

Construction of C^2 Pythagorean–hodograph interpolating splines by the homotopy method

Gudrun Albrecht

Mathematisches Institut, Technische Universität München, D–80290 München, Germany
E-mail: albrecht@mathematik.tu-muenchen.de

Rida T. Farouki

Department of Mechanical Engineering and Applied Mechanics, University of Michigan, Ann Arbor, MI 48109, USA
E-mail: farouki@engin.umich.edu

Received 6 February 1995; revised 21 September 1995

Communicated by T. Lyche

The complex representation of polynomial Pythagorean–hodograph (PH) curves allows the problem of constructing a C^2 PH quintic “spline” that interpolates a given sequence of points $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_N$ and end–derivatives \mathbf{d}_0 and \mathbf{d}_N to be reduced to solving a “tridiagonal” system of N quadratic equations in N complex unknowns. The system can also be easily modified to incorporate PH–spline *end conditions* that bypass the need to specify end–derivatives. Homotopy methods have been employed to compute all solutions of this system, and hence to construct a total of 2^{N+1} distinct interpolants for each of several different data sets. We observe empirically that all but one of these interpolants exhibits undesirable “looping” behavior (which may be quantified in terms of the *elastic bending energy*, i.e., the integral of the square of the curvature with respect to arc length). The remaining “good” interpolant, however, is invariably a *fairer* curve—having a smaller energy and a more even curvature distribution over its extent—than the corresponding “ordinary” C^2 cubic spline. Moreover, the PH spline has the advantage that its offsets are *rational* curves and its arc length is a *polynomial* function of the curve parameter.

1. Introduction

Pythagorean–hodograph (PH) curves are a novel class of parametric curves, suitable for use in computer–aided design (CAD) applications, that provide significant computational advantages over polynomial curves in general [9]. The properties that distinguish PH curves from general polynomial curves may be summarized as follows:

- the arc length of a PH curve can be computed *precisely*—i.e., without numerical quadrature—by simply evaluating a polynomial [5];

- the offsets to any PH curve (the loci of a point that maintains successive fixed distances from it) are *rational* curves, and can thus be represented without approximation in existing CAD systems [5];
- the “bending energy” of PH curves—i.e., the integral of the square of the curvature with respect to arc length—admits a closed-form expression in terms of rational functions, arctangents, and natural logarithms that is amenable to optimization with respect to free parameters [7];
- in the interpolation of discrete point/tangent data, PH curves typically yield *fairer* loci—exhibiting smoother curvature profiles—than the “ordinary” polynomial curves [7].

Further details on these attributes may be found in the cited references.

A crucial requirement in establishing the PH curves as a practical design medium is the provision of algorithms that offer intuitive geometric means to construct and manipulate these curves, while shielding the user from the formidable algebra that such functions nominally incur. Steps toward fulfillment of this goal—in terms of C^1 point/tangent data interpolation—were reported in [5,7,8] but thus far they fall short of the commonly-accepted minimum level of flexibility for any free-form design scheme, namely, C^2 interpolation of an ordered sequence $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_N$ of points.

Actually, systems of equations defining C^2 piecewise-PH quintic “splines” that interpolate points $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_N$ were already formulated in [5], but as a consequence of the *real* representation for PH quintics adopted therein, these equations transpired to be unduly cumbersome and were abandoned. The subsequent introduction of *complex* representations, and the dramatic simplifications in constructing and analyzing PH curves that ensue, prompts us to re-examine the C^2 interpolation problem from this perspective.

The organization of this paper is as follows. In section 2 we develop the complex-form C^2 continuity condition for two PH quintic arcs that share an endpoint. We further show that, when interpolating a sequence of $N + 1$ points together with derivatives at the first and last point, the satisfaction of C^2 continuity at each node is achieved upon solving a “tridiagonal” system of N quadratic equations in N complex variables (only three consecutive unknowns appear in each equation). Since the system apparently has no simple closed-form solution, and we wish to investigate the full spectrum of C^2 PH splines, we formulate homotopy methods in section 3 that yield all solutions.

The system may be modified, as described in section 4, to incorporate PH spline *end conditions* that circumvent the need to specify end-derivatives. As observed in the examples of section 5, the governing equations for PH splines are exceptionally well-conditioned: in double-precision arithmetic one typically obtains 14 or 15 accurate decimal digits. Among the 2^{N+1} distinct PH splines interpolating given data, we observe a unique “good” solution that exhibits a smaller bending energy—and a correspondingly *fairer* visual appearance—than the “ordinary” cubic spline. Finally, section 6 summarizes our results and identifies issues that are worthy of further investigation.

2. Complex-form C^2 continuity conditions

We shall find it convenient [6] to express plane polynomial curves as complex-valued polynomials $\mathbf{r}(t) = x(t) + iy(t)$ of a real parameter t . If the derivative or *hodograph* of such a curve is the *perfect square* of some complex polynomial $\mathbf{w}(t) = \pm [u(t) + iv(t)]$, i.e.,

$$\mathbf{r}'(t) = x'(t) + iy'(t) = \mathbf{w}^2(t) = u^2(t) - v^2(t) + i2u(t)v(t), \tag{1}$$

we say that $\mathbf{r}(t)$ is a *Pythagorean-hodograph curve* (in particular, it is a *regular PH curve* if the polynomial $\mathbf{w}(t)$ has relatively prime real and imaginary parts: $\text{gcd}(u, v) = 1$). Evidently the Cartesian components $x'(t)$ and $y'(t)$ of the hodograph $\mathbf{r}'(t)$ then satisfy the Pythagorean condition

$$x'^2(t) + y'^2(t) \equiv \sigma^2(t), \tag{2}$$

where $\sigma(t)$ denotes the polynomial $u^2(t) + v^2(t)$. Equation (2) gives rise to the attractive properties of PH curves enumerated in section 1.

The case where we take $\mathbf{w}(t)$ to be the quadratic

$$\mathbf{w}(t) = \pm [\mathbf{w}_0(1-t)^2 + \mathbf{w}_1 2(1-t)t + \mathbf{w}_2 t^2] \tag{3}$$

with Bernstein coefficients $\mathbf{w}_0, \mathbf{w}_1, \mathbf{w}_2$ yields a PH quintic

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

upon substituting into (1) and integrating. The complex Bézier control points $\mathbf{z}_k = x_k + iy_k$ may be expressed [6] in terms of these coefficients as

$$\begin{aligned} \mathbf{z}_1 &= \mathbf{z}_0 + \frac{1}{5} \mathbf{w}_0^2, \\ \mathbf{z}_2 &= \mathbf{z}_1 + \frac{1}{5} \mathbf{w}_0 \mathbf{w}_1, \\ \mathbf{z}_3 &= \mathbf{z}_2 + \frac{2 \mathbf{w}_1^2 + \mathbf{w}_0 \mathbf{w}_2}{15}, \\ \mathbf{z}_4 &= \mathbf{z}_3 + \frac{1}{5} \mathbf{w}_1 \mathbf{w}_2, \\ \mathbf{z}_5 &= \mathbf{z}_4 + \frac{1}{5} \mathbf{w}_2^2, \end{aligned} \tag{4}$$

where the constant of integration \mathbf{z}_0 is arbitrary. Note that the position of the final control point relative to the initial one is given by

$$\mathbf{z}_5 - \mathbf{z}_0 = \frac{1}{5} \mathbf{w}_0^2 + \frac{1}{5} \mathbf{w}_0 \mathbf{w}_1 + \frac{2 \mathbf{w}_1^2 + \mathbf{w}_0 \mathbf{w}_2}{15} + \frac{1}{5} \mathbf{w}_1 \mathbf{w}_2 + \frac{1}{5} \mathbf{w}_2^2. \tag{5}$$

The problem of Hermite interpolation to given end-point/derivative data by *individual* PH quintic arcs was thoroughly investigated in [8]. We are concerned

here with how a sequence of such interpolants may be pieced together to yield a curve that passes through given points with *second-order continuity*. Previous attempts [5] to solve this problem proved unsuccessful; we will now show that the complex formulation offers a simple and tractable approach.

Suppose $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_N$ is a sequence of $N + 1$ points (complex values) that we wish to interpolate by a C^2 piecewise-PH-quintic curve. Let $\mathbf{r}_k(t)$ for $t \in [0, 1]$ be the constituent arc of this curve between consecutive points \mathbf{p}_{k-1} and \mathbf{p}_k . We denote the Bézier control points of this arc by

$$\mathbf{z}_{k,0}, \mathbf{z}_{k,1}, \mathbf{z}_{k,2}, \mathbf{z}_{k,3}, \mathbf{z}_{k,4}, \mathbf{z}_{k,5}, \tag{6}$$

and if it is to have a Pythagorean hodograph, these points must be expressible in the form (4) in terms of three complex values

$$(-1)^{\epsilon_k} \mathbf{w}_{k,0}, (-1)^{\epsilon_k} \mathbf{w}_{k,1}, (-1)^{\epsilon_k} \mathbf{w}_{k,2} \tag{7}$$

for $k = 1, \dots, N$. To interpolate the points \mathbf{p}_{k-1} and \mathbf{p}_k , we clearly must have

$$\mathbf{z}_{k,0} = \mathbf{p}_{k-1} \quad \text{and} \quad \mathbf{z}_{k,5} = \mathbf{p}_k. \tag{8}$$

2.1. Continuity equations

Consider now the imposition of C^1 continuity at the juncture of arcs $\mathbf{r}_k(t)$ and $\mathbf{r}_{k+1}(t)$. We require that

$$\mathbf{r}'_k(1) = \mathbf{r}'_{k+1}(0)$$

for $k = 1, \dots, N - 1$. Since $\mathbf{r}'_k(t) = \mathbf{w}_k^2(t)$ and $\mathbf{r}'_{k+1}(t) = \mathbf{w}_{k+1}^2(t)$, where $\mathbf{w}_k(t)$ and $\mathbf{w}_{k+1}(t)$ are quadratics with Bernstein coefficients

$$(-1)^{\epsilon_k} \mathbf{w}_{k,0}, (-1)^{\epsilon_k} \mathbf{w}_{k,1}, (-1)^{\epsilon_k} \mathbf{w}_{k,2},$$

and

$$(-1)^{\epsilon_{k+1}} \mathbf{w}_{k+1,0}, (-1)^{\epsilon_{k+1}} \mathbf{w}_{k+1,1}, (-1)^{\epsilon_{k+1}} \mathbf{w}_{k+1,2},$$

we can achieve C^1 continuity for $k = 1, \dots, N - 1$ by taking

$$\mathbf{w}_{k+1,0} = (-1)^{i_k} \mathbf{w}_{k,2}, \quad \text{where} \quad i_k \in \{0, 1\}. \tag{9}$$

Consider next the imposition of C^2 continuity. The first derivatives of adjacent arcs k and $k + 1$ are given by

$$\begin{aligned} \mathbf{r}'_k(t) &= [\mathbf{w}_{k,0} (1 - t)^2 + \mathbf{w}_{k,1} 2(1 - t)t + \mathbf{w}_{k,2} t^2]^2, \\ \mathbf{r}'_{k+1}(t) &= [\mathbf{w}_{k+1,0} (1 - t)^2 + \mathbf{w}_{k+1,1} 2(1 - t)t + \mathbf{w}_{k+1,2} t^2]^2. \end{aligned} \tag{10}$$

By differentiating the above, it can be seen that the C^2 conditions

$$\mathbf{r}''_k(1) = \mathbf{r}''_{k+1}(0)$$

for $k = 1, \dots, N - 1$ are equivalent to

$$2 \mathbf{w}_{k,2} = \mathbf{w}_{k,1} + (-1)^{i_k} \mathbf{w}_{k+1,1}. \tag{11}$$

Note that the conditions (9) and (11), with $i_k = (\epsilon_k + \epsilon_{k+1}) \bmod 2$, yield C^0 and C^1 continuity, respectively, of the polynomial $\mathbf{w}_k(t)$ at $t = 1$ with the polynomial $\mathbf{w}_{k+1}(t)$ at $t = 0$. These conditions thus guarantee C^1 and C^2 continuity of the segments $\mathbf{r}_k(t)$ and $\mathbf{r}_{k+1}(t)$ at their juncture.

