \in [0, 1]$, then \mathcal{R}_{abs} is just the absolute value of the ordinary rotation number (23). If the curve has *inflections*, however, the integral (40) can only be evaluated by breaking it up at each of the values $t_1, \dots, t_N \in (0, 1)$ where κ changes sign.

Instead of using the methods of §4, we appeal directly to expression (39) for the tangent angle to compute \mathcal{R}_{abs} . Since (40) does not depend on the reference direction from which $\theta(t)$ is measured, we may drop the constant $\arg(\mathbf{k})$ and write $\theta(t) = 2[\arg(\mathbf{a} - t) + \arg(\mathbf{b} - t)]$. Note that for a regular curve, satisfying $\mathbf{r}'(t) \neq 0$ for all real t , \mathbf{a} and \mathbf{b} must have nonzero imaginary parts (and are not conjugates if the Hermite data is nondegenerate).

Proposition 4. *Let $\angle uzv$ be the angle subtended at vertex \mathbf{z} of a triangle in the complex plane whose other vertices are at values u and v on the real axis. Then for the PH quintics constructed according to Proposition 3 we have*

$$\mathcal{R}_{\text{abs}} = \frac{1}{\pi}(\angle 0\mathbf{a}1 + \angle 0\mathbf{b}1),$$

when \mathbf{a} and \mathbf{b} lie on the same side of the real axis, and

$$\mathcal{R}_{\text{abs}} = \frac{1}{\pi} \sum_{k=0}^N |\angle t_k \mathbf{a} t_{k+1} - \angle t_k \mathbf{b} t_{k+1}|$$

where $t_0 = 0$, $t_{N+1} = 1$, and t_1, \dots, t_N ($N \leq 2$) are the ordered roots of

$$(41) \quad \text{Im}(\mathbf{a} + \mathbf{b})t^2 - 2\text{Im}(\mathbf{a}\mathbf{b})t + \text{Im}(|\mathbf{a}|^2\mathbf{b} + |\mathbf{b}|^2\mathbf{a}) = 0$$

on $t \in (0, 1)$ when \mathbf{a} and \mathbf{b} lie on opposite sides of the real axis.

Proof. If \mathbf{a} and \mathbf{b} lie on one side of the real axis, $\arg(\mathbf{a} - t)$ and $\arg(\mathbf{b} - t)$ are either both positive and monotone increasing with t , or both negative and monotone decreasing with t . Thus, κ cannot vanish, and we may write

$$\begin{aligned} \mathcal{R}_{\text{abs}} &= \frac{1}{2\pi} [2|\arg(\mathbf{a} - 1) - \arg(\mathbf{a})| + 2|\arg(\mathbf{b} - 1) - \arg(\mathbf{b})|] \\ &= \frac{1}{\pi} (\angle 0\mathbf{a}1 + \angle 0\mathbf{b}1). \end{aligned}$$

If \mathbf{a} and \mathbf{b} lie on opposite sides of the real axis, there may be zero, one, or two inflections on $t \in (0, 1)$. In the complex formulation, the curvature can be expressed as $\kappa(t) = |\mathbf{r}'(t)|^{-3} \text{Im}(\overline{\mathbf{r}'(t)}\mathbf{r}''(t))$, and by substituting from (29) it may be verified that the numerator of this expression is proportional to the quadratic (41), whose simple roots identify inflections. Geometrically, the parameter values t corresponding to inflections are points on the real axis from which the ratio of the distances to \mathbf{a} and \mathbf{b} has the fixed value

$$|\mathbf{a} - t|/|\mathbf{b} - t| = \sqrt{-\text{Im}(\mathbf{a})/\text{Im}(\mathbf{b})}.$$

(Note that the discriminant $\Delta = -4\text{Im}(\mathbf{a})\text{Im}(\mathbf{b})|\overline{\mathbf{a}} - \mathbf{b}|^2$ of (41) is positive in this case, so the number of inflections on $t \in (0, 1)$ equals the number of sign variations in its Bernstein coefficients.) If inflections at t_1, \dots, t_N are identified, and we set $t_0 = 0$ and $t_{N+1} = 1$, then

$$\mathcal{R}_{\text{abs}} = \frac{1}{2\pi} \sum_{k=0}^N 2|\arg(\mathbf{a} - t_{k+1}) - \arg(\mathbf{a} - t_k) + \arg(\mathbf{b} - t_{k+1}) - \arg(\mathbf{b} - t_k)|,$$