We now define three sets of N complex values $\mathbf{c}_0, \dots, \mathbf{c}_{N-1}$ and $\mathbf{b}_1, \dots, \mathbf{b}_N$ and $\mathbf{a}_1, \dots, \mathbf{a}_N$ by setting

$$\mathbf{c}_{k-1} := \mathbf{w}_{k,0}, \quad \mathbf{b}_k := \mathbf{w}_{k,1}, \quad \mathbf{a}_k := \mathbf{w}_{k,2} \tag{12}$$

for $k = 1, \dots, N$. Consequently, the conditions (9) and (11) read

$$\mathbf{c}_k = (-1)^{i_k} \mathbf{a}_k \quad \text{and} \quad 2 \mathbf{a}_k = \mathbf{b}_k + (-1)^{i_k} \mathbf{b}_{k+1} \tag{13}$$

where $i_k \in \{0, 1\}$ for $k = 1, \dots, N - 1$. The above amount to $2N - 2$ equations for $3N$ unknowns, $\mathbf{c}_0, \dots, \mathbf{c}_{N-1}$ and $\mathbf{b}_1, \dots, \mathbf{b}_N$ and $\mathbf{a}_1, \dots, \mathbf{a}_N$. Now a further N equations arise from the condition that, for each arc $\mathbf{r}_k(t)$, the chosen values (12) and the endpoints (8) must be consistent with (5). This gives

$$3 \mathbf{c}_{k-1}^2 + \mathbf{c}_{k-1} \mathbf{a}_k + 3 \mathbf{a}_k^2 + 3 (\mathbf{c}_{k-1} + \mathbf{a}_k) \mathbf{b}_k + 2 \mathbf{b}_k^2 = 15 (\mathbf{p}_k - \mathbf{p}_{k-1}) \tag{14}$$

for $k = 1, \dots, N$. We thus have a total of $3N - 2$ equations, namely (13) and (14), for the $3N$ unknowns $\mathbf{c}_0, \dots, \mathbf{c}_{N-1}$ and $\mathbf{b}_1, \dots, \mathbf{b}_N$ and $\mathbf{a}_1, \dots, \mathbf{a}_N$.

As in the case of ‘‘ordinary’’ splines, we need additional *end conditions* in order to close the system. We defer the detailed treatment of end conditions to section 4, and at present consider only the simplest approach, namely, specifying values \mathbf{d}_0 and \mathbf{d}_N for the end derivatives:

$$\mathbf{c}_0^2 = \mathbf{r}'_0(0) = \mathbf{d}_0 \quad \text{and} \quad \mathbf{a}_N^2 = \mathbf{r}'_N(1) = \mathbf{d}_N, \tag{15}$$

i.e., we assign complex values *ab initio* to \mathbf{c}_0 and \mathbf{a}_N .

Equations (13) and (14) are equivalent to a far more cumbersome system—previously quoted as equations (38) and (39), respectively, in [5]—that arises from the *real* formulation of PH quintics. Note, in particular, that the linearity of equations (13) is specific to the complex formulation.

Suppose that \mathbf{c}_0 and \mathbf{a}_N are prescribed. Then from equations (13) we can substitute $\mathbf{c}_{k-1} = \frac{1}{2} [(-1)^{i_{k-1}} \mathbf{b}_{k-1} + \mathbf{b}_k]$ and $\mathbf{a}_k = \frac{1}{2} [\mathbf{b}_k + (-1)^{i_k} \mathbf{b}_{k+1}]$ into equations (14) in order to express the latter in terms of $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_N$ only. This gives the system of quadratic equations

$$3 \mathbf{b}_{k-1}^2 + 27 \mathbf{b}_k^2 + 3 \mathbf{b}_{k+1}^2 + 13 (-1)^{i_{k-1}} \mathbf{b}_{k-1} \mathbf{b}_k + 13 (-1)^{i_k} \mathbf{b}_k \mathbf{b}_{k+1} + (-1)^{i_{k-1} + i_k} \mathbf{b}_{k-1} \mathbf{b}_{k+1} = 60 (\mathbf{p}_k - \mathbf{p}_{k-1}) \tag{16}$$

for $k = 2, \dots, N - 1$, while for the cases $k = 1$ and $k = N$ we have

$$17 \mathbf{b}_1^2 + 3 \mathbf{b}_2^2 + 12 (-1)^{i_1} \mathbf{b}_1 \mathbf{b}_2 + 14 \mathbf{c}_0 \mathbf{b}_1 + 2 (-1)^{i_1} \mathbf{c}_0 \mathbf{b}_2 + 12 \mathbf{c}_0^2 = 60 (\mathbf{p}_1 - \mathbf{p}_0), \tag{17}$$

$$17 \mathbf{b}_N^2 + 3 \mathbf{b}_{N-1}^2 + 12 (-1)^{i_{N-1}} \mathbf{b}_N \mathbf{b}_{N-1} + 14 \mathbf{a}_N \mathbf{b}_N + 2 (-1)^{i_{N-1}} \mathbf{a}_N \mathbf{b}_{N-1} + 12 \mathbf{a}_N^2 = 60 (\mathbf{p}_N - \mathbf{p}_{N-1}) \tag{18}$$

where $i_k \in \{0, 1\}$ for $k = 1, \dots, N - 1$.

In section 2.2 and section 2.3 we will show that, without loss of generality, one may set $i_k = 0$ for $k = 1, \dots, N - 1$ in the system of equations (16)–(18), and that this system has only *finite* solutions. The reader interested in practical implementation details may proceed directly to section 3.

2.2. Equivalence of the 2^{N-1} equation systems

For each choice of $(i_1, \dots, i_{N-1}) \in \{0, 1\}^{N-1}$ in (16)–(18), we have a system of N simultaneous quadratic equations for the N quantities $\mathbf{b}_1, \dots, \mathbf{b}_N$, with at most 3 consecutive unknowns appearing in each equation. If these equations can be solved for $\mathbf{b}_1, \dots, \mathbf{b}_N$, the values of $\mathbf{c}_1, \dots, \mathbf{c}_{N-1}$ and $\mathbf{a}_1, \dots, \mathbf{a}_{N-1}$ follow directly from equations (13).

There are altogether 2^{N-1} such systems, and as they are all of similar structure—differing only in the signs of certain coefficients—the question naturally arises as to whether their solutions might be closely related. We will show that, in fact, the solutions to *any* of these systems yield a *unique* set of 2^{N+1} distinct PH splines, and it will therefore suffice to consider just *one* of the choices $(i_1, \dots, i_{N-1}) \in \{0, 1\}^{N-1}$.

By making the substitution¹

$$\mathbf{b}_k \rightarrow \prod_{j=k}^{N-1} (-1)^{i_j} \mathbf{b}_k \quad \text{for} \quad k = 1, \dots, N \tag{19}$$

in equations (16)–(18) and allowing for the sign ambiguity $\mathbf{c}_0 = \pm \sqrt{\mathbf{d}_0}$ in the end-conditions (15), one obtains the system

$$17 \mathbf{b}_1^2 + 3 \mathbf{b}_2^2 + 12 \mathbf{b}_1 \mathbf{b}_2 + 14 \mathbf{c}_0 \mathbf{b}_1 + 2 \mathbf{c}_0 \mathbf{b}_2 + 12 \mathbf{c}_0^2 = 60 (\mathbf{p}_1 - \mathbf{p}_0), \tag{20}$$

$$3 \mathbf{b}_{k-1}^2 + 27 \mathbf{b}_k^2 + 3 \mathbf{b}_{k+1}^2 + 13 \mathbf{b}_{k-1} \mathbf{b}_k + 13 \mathbf{b}_k \mathbf{b}_{k+1} + \mathbf{b}_{k-1} \mathbf{b}_{k+1} = 60 (\mathbf{p}_k - \mathbf{p}_{k-1}) \quad \text{for} \quad k = 2, \dots, N - 1, \tag{21}$$

$$17 \mathbf{b}_N^2 + 3 \mathbf{b}_{N-1}^2 + 12 \mathbf{b}_N \mathbf{b}_{N-1} + 14 \mathbf{a}_N \mathbf{b}_N + 2 \mathbf{a}_N \mathbf{b}_{N-1} + 12 \mathbf{a}_N^2 = 60 (\mathbf{p}_N - \mathbf{p}_{N-1}), \tag{22}$$

¹ For $k = N$, this becomes $\mathbf{b}_N \rightarrow \mathbf{b}_N$.

which evidently corresponds to the special choice $(i_1, \dots, i_{N-1}) = (0, \dots, 0)$ in equations (16)–(18).

This means that any solution $\{\mathbf{b}_k\}_{k=1, \dots, N}$ of (20)–(22) induces a solution

$$\left\{ \tilde{\mathbf{b}}_k := \prod_{j=k}^{N-1} (-1)^j \mathbf{b}_k \right\}_{k=1, \dots, N} \tag{23}$$

of the “generic” system (16)–(18). Since we also know by Bezout’s theorem—see, for example, [11, p. 230]—that the systems (16)–(18) all have the same number of solutions (counted with multiplicity over the projective domain), namely 2^N , we can restrict our attention to the solution of just one of these systems—equations (20)–(22), say.

We now denote by $\mathbf{c}_k, \mathbf{a}_k$ the quantities from (13) related to the solution $\mathbf{b}_1, \dots, \mathbf{b}_N$ of the system (20)–(22) defined by $(i_1, \dots, i_{N-1}) = (0, \dots, 0)$, and similarly by $\tilde{\mathbf{c}}_k, \tilde{\mathbf{a}}_k$ those related to the corresponding solution $\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_N$ of the system (16)–(18) for an arbitrary choice of (i_1, \dots, i_{N-1}) . Clearly, we have $\mathbf{c}_0^2 = \tilde{\mathbf{c}}_0^2 = \mathbf{d}_0$ and $\mathbf{a}_N^2 = \tilde{\mathbf{a}}_N^2 = \mathbf{d}_N$.

To obtain the Bézier control points (6) of the arc $\mathbf{r}_k(t)$ we need, according to expressions (4), the complex values

$$\mathbf{c}_{k-1}^2, \quad \mathbf{c}_{k-1}\mathbf{b}_k, \quad \mathbf{b}_k^2, \quad \mathbf{c}_{k-1}\mathbf{a}_k, \quad \mathbf{b}_k\mathbf{a}_k, \quad \mathbf{a}_k^2 \tag{24}$$

in the case of the system (20)–(22), and

$$\tilde{\mathbf{c}}_{k-1}^2, \quad \tilde{\mathbf{c}}_{k-1}\tilde{\mathbf{b}}_k, \quad \tilde{\mathbf{b}}_k^2, \quad \tilde{\mathbf{c}}_{k-1}\tilde{\mathbf{a}}_k, \quad \tilde{\mathbf{b}}_k\tilde{\mathbf{a}}_k, \quad \tilde{\mathbf{a}}_k^2 \tag{25}$$

in the case of (16)–(18). But from (13) and (23) we have

$$\tilde{\mathbf{c}}_k = \prod_{j=k+1}^{N-1} (-1)^j \mathbf{c}_k \quad \text{and} \quad \tilde{\mathbf{a}}_k = \prod_{j=k}^{N-1} (-1)^j \mathbf{a}_k$$

for $k = 1, \dots, N - 1$, and using these relations it is easy to verify that the corresponding quantities in (24) and (25) are identical. Thus, not only do the systems (16)–(18) have closely related solutions, but in all cases these solutions yield *exactly the same set of interpolating C^2 PH splines*.

Hence, by writing $\mathbf{a}_0 := \mathbf{c}_0$, we may completely solve the C^2 PH spline interpolation problem with prescribed end derivatives $\mathbf{a}_0^2 = \mathbf{d}_0$ and $\mathbf{a}_N^2 = \mathbf{d}_N$ by finding all solutions to the following system of quadratic equations:

$$\begin{aligned} &17 \mathbf{b}_1^2 + 3 \mathbf{b}_2^2 + 12 \mathbf{b}_1 \mathbf{b}_2 + 14 \mathbf{a}_0 \mathbf{b}_1 + 2 \mathbf{a}_0 \mathbf{b}_2 + 12 \mathbf{a}_0^2 \\ &= 60 (\mathbf{p}_1 - \mathbf{p}_0), \\ &3 \mathbf{b}_{k-1}^2 + 27 \mathbf{b}_k^2 + 3 \mathbf{b}_{k+1}^2 + 13 \mathbf{b}_{k-1} \mathbf{b}_k + 13 \mathbf{b}_k \mathbf{b}_{k+1} + \mathbf{b}_{k-1} \mathbf{b}_{k+1} \\ &= 60 (\mathbf{p}_k - \mathbf{p}_{k-1}) \quad \text{for} \quad k = 2, \dots, N - 1, \\ &17 \mathbf{b}_N^2 + 3 \mathbf{b}_{N-1}^2 + 12 \mathbf{b}_N \mathbf{b}_{N-1} + 14 \mathbf{a}_N \mathbf{b}_N + 2 \mathbf{a}_N \mathbf{b}_{N-1} + 12 \mathbf{a}_N^2 \\ &= 60 (\mathbf{p}_N - \mathbf{p}_{N-1}). \end{aligned} \tag{26}$$

Note that, because of the choice $(i_1, \dots, i_{N-1}) = (0, \dots, 0)$, we need only be concerned with $\mathbf{a}_0, \dots, \mathbf{a}_N$ and $\mathbf{b}_1, \dots, \mathbf{b}_N$, and equations (13) become

$$\mathbf{a}_k = \frac{1}{2} (\mathbf{b}_k + \mathbf{b}_{k+1}) \quad \text{for} \quad k = 1, \dots, N - 1. \tag{27}$$

2.3. Absence of solutions at infinity

Each of the quadratic equations (26) can be interpreted geometrically as a hyperquadric—i.e., a hypersurface of second order—in N -dimensional complex projective space, \mathbf{P}^N . The homogeneous representation of these hyperquadrics in \mathbf{P}^N is obtained by introducing the homogenizing variable \mathbf{b}_{N+1} via the substitution $\mathbf{b}_k \rightarrow \mathbf{b}_k/\mathbf{b}_{N+1}$ for $k = 1, \dots, N$ in each of the equations (26) and then multiplying through by \mathbf{b}_{N+1}^2 .

The problem of finding the solutions of the system (26) is thus equivalent to determining the intersection points of the corresponding hyperquadrics. Whereas such intersections might, in general, lie in both the affine space \mathbf{R}^N (i.e., $\mathbf{b}_{N+1} \neq 0$) and in the hyperplane at infinity ($\mathbf{b}_{N+1} = 0$), we will show that our particular system (26) never has solutions at infinity.

To establish this, we must verify that the system

$$\begin{aligned} \mathbf{H}_1(\mathbf{b}_1, \dots, \mathbf{b}_N) &:= 17\mathbf{b}_1^2 + 3\mathbf{b}_2^2 + 12\mathbf{b}_1\mathbf{b}_2 = 0, \\ \mathbf{H}_k(\mathbf{b}_1, \dots, \mathbf{b}_N) &:= 3\mathbf{b}_{k-1}^2 + 27\mathbf{b}_k^2 + 3\mathbf{b}_{k+1}^2 + 13\mathbf{b}_{k-1}\mathbf{b}_k + 13\mathbf{b}_k\mathbf{b}_{k+1} \\ &\quad + \mathbf{b}_{k-1}\mathbf{b}_{k+1} = 0 \quad \text{for} \quad k = 2, \dots, N - 1, \\ \mathbf{H}_N(\mathbf{b}_1, \dots, \mathbf{b}_N) &:= 17\mathbf{b}_N^2 + 3\mathbf{b}_{N-1}^2 + 12\mathbf{b}_N\mathbf{b}_{N-1} = 0, \end{aligned} \tag{28}$$

obtained by setting $\mathbf{b}_{N+1} = 0$ in the homogeneous version of (26), admits only the trivial solution² $\mathbf{b}_1 = \dots = \mathbf{b}_N = 0$.

We can find solutions of (28) by successively eliminating the variables $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{N-1}$, and thus reducing the system of N quadratic equations to one equation of order 2^N in \mathbf{b}_N . According to [10, p. 87], this may be done by successively building Sylvester’s resultants³

$$\mathbf{R}_k = \text{Resultant}_{\mathbf{b}_{k-1}}(\mathbf{R}_{k-1}, \mathbf{H}_k) \tag{29}$$

with respect to \mathbf{b}_{k-1} for $k = 2, \dots, N$, where $\mathbf{R}_1 = \mathbf{H}_1$ and $\mathbf{R}_N = 0$ represents the above-mentioned equation of order 2^N in \mathbf{b}_N .

Now the first step is to show that \mathbf{R}_k for $k = 1, \dots, N - 1$ is of the form

$$\mathbf{R}_k = \sum_{r=0}^{2^k} \alpha_{k,r} \mathbf{b}_k^{2^k-r} \mathbf{b}_{k+1}^r \tag{30}$$

for certain real coefficients $\alpha_{k,r}$. Equation (30) is proved by induction on k . $\mathbf{R}_1 = \mathbf{H}_1$ is obviously of the form (30). By assuming (30) to be valid for a certain k in the

² The values $(\mathbf{b}_1, \dots, \mathbf{b}_N, \mathbf{b}_{N+1}) = (0, \dots, 0, 0)$ do not identify a proper point of \mathbf{P}^N .

³ For the formal definition of Sylvester’s resultant see e.g. [10, p. 86].

induction step, we have to show that \mathbf{R}_{k+1} has the form

$$\mathbf{R}_{k+1} = \sum_{r=0}^{2^{k+1}} \alpha_{k+1,r} \mathbf{b}_{k+1}^{2^{k+1}-r} \mathbf{b}_{k+2}^r \tag{31}$$

for certain coefficients $\alpha_{k+1,r}$. By definition, $\mathbf{R}_{k+1} = \text{Resultant}_{\mathbf{b}_k}(\mathbf{R}_k, \mathbf{H}_{k+1})$ is the resultant of the polynomials

$$\begin{aligned} \mathbf{R}_k &= \sum_{r=0}^{2^k} \alpha_{k,r} \mathbf{b}_k^{2^k-r} \mathbf{b}_{k+1}^r, \\ \mathbf{H}_{k+1} &= 3 \mathbf{b}_k^2 + (13 \mathbf{b}_{k+1} + \mathbf{b}_{k+2}) \mathbf{b}_k \\ &\quad + 27 \mathbf{b}_{k+1}^2 + 13 \mathbf{b}_{k+1} \mathbf{b}_{k+2} + 3 \mathbf{b}_{k+2}^2 \end{aligned}$$

with respect to \mathbf{b}_k . By using the abbreviations

$$\mathbf{l}_j^i = \gamma_{ij} \mathbf{b}_{k+1}^i \quad \text{and} \quad \mathbf{q}_j^i = \sum_{r=0}^i \delta_{ijr} \mathbf{b}_{k+1}^{i-r} \mathbf{b}_{k+2}^r \tag{32}$$

with suitable real values γ_{ij} and δ_{ijr} for $i = 0, \dots, 2^{k+1}$ and $j = 0, \dots, 2^k + 1$, \mathbf{R}_{k+1} may be written (see, e.g., [10, p. 86]) as the determinant

$$\mathbf{R}_{k+1} = \begin{vmatrix} \mathbf{q}_0^0 & \mathbf{q}_0^1 & \mathbf{q}_0^2 & & & & & & \\ & \mathbf{q}_0^0 & \mathbf{q}_0^1 & \mathbf{q}_0^2 & & & & & \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \\ \dots & \dots & \dots & \dots & \dots & \dots & \mathbf{q}_0^0 & \mathbf{q}_0^1 & \mathbf{q}_0^2 \\ \mathbf{l}_0^0 & \mathbf{l}_0^1 & \mathbf{l}_0^2 & \mathbf{l}_0^3 & \dots & \dots & \mathbf{l}_0^{2^k} & & \\ & \mathbf{l}_0^0 & \mathbf{l}_0^1 & \mathbf{l}_0^2 & \mathbf{l}_0^3 & \dots & \dots & \mathbf{l}_0^{2^k} & \end{vmatrix} =: \mathbf{R}_{k+1}^{(0)}$$

of order $2^k + 2$. Since elementary row operations do not alter the value of a determinant, we now subtract suitable multiples of the first and second rows from the $(2^k + 1)$ -th and $(2^k + 2)$ -th rows, respectively, and obtain

$$\mathbf{R}_{k+1}^{(1)} := \begin{vmatrix} \mathbf{q}_0^0 & \mathbf{q}_0^1 & \mathbf{q}_0^2 & & & & & & \\ & \mathbf{q}_0^0 & \mathbf{q}_0^1 & \mathbf{q}_0^2 & & & & & \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \\ \dots & \dots & \dots & \dots & \dots & \dots & \mathbf{q}_0^0 & \mathbf{q}_0^1 & \mathbf{q}_0^2 \\ & \mathbf{q}_1^1 & \mathbf{q}_1^2 & \mathbf{l}_0^3 & \dots & \dots & \mathbf{l}_0^{2^k} & & \\ & & \mathbf{q}_1^1 & \mathbf{q}_1^2 & \mathbf{l}_0^3 & \dots & \dots & \mathbf{l}_0^{2^k} & \end{vmatrix}$$