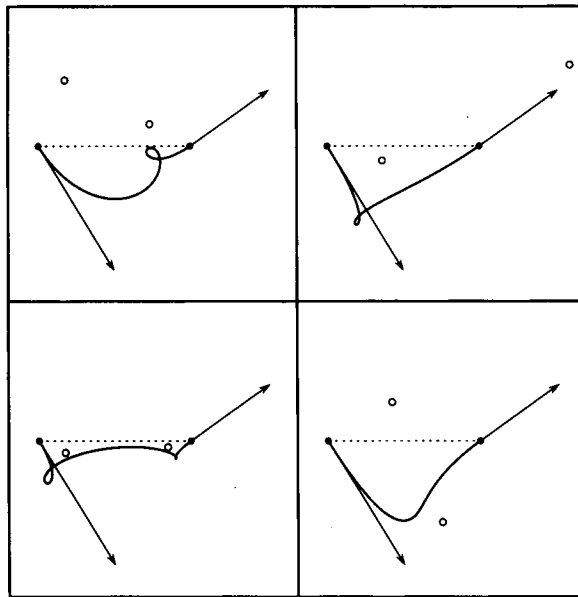


FIGURE 4. The four PH quintic interpolants to prescribed Hermite data on the unit interval (the dashed line). The open dots indicate the locations of the roots \mathbf{a} and \mathbf{b} of the complex hodograph given by expression (29)

and since $\arg(\mathbf{a} - t_{k+1}) - \arg(\mathbf{a} - t_k)$ and $\arg(\mathbf{b} - t_{k+1}) - \arg(\mathbf{b} - t_k)$ must be of opposite sign, and equal in absolute value to $\angle t_k \mathbf{a} t_{k+1}$ and $\angle t_k \mathbf{b} t_{k+1}$, we have

$$\mathcal{R}_{\text{abs}} = \frac{1}{\pi} \sum_{k=0}^N |\angle t_k \mathbf{a} t_{k+1} - \angle t_k \mathbf{b} t_{k+1}|. \quad \square$$

Note that $\angle 0 \mathbf{a} 1$ and $\angle 0 \mathbf{b} 1$ correspond to $|\pi - \arg(\mu_1)|$ and $|\pi - \arg(\mu_2)|$, μ_1 and μ_2 being the roots of (34). Typically, configurations where \mathbf{a} and \mathbf{b} lie on opposite sides of the real axis and are not close to the interval $t \in (0, 1)$ give the smallest \mathcal{R}_{abs} values; see the example shown in Figure 4. The upper bound $\mathcal{R}_{\text{abs}} = 2$ is approached when either: (i) \mathbf{a} and \mathbf{b} are near to and on the same side of $t \in (0, 1)$; or (ii) \mathbf{a} and \mathbf{b} are near to but on opposite sides of *distinct* subintervals $t \in (t_k, t_{k+1})$ delineated by the inflections.

7. CONCLUDING REMARKS

By means of the PH quintic Hermite interpolation scheme, smooth (C^1) piecewise PH curves may be constructed that interpolate sequences of point/tangent data. Such curves then have exactly computable arc lengths and offset curves; the details of these computations may be found in earlier papers [5, 10].

If only points are prescribed, one way of associating a “tangent vector” (derivative) with each is by solving the tridiagonal linear system that arises in imposing curvature continuity at each node of an ordinary C^2 cubic spline [2]. Figure 5 compares the piecewise PH quintic interpolating data for a model airfoil constructed thus (the “good” interpolants between consecutive points being

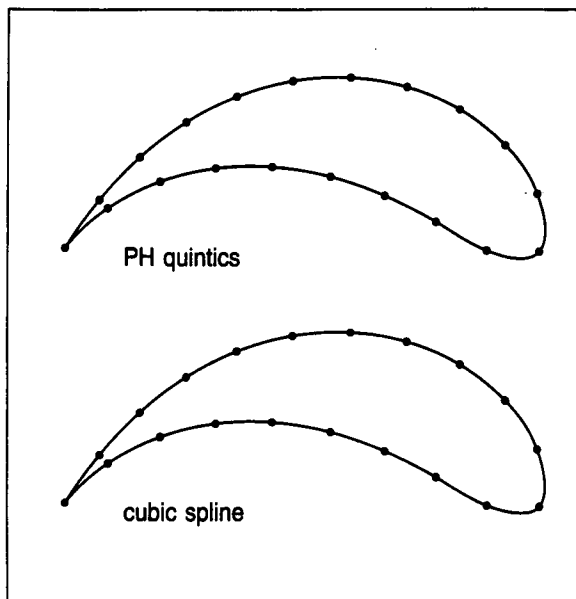


FIGURE 5. Comparison of piecewise PH quintic (upper) and cubic spline (lower) fits to point data for an airfoil, with specified initial/final derivatives at the trailing edge (the tangent vectors for the piecewise PH curve are taken from the cubic spline). The two curves are essentially indistinguishable, but the piecewise PH interpolant allows an exact representation of its offsets

identified as described in §6) with the corresponding cubic spline. The curves are virtually indistinguishable, as is generally the case when the given data does not suggest inflections or strong curvature variations.

Other methods for selecting tangent vectors, such as those used [12, 13] in “shape-preserving” interpolation schemes, are available. Note that if precise C^2 continuity is desired of piecewise PH interpolants, one must contend with systems of nonlinear equations [5], and the resulting curves would not be true “splines” in the sense of minimizing an (approximate) “energy” integral [2]. Nevertheless, for many practical curve design/approximation problems, exact continuity of the curvature is not essential or necessarily even desirable: while κ is strictly continuous for an ordinary C^2 cubic spline, this does not guarantee a *fair* curve, in the sense that κ may exhibit a “bumpy” global behavior. The example shown in Figure 6 (next page), for which we use the “not-a-knot” condition [2] instead of specifying end derivatives, illustrates this point.