By continuing this process (with $\mathbf{R}_{k+1}^{(j)}$ generated by subtracting a suitable multiple of rows j and $j + 1$ of $\mathbf{R}_{k+1}^{(j-1)}$ from rows $2^k + 1$ and $2^k + 2$, respectively) after

$$\begin{aligned}
 \mathbf{f}_N(\mathbf{b}_1, \dots, \mathbf{b}_N) &= 17 \mathbf{b}_N^2 + 3 \mathbf{b}_{N-1}^2 + 12 \mathbf{b}_N \mathbf{b}_{N-1} \\
 &\quad + 14 \mathbf{a}_N \mathbf{b}_N + 2 \mathbf{a}_N \mathbf{b}_{N-1} + 12 \mathbf{a}_N^2 \\
 &\quad - 60 (\mathbf{p}_N - \mathbf{p}_{N-1}) = 0
 \end{aligned}
 \tag{33}$$

of N quadratic equations in N complex unknowns $\mathbf{b}_1, \dots, \mathbf{b}_N$. If \mathbf{a}_0 and \mathbf{a}_N are given, we know by Bezout’s theorem that there are 2^N solutions altogether (counted with multiplicity over the projective domain; see section 2).

Now the values \mathbf{a}_0 and \mathbf{a}_N appearing in $\mathbf{f}_1(\mathbf{b}_1, \dots, \mathbf{b}_N)$ and $\mathbf{f}_N(\mathbf{b}_1, \dots, \mathbf{b}_N)$ are determined by the given end-derivatives \mathbf{d}_0 and \mathbf{d}_N , they must satisfy

$$\mathbf{a}_0^2 = \mathbf{d}_0 \quad \text{and} \quad \mathbf{a}_N^2 = \mathbf{d}_N.
 \tag{34}$$

Since the above conditions yield *two* complex values for each of \mathbf{a}_0 and \mathbf{a}_N , it might appear that there are 2^{N+2} sets of values $\mathbf{a}_0, \dots, \mathbf{a}_N$ and $\mathbf{b}_1, \dots, \mathbf{b}_N$ satisfying (33) and (34)—together with equations (27)—that define C^2 piecewise-PH-quintic interpolants to the data $\mathbf{p}_0, \dots, \mathbf{p}_N$ and $\mathbf{d}_0, \mathbf{d}_N$.

However, if values $\mathbf{a}_0, \dots, \mathbf{a}_N$ and $\mathbf{b}_1, \dots, \mathbf{b}_N$ satisfy these equations, then so also do $-\mathbf{a}_0, \dots, -\mathbf{a}_N$ and $-\mathbf{b}_1, \dots, -\mathbf{b}_N$, and it is clear from expressions (10), (12), and (13) that these two solutions yield *exactly the same curve*. As the two values defined by (34) for each of \mathbf{a}_0 and \mathbf{a}_N differ merely by the factor -1 , inserting all four combinations of these values into (33) would result in a systematic duplication of the interpolant curves. By choosing a specific value for (say) \mathbf{a}_N in combination with the two possible values of \mathbf{a}_0 , one generates 2^{N+1} distinct interpolants without replication.

We have implemented two predictor–corrector path–following programs to solve the system (33). The first, described in section 3.1, is “customized” to the problem at hand: it makes explicit use of complex arithmetic and the tridiagonality of the Jacobian. For simplicity, this code uses only constant, linear prediction steps in the homotopy parameter. The second program, described in section 3.2, was adapted from an established code [16]. It employs a higher–order adaptive stepping procedure with respect to arc–length along the homotopy path, and is implemented in real arithmetic.

3.1. Customized predictor–corrector path–following

By deleting certain terms from each of the equations $\mathbf{f}_j(\mathbf{b}_1, \dots, \mathbf{b}_N) = 0$ for $j = 1, \dots, N$, the system (33) can be reduced to an “initial system” that admits a simple closed–form solution (see, e.g., [15]). In particular, we take

$$\begin{aligned}
 \mathbf{g}_1(\mathbf{b}_1, \dots, \mathbf{b}_N) &= 17 \mathbf{b}_1^2 + 14 \mathbf{a}_0 \mathbf{b}_1 + 12 \mathbf{a}_0^2 \\
 &\quad - 60 (\mathbf{p}_1 - \mathbf{p}_0) = 0, \\
 \mathbf{g}_j(\mathbf{b}_1, \dots, \mathbf{b}_N) &= 27 \mathbf{b}_j^2 - 60 (\mathbf{p}_j - \mathbf{p}_{j-1}) = 0 \\
 &\quad \text{for} \quad j = 2, \dots, N - 1,
 \end{aligned}$$

$$\mathbf{g}_N(\mathbf{b}_1, \dots, \mathbf{b}_N) = 17 \mathbf{b}_N^2 + 14 \mathbf{a}_N \mathbf{b}_N + 12 \mathbf{a}_N^2 - 60 (\mathbf{p}_N - \mathbf{p}_{N-1}) = 0 \tag{35}$$

as our starting system—it has the straightforward solutions

$$\begin{aligned} \mathbf{b}_1 &= \frac{-7\mathbf{a}_0 \pm \sqrt{1020(\mathbf{p}_1 - \mathbf{p}_0) - 155\mathbf{a}_0^2}}{17}, \\ \mathbf{b}_j &= \pm \frac{2}{3} \sqrt{5(\mathbf{p}_j - \mathbf{p}_{j-1})} \quad \text{for } j = 2, \dots, N-1, \\ \mathbf{b}_N &= \frac{-7\mathbf{a}_N \pm \sqrt{1020(\mathbf{p}_N - \mathbf{p}_{N-1}) - 155\mathbf{a}_N^2}}{17}. \end{aligned} \tag{36}$$

By considering all combinations of signs in the above expressions we obtain, in general, 2^N distinct sets of starting values for the quantities $\mathbf{b}_0, \dots, \mathbf{b}_N$.

The homotopy method amounts to “continuously deforming” the initial equation system (35) into the desired system (33) while steadily tracking the motion of the solutions from their initial values (36) during the deformation process. We shall use the simple homotopy defined for $j = 1, \dots, N$ by

$$\begin{aligned} \mathbf{h}_j(\mathbf{b}_1, \dots, \mathbf{b}_N, \lambda) &= \lambda \mathbf{f}_j(\mathbf{b}_1, \dots, \mathbf{b}_N) \\ &+ (1 - \lambda) \mathbf{z} \mathbf{g}_j(\mathbf{b}_1, \dots, \mathbf{b}_N) = 0 \end{aligned} \tag{37}$$

where λ is a *real* “homotopy parameter” and \mathbf{z} is a non-zero complex number. As λ increases from 0 to 1, expressions (37) yield a continuous transformation from the start system (35) with known solutions (36) to the system (33) that we wish to solve. Corresponding to each λ value, the system (37) will possess 2^N solution sets $\mathbf{b}_1(\lambda), \dots, \mathbf{b}_N(\lambda)$ that we can track from the known solutions (36) at $\lambda = 0$ to the desired ones at $\lambda = 1$.

The complex number \mathbf{z} in (37) is a point on the unit circle, $\mathbf{z} = \exp(i\phi)$, corresponding to a randomly-chosen angle ϕ . Morgan and Sommese [14] and Zulehner [17] show that, for all but a finite set of angles ϕ , the 2^N solution paths $\mathbf{b}_1(\lambda), \dots, \mathbf{b}_N(\lambda)$ are regular for $0 < \lambda < 1$, i.e., the *Jacobian matrix* \mathbf{M} with elements

$$\mathbf{M}_{jk} = \frac{\partial \mathbf{h}_j}{\partial \mathbf{b}_k} \quad \text{for } 1 \leq j, k \leq N \tag{38}$$

is non-singular for each position along these solution paths. For the functions (37), the Jacobian matrix is tridiagonal—its non-zero elements are

$$\begin{aligned} \mathbf{M}_{k,k-1} &= \lambda (6 \mathbf{b}_{k-1} + 13 \mathbf{b}_k + \mathbf{b}_{k+1}), \\ \mathbf{M}_{kk} &= 13 \lambda (\mathbf{b}_{k-1} + \mathbf{b}_{k+1}) + 54 (\lambda + (1 - \lambda)\mathbf{z}) \mathbf{b}_k, \\ \mathbf{M}_{k,k+1} &= \lambda (\mathbf{b}_{k-1} + 13 \mathbf{b}_k + 6 \mathbf{b}_{k+1}) \end{aligned}$$

for rows $k = 2, \dots, N - 1$, while for rows 1 and N we have:

$$\begin{aligned} \mathbf{M}_{11} &= 12 \lambda \mathbf{b}_2 + (\lambda + (1 - \lambda)\mathbf{z}) (34 \mathbf{b}_1 + 14 \mathbf{a}_0), \\ \mathbf{M}_{12} &= \lambda (12 \mathbf{b}_1 + 6 \mathbf{b}_2 + 2 \mathbf{a}_0), \\ \mathbf{M}_{N,N-1} &= \lambda (6 \mathbf{b}_{N-1} + 12 \mathbf{b}_N + 2 \mathbf{a}_N), \\ \mathbf{M}_{NN} &= 12 \lambda \mathbf{b}_{N-1} + (\lambda + (1 - \lambda)\mathbf{z}) (34 \mathbf{b}_N + 14 \mathbf{a}_N). \end{aligned}$$

In order to trace solution paths from the initial positions (36) at $\lambda = 0$ to the desired solutions at $\lambda = 1$, we employ a predictor–corrector method. For a step $\Delta\lambda$ in the homotopy parameter, the *prediction* stage consists of computing increments $\Delta\mathbf{b}_1, \dots, \Delta\mathbf{b}_N$ to the current $\mathbf{b}_1, \dots, \mathbf{b}_N$ values that correspond to a linear expansion of equations (37) about the current position. In other words, we step a distance $\Delta\lambda$ along the *tangent direction* to the one–dimensional locus in $\mathbf{C}^N \times \mathbf{R}$ defined by equations (37). This gives rise to the tridiagonal (linear) system of equations

$$\sum_{k=1}^N \mathbf{M}_{jk} \Delta\mathbf{b}_k = (\mathbf{z} \mathbf{g}_j - \mathbf{f}_j) \Delta\lambda, \quad j = 1, \dots, N \tag{39}$$

for the increments $\Delta\mathbf{b}_1, \dots, \Delta\mathbf{b}_N$, where it is understood that the Jacobian matrix elements (38) and the functions \mathbf{f}_j and \mathbf{g}_j are to be evaluated *at the current position* $(\mathbf{b}_1, \dots, \mathbf{b}_N, \lambda)$.

Owing to the curvature of the solution locus, the prediction step generally incurs an appreciable deviation from the precise solution of (37) at $\lambda + \Delta\lambda$. To compensate for this, we employ Newton–Raphson *correction* iterations

$$\mathbf{b}_k^{(r+1)} = \mathbf{b}_k^{(r)} + \delta\mathbf{b}_k \quad \text{for} \quad k = 1, \dots, N$$

and $r = 1, 2, \dots$, where the starting approximation $\mathbf{b}_1^{(0)}, \dots, \mathbf{b}_N^{(0)}$ corresponds to the outcome of the prediction step. At the r –th iteration, the increments $\delta\mathbf{b}_1, \dots, \delta\mathbf{b}_N$ are solutions to the tridiagonal system

$$\sum_{k=1}^N \mathbf{M}_{jk}^{(r)} \delta\mathbf{b}_k = -\mathbf{h}_j^{(r)}, \quad j = 1, \dots, N, \tag{40}$$

the matrix elements $\mathbf{M}_{jk}^{(r)}$ and functions $\mathbf{h}_j^{(r)}$ being evaluated at $(\mathbf{b}_1^{(r)}, \dots, \mathbf{b}_N^{(r)}, \lambda + \Delta\lambda)$. The correction iterations proceed until the quantity

$$\epsilon_{\mathbf{h}}^{(r)} = \sqrt{\sum_{j=1}^N |\mathbf{h}_j(\mathbf{b}_1^{(r)}, \dots, \mathbf{b}_N^{(r)}, \lambda + \Delta\lambda)|^2} \tag{41}$$

diminishes below a prescribed tolerance value: typically, only a few iterations are needed when this value is set to $\sim 10^{-10}$ and $\Delta\lambda$ is not too large.

Note that both the prediction step (39) and correction iterations (40) each incur only a *tridiagonal linear system*, which can be solved [4] in just $3N - 3$ additions and multiplications and $2N - 1$ divisions. Once a solution $\mathbf{b}_1, \dots, \mathbf{b}_N$ to (33) has been

determined to sufficient accuracy, for given \mathbf{a}_0 and \mathbf{a}_N values, one computes the remaining values $\mathbf{a}_1, \dots, \mathbf{a}_{N-1}$ from (27). Finally, integrating the hodographs (10) with $\mathbf{r}_1(0) = \mathbf{p}_0$ and $\mathbf{r}_k(1) = \mathbf{r}_{k+1}(0)$ for $k = 1, \dots, N - 1$, completes the construction of a C^2 PH–quintic spline. This process is repeated for each of the 2^N sets of starting values (36) and the two values of (say) \mathbf{a}_0 .

3.2. General predictor–corrector path–following

We now briefly describe a homotopy scheme for *general* systems of N polynomial equations in N complex unknowns, and we compare it to the “customized” one described above. Among the many available predictor–corrector path–following methods [2], we selected the one implemented in the HOMPACK [2,16] software package. If d_j is the degree of the polynomial f_j in (33), the “initial” system reads

$$\mathbf{g}_j(\mathbf{b}_1, \dots, \mathbf{b}_N) = \mathbf{u}_j^d \mathbf{b}_j^{d_j} - \mathbf{v}_j^{d_j} = 0 \quad \text{for } j = 1, \dots, N \quad (42)$$

where \mathbf{u}_j and \mathbf{v}_j are randomly chosen complex values. For each j , equation (42) has the straightforward solutions

$$\frac{\mathbf{v}_j}{\mathbf{u}_j} \exp\left(i \frac{2\pi k}{d_j}\right) \quad \text{for } k = 0, \dots, d_j - 1.$$

The homotopy map is then given by

$$\mathbf{h}_j(\lambda, \mathbf{b}_1, \dots, \mathbf{b}_N) = \lambda \mathbf{f}_j(\mathbf{b}_1, \dots, \mathbf{b}_N) + (1 - \lambda) \mathbf{g}_j(\mathbf{b}_1, \dots, \mathbf{b}_N) = 0 \quad (43)$$

where $\lambda \in [0, 1]$. By introducing the *real* $2N$ –dimensional vectors⁵

$$B := \begin{pmatrix} \text{Re}(\mathbf{b}_1) \\ \text{Im}(\mathbf{b}_1) \\ \vdots \\ \text{Re}(\mathbf{b}_N) \\ \text{Im}(\mathbf{b}_N) \end{pmatrix}, \quad F := \begin{pmatrix} \text{Re}(\mathbf{f}_1) \\ \text{Im}(\mathbf{f}_1) \\ \vdots \\ \text{Re}(\mathbf{f}_N) \\ \text{Im}(\mathbf{f}_N) \end{pmatrix}, \quad G := \begin{pmatrix} \text{Re}(\mathbf{g}_1) \\ \text{Im}(\mathbf{g}_1) \\ \vdots \\ \text{Re}(\mathbf{g}_N) \\ \text{Im}(\mathbf{g}_N) \end{pmatrix}, \quad H := \begin{pmatrix} \text{Re}(\mathbf{h}_1) \\ \text{Im}(\mathbf{h}_1) \\ \vdots \\ \text{Re}(\mathbf{h}_N) \\ \text{Im}(\mathbf{h}_N) \end{pmatrix}$$

the homotopy path, which we parametrize with respect to its arc length s , may be written as

$$\begin{pmatrix} \lambda(s) \\ B(s) \end{pmatrix}, \quad (44)$$

and it satisfies

$$H(\lambda(s), B(s)) \equiv 0.$$

The algorithm comprises four phases–prediction, correction, estimation of step

⁵ $\text{Re}(x)$ and $\text{Im}(x)$ denote the real and imaginary parts of the complex value x . The complex continuation system (43) is thus replaced by a real one, and the problem is solved in $2N$ –dimensional real space. For details, see [13, pp. 284–290].

size, and computation of the solution at $\lambda = 1$. For the latter two phases, we refer the reader to [16]; here we briefly describe only the prediction and correction step to allow a comparison with the customized algorithm. For the prediction of the i -th point⁶ on the homotopy path (44) the cubic Hermite interpolant, uniquely determined by the previous two points and tangent vectors,

$$(\lambda(s_{i-2}), B(s_{i-2})), \left(\frac{d\lambda}{ds}(s_{i-2}), \frac{dB}{ds}(s_{i-2}) \right)$$

and

$$(\lambda(s_{i-1}), B(s_{i-1})), \left(\frac{d\lambda}{ds}(s_{i-1}), \frac{dB}{ds}(s_{i-1}) \right),$$

is formed and evaluated at $s_i = s_{i-1} + \Delta s$. The resulting point, denoted by $(\lambda^{(0)}(s_i), B^{(0)}(s_i))$, is then used as a starting approximation for Newton–Raphson correction iterations

$$\begin{pmatrix} \lambda^{(r+1)}(s_i) \\ B^{(r+1)}(s_i) \end{pmatrix} = \begin{pmatrix} \lambda^{(r)}(s_i) \\ B^{(r)}(s_i) \end{pmatrix} + \begin{pmatrix} \delta\lambda \\ \delta B \end{pmatrix} \tag{45}$$

for $r = 0, 1, 2, \dots$. At the $(r + 1)$ -th iteration, the increments $(\delta\lambda, \delta B)$ are solutions to the linear system

$$DH_i^{(r)} \begin{pmatrix} \delta\lambda \\ \delta B \end{pmatrix} = - H_i^{(r)}, \tag{46}$$

where $H_i^{(r)}$ and the $2N \times (2N + 1)$ -dimensional Jacobian $DH_i^{(r)}$ (which has full rank along the homotopy paths for randomly chosen coefficients $\mathbf{u}_j, \mathbf{v}_j$ in (42) – see, for example, [16, p. 302]) are evaluated at $(\lambda^{(r)}(s_i), B^{(r)}(s_i))$.

To determine the general solution of the linear system (46), we require a particular solution $P^{(r)} \in \mathbf{R}^{2N+1}$ of the inhomogeneous system and a vector $V^{(r)} \in \mathbf{R}^{2N+1}$ in the one-dimensional kernel of $DH_i^{(r)}$. They are calculated by *QR*-factorization of the matrix $DH_i^{(r)}$ in each step. From the totality of solutions $P^{(r)} + \mu V^{(r)}$ with $\mu \in \mathbf{R}$, the one with the minimum norm is then selected for the increments $(\delta\lambda, \delta B)$. Once such a point $(\lambda^{(k+1)}(s_i), B^{(k+1)}(s_i))$ has converged sufficiently close to the homotopy path (i.e., within a prescribed error bound), it is accepted as the i -th point on the path, and the vector $V^{(k-1)} / \|V^{(k-1)}\|$ calculated in the k -th iteration (45) is accepted as the unit tangent vector at this point.

Although the general predictor–corrector path following method is far more sophisticated than the “customized” scheme described above, the latter was found to be more efficient and very reliable in practice. There was no discernible difference in the accuracy of the results generated by the two methods.

⁶ For $i = 2$, only the tangent is used.

4. End conditions

In many contexts, the need to specify end-derivatives \mathbf{d}_0 and \mathbf{d}_N in addition to the points $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_N$ is a rather artificial requirement. It is desirable to formulate suitable *end conditions* for PH splines that circumvent this need.

Three forms of end condition are commonly enforced on the “ordinary” C^2 cubic spline—(i) quadratic end-spans; (ii) the periodic end condition; and (iii) the “not-a-knot” end condition (so-named because enforcing C^3 continuity at the junctures of the first and last pair of cubic segments implies that these pairs actually correspond to *single* cubic curves).

We shall formulate analogs of conditions (i) and (ii) for the PH splines. It is also easy (but somewhat more cumbersome) to derive conditions that ensure C^3 continuity between the first and last pair of quintic segments for PH splines, though it would perhaps be misleading to call these “not-a-knot” conditions—for brevity we shall omit the details.