As noted in §1, a primary use envisaged for the Hermite interpolation scheme is one-for-one replacement of the *de facto* “standard” Bézier cubics by PH quintics having similar “shape” properties. This allows one to take advantage of the exact rectification and offset properties of PH curves in applications where such requirements are of paramount importance. In fact, empirical evidence suggests that the “good” PH quintics have better shape (i.e., a more even curvature distribution) than their cubic counterparts.

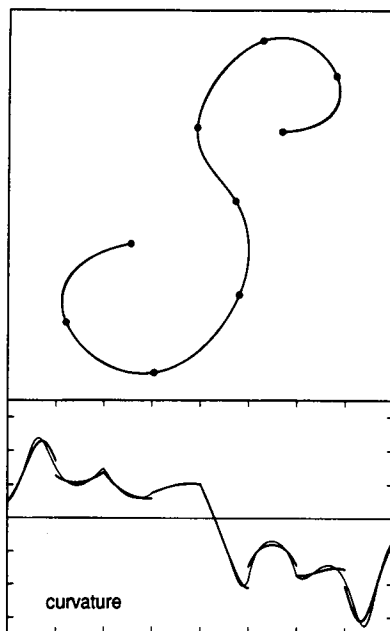


FIGURE 6. Comparison of piecewise PH quintic (bold curve) and C^2 cubic spline (light curve) fits to point data, and the curvature distribution along them—the two curves are superposed and thus virtually indistinguishable. While the curvature of the piecewise PH curve is not precisely continuous, the small jumps in κ at the nodes are perhaps insignificant in view of the overall “bumpy” variation of κ

BIBLIOGRAPHY

1. P. Bézier, *The mathematical basis of the UNISURF CAD system*, Butterworths, London, 1986.
2. C. de Boor, *A practical guide to splines*, Springer, New York, 1978.
3. M. P. do Carmo, *Differential geometry of curves and surfaces*, Prentice-Hall, Englewood Cliffs, NJ, 1976.
4. G. Farin, *Curves and surfaces for computer aided geometric design*, 3rd ed., Academic Press, Boston, 1993.
5. R. T. Farouki, *Pythagorean-hodograph curves in practical use*, Geometry Processing for Design and Manufacturing (R. E. Barnhill, ed.), SIAM, Philadelphia, PA, 1992, pp. 3–33.
6. ———, *The conformal map $z \rightarrow z^2$ of the hodograph plane*, Comput. Aided Geom. Design **11** (1994), 363–390.
7. R. T. Farouki and C. A. Neff, *Analytic properties of plane offset curves*, Comput. Aided Geom. Design **7** (1990), 83–99; *Algebraic properties of plane offset curves*, *ibid.*, 100–127.
8. R. T. Farouki and V. T. Rajan, *On the numerical condition of polynomials in Bernstein form*, Comput. Aided Geom. Design **4** (1987), 191–216.
9. ———, *Algorithms for polynomials in Bernstein form*, Comput. Aided Geom. Design **5** (1988), 1–26.
10. R. T. Farouki and T. Sakkalis, *Pythagorean hodographs*, IBM J. Res. Develop. **34** (1990), 736–752.
11. F. R. Gantmacher, *The theory of matrices*, Vol. 2, Chelsea, New York, 1959, pp. 174–176.

12. T. N. T. Goodman and K. Unsworth, *Shape preserving interpolation by parametrically defined curves*, SIAM J. Numer. Anal. **25** (1988), 1453–1465.
13. ———, *Shape preserving interpolation by curvature continuous parametric curves*, Comput. Aided Geom. Design **5** (1988), 323–340.
14. P. Henrici, *Applied and computational complex analysis*, Vol. 1, Wiley, New York, 1974, pp. 443–444.
15. E. Kreyszig, *Differential geometry*, University of Toronto Press, Toronto, 1959.
16. K. K. Kubota, *Pythagorean triples in unique factorization domains*, Amer. Math. Monthly **79** (1972), 503–505.
17. R. S. Millman and G. D. Parker, *Elements of differential geometry*, Prentice-Hall, Englewood Cliffs, NJ, 1977.
18. B. Pham, *Offset curves and surfaces: a brief survey*, Comput. Aided Design **24** (1992), 223–229.
19. I. J. Schoenberg, *On variation-diminishing approximation methods*, On Numerical Approximation (R. E. Langer, ed.), Univ. of Wisconsin Press, Madison, WI, 1959, pp. 249–274.
20. H. Whitney, *On regular closed curves in the plane*, Compositio Math. **4** (1937), 276–284.

DEPARTMENT OF MECHANICAL ENGINEERING AND APPLIED MECHANICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MICHIGAN 48109
E-mail address: farouki@engin.umich.edu

IBM THOMAS J. WATSON RESEARCH CENTER, P.O. BOX 218, YORKTOWN HEIGHTS, NEW YORK 10598
E-mail address: caneff@watson.ibm.com