4.1. Cubic end spans

The simplest means of avoiding the need to explicitly specify \mathbf{d}_0 and \mathbf{d}_N is to force the initial and final spans to be *cubic* PH-curve segments.⁷ A cubic PH segment arises if the quadratic (3) defining the hodograph (1) is actually a *degree-elevated linear polynomial*, which corresponds [6] to the condition $\mathbf{w}_1 = \frac{1}{2}(\mathbf{w}_0 + \mathbf{w}_2)$. In terms of the quantities $\mathbf{c}_0, \dots, \mathbf{c}_{N-1}$ and $\mathbf{b}_1, \dots, \mathbf{b}_N$ and $\mathbf{a}_1, \dots, \mathbf{a}_N$ defined by (12) this gives $\mathbf{b}_1 = \frac{1}{2}(\mathbf{c}_0 + \mathbf{a}_1)$ and $\mathbf{b}_N = \frac{1}{2}(\mathbf{c}_{N-1} + \mathbf{a}_N)$, and by using the relations (13) we can solve explicitly for \mathbf{c}_0 and \mathbf{a}_N as

$$\mathbf{c}_0 = \frac{1}{2} [3\mathbf{b}_1 - (-1)^i \mathbf{b}_2] \quad \text{and} \quad \mathbf{a}_N = \frac{1}{2} [3\mathbf{b}_N - (-1)^{i_{N-1}} \mathbf{b}_{N-1}]. \quad (47)$$

Substituting into (17) and (18), the latter equations can thus be replaced by

$$\begin{aligned} 13\mathbf{b}_1^2 + \mathbf{b}_2^2 - 2(-1)^i \mathbf{b}_1 \mathbf{b}_2 &= 12(\mathbf{p}_1 - \mathbf{p}_0), \\ 13\mathbf{b}_N^2 + \mathbf{b}_{N-1}^2 - 2(-1)^{i_{N-1}} \mathbf{b}_N \mathbf{b}_{N-1} &= 12(\mathbf{p}_N - \mathbf{p}_{N-1}). \end{aligned} \quad (48)$$

For each instance of $(i_1, \dots, i_{N-1}) \in \{0, 1\}^{N-1}$ the above equations, together with (16), again comprise a tridiagonal system of N quadratic equations for the N unknowns $\mathbf{b}_1, \dots, \mathbf{b}_N$.

Now by substituting expression (19) into this system we obtain

$$\begin{aligned} 13\mathbf{b}_1^2 + \mathbf{b}_2^2 - 2\mathbf{b}_1 \mathbf{b}_2 &= 12(\mathbf{p}_1 - \mathbf{p}_0), \\ 13\mathbf{b}_N^2 + \mathbf{b}_{N-1}^2 - 2\mathbf{b}_N \mathbf{b}_{N-1} &= 12(\mathbf{p}_N - \mathbf{p}_{N-1}) \end{aligned} \quad (49)$$

together with equations (21); this new system also corresponds to the special choice $(i_1, \dots, i_{N-1}) = (0, \dots, 0)$ in (16) and (48). The same reasoning as in section 2.2 (with $\tilde{\mathbf{c}}_0 = \prod_{j=1}^{N-1} (-1)^j \mathbf{c}_0$ and $\tilde{\mathbf{a}}_N = \mathbf{a}_N$) then leads, again, to the conclusion that

⁷ These are actually portions of a unique curve, known as Tschirnhausen’s cubic [9].

all the systems defined by (16) and (48) yield exactly the same set of interpolating C^2 piecewise-PH quintics.

We may thus restrict our attention to just one of these systems—e.g., that defined by equations (21) and (49), which can be solved by the methods described in section 3. For the starting system of section 3.1, $\mathbf{g}_1(\mathbf{b}_1, \dots, \mathbf{b}_N)$ and $\mathbf{g}_N(\mathbf{b}_1, \dots, \mathbf{b}_N)$ are defined by retaining only the first terms on the left in each of equations (49), while $\mathbf{g}_j(\mathbf{b}_1, \dots, \mathbf{b}_N)$ for $j = 2, \dots, N - 1$ are as in (35).

Once the $\mathbf{b}_1, \dots, \mathbf{b}_N$ values are known, we may obtain $\mathbf{a}_0 = \mathbf{c}_0$ and $\mathbf{a}_1, \dots, \mathbf{a}_N$ from equations (13) and (47). Since equations (47) are just *linear*, as compared to the *quadratic* equations (34) for the case of specified end-derivatives, there are in general just 2^{N-1} (rather than 2^{N+1}) distinct PH splines having cubic end spans. Although the end-spans are now actually cubics, they will normally be represented in degree-elevated form as PH quintics.

For open curves, the cubic end-span formulation is the most convenient means of avoiding the need to choose values for the end-derivatives $\mathbf{d}_0, \mathbf{d}_N$. It can generally be relied upon to produce more pleasing loci than even the most well-intentioned “guesses” for \mathbf{d}_0 and \mathbf{d}_N .

4.2. Periodic end condition

Periodic end conditions are necessary when a closed C^2 curve is to be created, which may be indicated by coincidence of the initial and final point: $\mathbf{p}_N = \mathbf{p}_0$. At this common juncture of the first and last span, the continuity conditions $\mathbf{r}'_N(1) = \mathbf{r}'_1(0)$ and $\mathbf{r}''_N(1) = \mathbf{r}''_1(0)$ yield the equations

$$\mathbf{a}_N^2 = \mathbf{c}_0^2 \quad \text{and} \quad \mathbf{a}_N^2 - \mathbf{a}_N \mathbf{b}_N = \mathbf{c}_0 \mathbf{b}_1 - \mathbf{c}_0^2,$$

which may be solved for \mathbf{c}_0 and \mathbf{a}_N in terms of \mathbf{b}_1 and \mathbf{b}_N :

$$\mathbf{c}_0 = \frac{1}{2} [\mathbf{b}_1 + (-1)^{i_0} \mathbf{b}_N] \quad \text{and} \quad \mathbf{a}_N = \frac{1}{2} [\mathbf{b}_N + (-1)^{i_0} \mathbf{b}_1], \quad (50)$$

where $i_0 \in \{0, 1\}$. By substituting the latter expressions into (17) and (18) we thus obtain

$$\begin{aligned} 3 \mathbf{b}_N^2 + 27 \mathbf{b}_1^2 + 3 \mathbf{b}_2^2 + 13 (-1)^{i_1} \mathbf{b}_1 \mathbf{b}_2 + 13 (-1)^{i_0} \mathbf{b}_N \mathbf{b}_1 \\ + (-1)^{i_0+i_1} \mathbf{b}_N \mathbf{b}_2 = 60 (\mathbf{p}_1 - \mathbf{p}_0), \\ 3 \mathbf{b}_{N-1}^2 + 27 \mathbf{b}_N^2 + 3 \mathbf{b}_1^2 + 13 (-1)^{i_{N-1}} \mathbf{b}_{N-1} \mathbf{b}_N + 13 (-1)^{i_0} \mathbf{b}_N \mathbf{b}_1 \\ + (-1)^{i_0+i_{N-1}} \mathbf{b}_{N-1} \mathbf{b}_1 = 60 (\mathbf{p}_N - \mathbf{p}_{N-1}). \end{aligned} \quad (51)$$

For each instance of $(i_1, \dots, i_{N-1}) \in \{0, 1\}^{N-1}$ the above equations, together with (16), amount to a system of N quadratic equations for the unknowns $\mathbf{b}_1, \dots, \mathbf{b}_N$. By again substituting (19) into (16) and (51) we obtain the system comprising

$$\begin{aligned} 3 \mathbf{b}_N^2 + 27 \mathbf{b}_1^2 + 3 \mathbf{b}_2^2 + 13 \mathbf{b}_1 \mathbf{b}_2 \pm 13 \mathbf{b}_N \mathbf{b}_1 \pm \mathbf{b}_N \mathbf{b}_2 \\ = 60 (\mathbf{p}_1 - \mathbf{p}_0), \end{aligned}$$

$$\begin{aligned}
 3 \mathbf{b}_{N-1}^2 + 27 \mathbf{b}_N^2 + 3 \mathbf{b}_1^2 + 13 \mathbf{b}_{N-1} \mathbf{b}_N \pm 13 \mathbf{b}_N \mathbf{b}_1 \pm \mathbf{b}_{N-1} \mathbf{b}_1 \\
 = 60 (\mathbf{p}_N - \mathbf{p}_{N-1})
 \end{aligned} \tag{52}$$

together with equations (21)—this may also be identified with the special choice $(i_1, \dots, i_{N-1}) = (0, \dots, 0)$ in equations (16) and (51).

Again, we need only solve the system defined by (21) and (52) to generate all distinct solutions of the C^2 PH quintic interpolation problem (reasoning as in section 2.2, with $\tilde{\mathbf{c}}_0 = \prod_{j=1}^{N-1} (-1)^j \frac{1}{2} [\mathbf{b}_1 + \prod_{j=0}^{N-1} (-1)^j \mathbf{b}_N]$ and $\tilde{\mathbf{a}}_N = \frac{1}{2} [\mathbf{b}_N + \prod_{j=0}^{N-1} (-1)^j \mathbf{b}_1]$). Note, however, that this system—and hence its Jacobian—is no longer tridiagonal. Thus, a general purpose linear-equations solver is required to execute each prediction step and each correction iteration.

Note that we choose the plus or minus signs in (52) according to whether $i_0 = 0$ or $i_0 = 1$ in (50). For the starting system of section 3.1, we retain only the terms $27 \mathbf{b}_1^2$ and $27 \mathbf{b}_N^2$ on the left in (52). Once $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_N$ are known, the corresponding values $\mathbf{a}_0 = \mathbf{c}_0$ and $\mathbf{a}_1, \dots, \mathbf{a}_N$ are given by (13) and (50). As in the case of specified end-derivatives, there are in general 2^{N+1} distinct C^2 periodic PH splines interpolating points $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_N$ (with $\mathbf{p}_N = \mathbf{p}_0$).

5. Empirical results

The “customized” homotopy method of section 3.1 was coded up in a double-precision, complex-arithmetic FORTRAN program. The use of complex arithmetic, and the fact that little more than a tridiagonal solver is needed (except in the case of periodic end conditions) allowed a remarkably compact and bug-free implementation. Having no prior experience with the homotopy method, the authors were pleasantly surprised with its accuracy, efficiency, and reliability.

The “general” homotopy method discussed in section 3.2 is available in the POLSYS polynomial system solver, a suite of FORTRAN subroutines in the HOMPACK software package [16,2]. This solver, originally designed for solving polynomial systems with real coefficients, was modified to deal with systems having complex coefficients. The option of projective transformation available in POLSYS for eliminating divergent homotopy paths (see, e.g., [16, p. 303] and [15, p. 124], which is useful for systems with solutions at infinity, was not needed for the system (33) we are concerned with because of the absence of solutions in the hyperplane at infinity (see section 2.3).

For the particular system (33) we observed that the customized program and the POLSYS code produced solutions of comparable accuracy, though the former was less demanding in terms of cpu time and memory. The following discussion and examples, therefore, pertain to results generated by the customized program.

Numerous tests have shown that the homotopy system (37) is, in general, exceedingly well-conditioned, allowing PH splines to be constructed to a very high level of precision (see the data in table 1 below). We have employed constant steps in the homotopy parameter, $\Delta\lambda = 0.02$, for the examples described herein. For most

Table 1
Comparison of nodal derivatives and “gaps” for Example 1.

node	
1	Δz_1 (0.000000000000000, 0.000000000000000)
	$r'_1(1)$ (1.360779002855208, -2.652974680926679)
	$r'_2(0)$ (1.360779002855208, -2.652974680926679)
	$r''_1(1)$ (5.632083355923054, -0.182314923482134)
	$r''_2(0)$ (5.632083355923054, -0.182314923482134)
	2
$r'_2(1)$ (2.048883028548189, 2.211132842356474)	
$r'_3(0)$ (2.048883028548189, 2.211132842356474)	
$r''_2(1)$ (-7.081869673180421, 5.944174277861420)	
$r''_3(0)$ (-7.081869673180421, 5.944174277861420)	
3	
	$r'_3(1)$ (2.842303868097108, 1.968545850008491)
	$r'_4(0)$ (2.842303868097108, 1.968545850008490)
	$r''_3(1)$ (8.462769563261957, -7.740519332527053)
	$r''_4(0)$ (8.462769563261962, -7.740519332527053)
	4
$r'_4(1)$ (-0.755435263095192, -2.914313055871553)	
$r'_5(0)$ (-0.755435263095192, -2.914313055871553)	
$r''_4(1)$ (-9.583644688862023, 5.647009445446316)	
$r''_5(0)$ (-9.583644688862023, 5.647009445446313)	
5	
	$r'_5(1)$ (-1.735966775898855, -1.230156974065111)
	$r'_6(0)$ (-1.735966775898856, -1.230156974065111)
	$r''_5(1)$ (4.743411905896337, -3.811925035449617)
	$r''_6(0)$ (4.743411905896337, -3.811925035449615)

instances this choice transpired to be very conservative—considerably larger steps could have been safely used. However, we did find that large steps ($\Delta\lambda > 0.1$) could incur “jumps” between distinct solution paths during the predictor–corrector steps, which meant that not all of the distinct interpolants would be generated upon commencing with distinct solutions of the “start” system (35).

The well–conditioned nature of the system (37) allowed the use of a tight tolerance on the norm (41) that governs the Newton–Raphson correction iterations. We consistently employed 10^{-12} for this tolerance value without encountering convergence difficulties. Note that the transformation of the “initial” system (35) into the “desired” system (33) through the homotopy (37) is not, in general, a monotone process: Figure 1 shows the variation of norms analogous to (41) for the \mathbf{f} and \mathbf{g} functions in a typical instance.

We have observed empirically that, among the 2^{N+1} distinct C^2 PH splines interpolating prescribed data (or 2^{N-1} for the cubic end–span case), all but one of these loci exhibits undesirable loops that do not agree with the intuitive “shape” of that

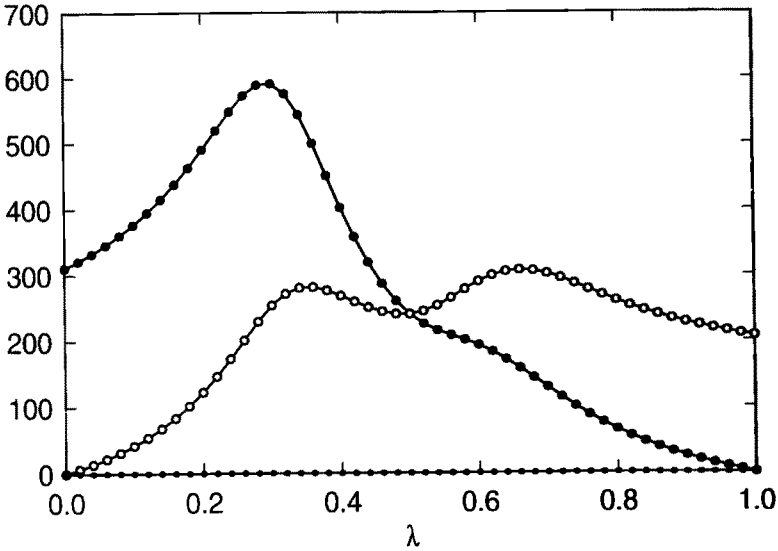


Figure 1. Variation of the f , g , and h function norms (solid dots, open dots, and smaller solid dots) during 50 equal steps of the continuation method.

data. The remaining “good” PH spline, however, is invariably a smoother or *fairer* locus than the ordinary cubic spline.

To quantify these perceptions, two measures are employed to identify the “good” PH spline: the *absolute rotation index* and the *energy integral*

$$R_{\text{abs}} = \frac{1}{2\pi} \int |\kappa| ds \quad \text{and} \quad E = \int \kappa^2 ds, \tag{53}$$

κ and s being the curvature and arc length. For PH splines, these integrals may be reduced analytically, as described in [7] and [8]. For the ordinary cubic spline, however, numerical quadrature must be employed.

In all examples we have tested, the PH spline with the minimum R_{abs} proved to also have the minimum E . We do not, however, believe that this is necessarily always the case, and the user may be called upon to judge as to which of these two quantities identifies the “better” interpolant—see [7] for an example in the simpler context of two-point Hermite interpolation.

Example 1

We interpolate the points $\mathbf{p}_0 = (-2.1, 1.8)$, $\mathbf{p}_1 = (-3.1, 0.0)$, $\mathbf{p}_2 = (-0.3, -0.8)$, $\mathbf{p}_3 = (0.7, 2.2)$, $\mathbf{p}_4 = (3.4, 0.5)$, $\mathbf{p}_5 = (1.1, -0.6)$, $\mathbf{p}_6 = (2.3, -2.4)$ by a C^2 PH-quintic spline with cubic end spans. Among the 32 distinct solutions, the “good” PH spline, as compared in figure 2 with the corresponding “ordinary” cubic spline using quadratic end spans, evidently has a much more pleasing shape. Inspection of the curvature profiles, shown in figure 3, corroborates this claim.

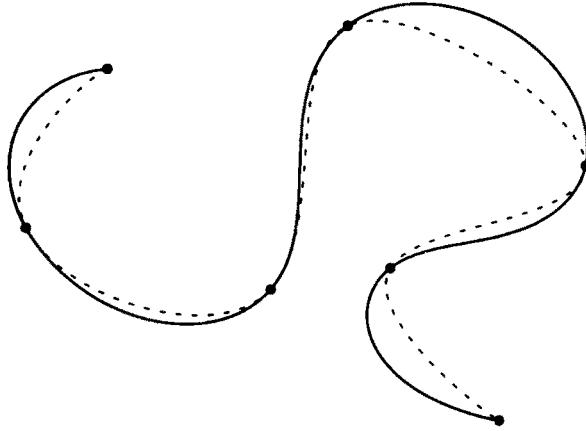


Figure 2. Comparison of PH quintic spline with cubic end-spans (solid) and “ordinary” cubic spline with quadratic end-spans (dotted) for Example 1.

Correspondingly, the PH spline has energy $E = 9.39$, while $E = 13.40$ for the ordinary cubic spline.

We use this example to illustrate the remarkable level of accuracy that can be achieved in the construction of PH splines. Once $\mathbf{a}_0, \dots, \mathbf{a}_N$ and $\mathbf{b}_1, \dots, \mathbf{b}_N$ have been computed by the homotopy method, we make the assignments (12) and compute the control points (6) for each PH quintic span $k = 1, \dots, N$ using equations (4). These control points are then used to test continuity of the left- and

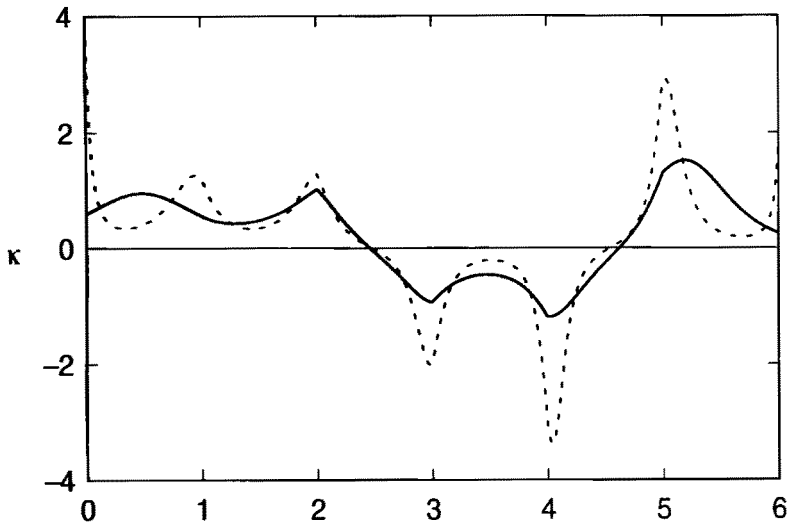


Figure 3. Comparison of curvature profiles for the PH quintic spline (solid) and the “ordinary” cubic spline (dotted) interpolating the data of Example 1.

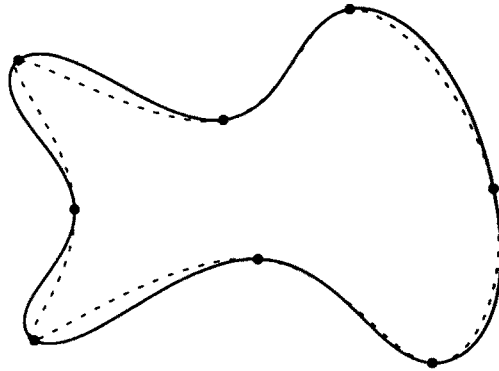


Figure 4. Comparison of PH quintic (solid) and “ordinary” cubic (dotted) periodic spline interpolants for the point data of Example 2.

right-derivatives at each node, using the formulae

$$\begin{aligned} \mathbf{r}'_k(1) &= 5(\mathbf{z}_{k,5} - \mathbf{z}_{k,4}), \\ \mathbf{r}'_{k+1}(0) &= 5(\mathbf{z}_{k+1,1} - \mathbf{z}_{k+1,0}), \\ \mathbf{r}''_k(1) &= 20(\mathbf{z}_{k,5} - 2\mathbf{z}_{k,4} + \mathbf{z}_{k,3}), \\ \mathbf{r}''_{k+1}(0) &= 20(\mathbf{z}_{k+1,2} - 2\mathbf{z}_{k+1,1} + \mathbf{z}_{k+1,0}), \end{aligned}$$

for $k = 1, \dots, N - 1$. The results are enumerated in table 1—it can be seen that, despite the “numerical” nature of the solution procedure, the nodal derivatives agree to ~ 15 significant digits in all instances. For standard double-precision arithmetic, with rounding and a mantissa of $d = 53$ bits, the *machine unit* is

$$\eta = \frac{1}{2} 2^{-(d-1)} \approx 1.1 \times 10^{-16},$$

and hence the computed solutions are accurate to within about one decimal digit of the theoretical limit imposed by the floating-point number system!

As another measure of accuracy for the constructed PH splines, we examine the magnitude of the “gaps”

$$\Delta \mathbf{z}_k = \mathbf{z}_{k+1,0} - \mathbf{z}_{k,5} \tag{54}$$

for $k = 1, \dots, N - 1$ that arise between consecutive arcs owing to the fact that the final control point $\mathbf{z}_{k,5}$ of arc k , as given by expressions (4), need not coincide precisely with the initial control point $\mathbf{z}_{k+1,0}$ of arc $k + 1$. Again, we see from table 1 that the effects of numerical errors are extremely subdued, and these gaps are consequently insignificant.

The data presented in Table 1 are by no means “special”—they are, in fact, representative of what we have observed in numerous examples.

Example 2

We wish to construct periodic curves that interpolate the point data $\mathbf{p}_0 = (2.3, -0.2)$, $\mathbf{p}_1 = (1.0, 1.5)$, $\mathbf{p}_2 = (-0.2, 0.5)$, $\mathbf{p}_3 = (-2.1, 1.1)$, $\mathbf{p}_4 = (-1.6, -0.3)$,

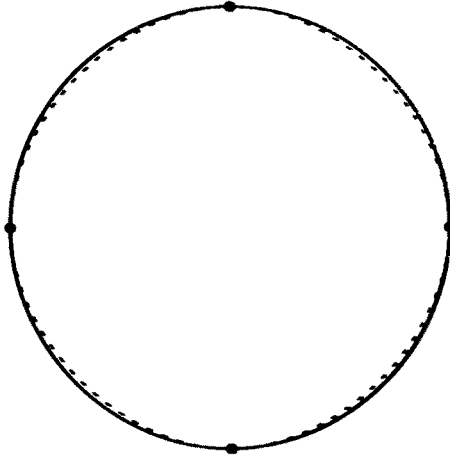


Figure 5. Comparison of periodic PH quintic spline (solid) and “ordinary” cubic spline (dotted) approximations to the unit circle—see Example 3.

$\mathbf{p}_5 = (-2.0, -1.5)$, $\mathbf{p}_6 = (0.1, -0.8)$, $\mathbf{p}_7 = (1.7, -1.8)$, and $\mathbf{p}_8 = (2.3, -0.2) = \mathbf{p}_0$. The resulting “good” PH spline is compared with the “ordinary” cubic spline in figure 4. The discrepancy between the curvature distributions is even more pronounced than in Example 1—the ordinary cubic spline has a maximum curvature more than 10 times greater than that of the PH spline! This difference is apparent in the curve energies: $E = 28.0$ for the PH spline, whereas $E = 148.5$ for the ordinary cubic spline.

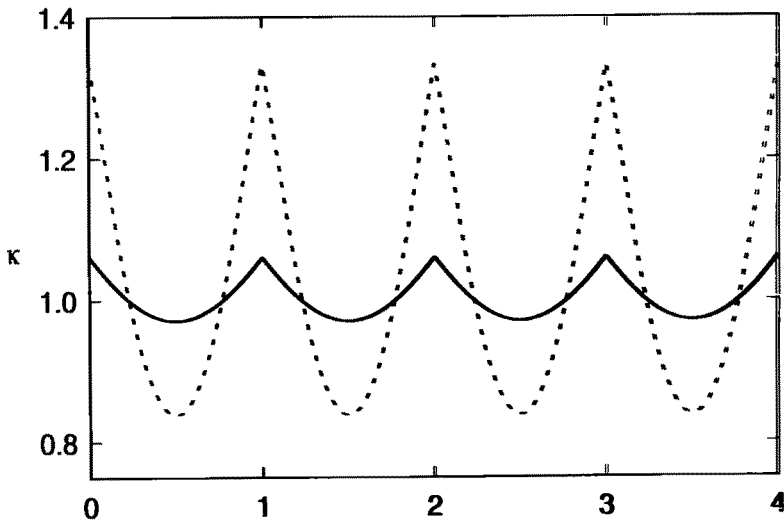


Figure 6. Curvature profiles of the periodic PH quintic (solid) and “ordinary” cubic (dotted) spline approximations to the unit circle—see Example 3.

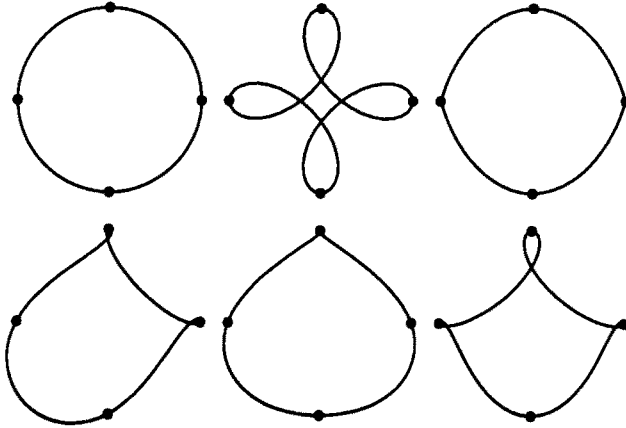


Figure 7. The six different shapes among the 32 distinct PH quintic splines that interpolate four equally-spaced points on the unit circle (see Example 3).

Example 3

Here we desire a periodic curve that interpolates four equally-spaced points, $\mathbf{p}_0 = (1, 0)$, $\mathbf{p}_1 = (0, 1)$, $\mathbf{p}_2 = (-1, 0)$, $\mathbf{p}_3 = (0, -1)$, and $\mathbf{p}_4 = (1, 0) = \mathbf{p}_0$, on the unit circle. The resulting PH spline is compared with the “ordinary” periodic cubic spline in figure 5. The corresponding curvature plots in figure 6 indicate that the former is a much better approximation of the unit circle than the latter (for the ordinary cubic spline, the curvature varies from a minimum of 0.84 to a maximum of 1.33, while that of the PH spline has a much narrower range—from 0.97 to 1.06). The energy integral clearly has the value $E = 2\pi$ for an exact circle: for the PH spline, we have $E/2\pi = 1.0034$, whereas the ordinary cubic spline gives $E/2\pi = 1.0376$.

The present example is an interesting context in which to illustrate the multiplicity of solutions to the problem of PH spline interpolation. Although, for $N = 4$ and periodic boundary conditions, there are 32 distinct solutions, only six of these are actually of different shape (see figure 7) owing to the symmetry of the data. Other than the “good” solution in the top left-hand corner, all the interpolants exhibit loops—in some cases, they are so “tight” as to be virtually hidden by the dots representing the interpolation points. Correspondingly, the bending energies and absolute rotation indices are much higher than that of the “good” solution.

6. Closure

We have developed a precise formulation for the problem of interpolation by C^2 PH splines, using complex representations, and we have shown that it yields highly accurate solutions by means of standard numerical methods. The formulation is simple to implement, and admits flexibility in terms of spline end conditions. Empirical results indicate that PH splines have exceptionally attractive shape

properties compared with the “ordinary” cubic splines, a fact that has previously been noted [7] in the context of individual Hermite arcs. Furthermore, the arc lengths and offset curves of the PH splines admit exact representations.

Perhaps the greatest shortcoming of the method, as it currently stands, is the need to sift through 2^{N+1} solutions in order to select the sole “good” interpolant. In contexts where a high degree of interactivity is required, and where N is large, this burden may prove intolerable. We remain hopeful that the theory will further unfold to a point where this “wasteful” aspect may be obviated. In the interim, in such contexts where speed is not of paramount concern and the superior shape properties or exact arc lengths and offsets of PH curves are highly valued, we believe the formulation is ready for practical use and thus worthy of a permanent place [12] among the plethora of available curve interpolation/design schemes.

Several important extensions of the C^2 PH spline method described herein come immediately to mind. For brevity, we only mention two here:

- As with “ordinary” splines, it would be useful to extend the formulation to accommodate the case of *non-uniform knots*—which are better suited to interpolating unevenly-spaced sequences of points.
- Unlike the “ordinary” splines, extending the method to space curves is *not* a trivial matter—the complex representation, which was instrumental in simplifying the defining equations for *planar* PH splines, has no natural generalization to the three-dimensional case.

We hope to turn our attention to these issues in due course.

References

- [1] E.L. Allgower and K. Georg, *Numerical Continuation Methods: An Introduction* (Springer, Berlin, 1990).
- [2] E.L. Allgower and K. Georg, Continuation and path following, *Acta Numerica* (1993) 1–64.
- [3] J.F. Canny, *The Complexity of Robot Motion Planning* (MIT Press, Cambridge, MA, 1988).
- [4] G. Dahlquist and A. Björck, *Numerical Methods* (Prentice-Hall, Englewood Cliffs, NJ, 1974).
- [5] R.T. Farouki, Pythagorean-hodograph curves in practical use, in *Geometry Processing for Design and Manufacturing* (R.E. Barnhill, ed.), (SIAM, Philadelphia, 1992) 3–33.
- [6] R.T. Farouki, The conformal map $z \rightarrow z^2$ of the hodograph plane, *Comput. Aided Geom. Design* 11 (1994) 363–390.
- [7] R.T. Farouki, The elastic bending energy of Pythagorean-hodograph curves, *Comput. Aided Geom. Design* 13 (1996) 227–241.
- [8] R.T. Farouki and C.A. Neff, Hermite interpolation by Pythagorean-hodograph quintics, *Math. Comp.* 64 (1995) 1589–1609.
- [9] R.T. Farouki and T. Sakkalis, Pythagorean hodographs, *IBM J. Res. Develop.* 34 (1990) 736–752.
- [10] W. Gröbner, *Algebraische Geometrie I* (Bibliographisches Institut, Mannheim, 1968).
- [11] W. Gröbner, *Algebraische Geometrie II* (Bibliographisches Institut, Mannheim, 1970).
- [12] T. Lyche, There’s always more room in the “Spline Zoo,” personal communication (1989).

- [13] A.P. Morgan, *Solving Polynomial Systems Using Continuation for Engineering and Scientific Problems* (Prentice–Hall, Englewood Cliffs, NJ, 1987).
- [14] A.P. Morgan and A. Sommese, A homotopy for solving general polynomial systems that respects m -homogeneous structures, *Appl. Math. Comp.* 24 (1987) 101–113.
- [15] A.P. Morgan and A. Sommese, Computing all solutions to polynomial systems using homotopy continuation, *Appl. Math. Comp.* 24 (1987) 115–138.
- [16] L.T. Watson, S.C. Billups, and A.P. Morgan, ALGORITHM 652 HOMPACK: A suite of codes for globally convergent homotopy algorithms, *ACM Trans. Math. Software* 13 (1987) 281–310.
- [17] W. Zulehner, A simple homotopy method for determining all isolated solutions to polynomial systems, *Math. Comp.* 181 (1988) 167–177.