Planar Curve Interpolation by Piecewise Conics of Arbitrary Type

Robert Schaback

Abstract: Five points in general position in $\mathbb{R}^2$ always lie on a unique conic, and three points plus two tangents also have a unique interpolating conic, the type of which depends on the data. These well–known facts from projective geometry are generalized: an odd number $2n+1 \geq 5$ of points in $\mathbb{R}^2$, if they can be interpolated at all by a smooth curve with nonvanishing curvature, will have a unique GC^2 interpolant consisting of pieces of conics of varying type. This interpolation process reproduces conics of arbitrary type and preserves strict convexity. Under weak additional assumptions its approximation order is $\mathcal{O}(h^5)$, where h is the maximal distance of adjacent data points $f(t_i)$ sampled from a smooth and regular planar curve f with nonvanishing curvature. Two algorithms for the construction of the interpolant are suggested, and some examples are presented.

Keywords: Geometric rational curve interpolation, shape preservation, convexity, parametric splines, conics, approximation order.

MSC Classification: 41A05, 41A15, 41A20, 65D05, 65D07, 65D17, 68U07

1 Introduction

Geometric two-point Hermite interpolation of planar curve data by a GC^2 piecewise cubic polynomial interpolant was considered by deBoor, Höllig, and Sabin [2]. The approximation order of their interpolation process is $\mathcal{O}(h^6)$ for $h \to 0$ with respect to the maximal distance h of adjacent data points, but the interpolant does not always exist. The method of Goodman, Ong, and Unsworth ([4], [5], [6]) avoids this drawback by using rational cubics and is tailored to reproduce arcs of circles. However, its approximation order is not known but does not exceed 4, as can be shown by looking at special cases. For planar Lagrange data, a GC^2 and $\mathcal{O}(h^4)$ interpolant to convex data, using piecewise quadratic polynomials, was given in [9] with extensions in [10] and [11]. The paper [12] contains GC^{k-1} methods of order $\mathcal{O}(h^{2k})$ at the price of using piecewise polynomials of degree $2k+1$. Because the high degree causes unwanted wiggles in curvature plots, one should keep the degree of interpolating curves as small as possible. Furthermore, these high-order methods are not generally convexity-preserving. Therefore one should try to achieve the highest possible approximation order using local pieces of low degree without sacrificing shape-preserving properties. An overview of rational geometric interpolation schemes is provided by [13], while non-interpolating rational spline construction methods are given by Boehm [1] and Farin [3]. A cubic rational GC^2 spline interpolant for space curves was designed by Höllig [7].

[0] File: ~/tex/fertig/rat/trapp.tex, TEXed on August 31, 1992, Status : Revised for Constructive Approximation

The goal of this paper is to provide a convexity-preserving method for curves in $I\!R^2$ that uses piecewise quadratic rational functions, reproduces arbitrary conics and has approximation order $\mathcal{O}(h^5)$. Specifically, we assume an odd number $2n+1 \geq 3$ of data points $y_0, y_1, \ldots, y_{2n}$ in $I\!R^2$ to be given. Then we want to interpolate the data by a geometrically C^2 curve, consisting of pieces

$$r_j(t) := \frac{\sum_{i=0}^{2} b_{2j+i}\beta_i(t)w_{ij}}{\sum_{i=0}^{2} \beta_i(t)w_{ij}}, \quad t \in [0,1], \ 0 \leq j \leq n-1 \tag{1.1}$$

of rational quadratic curves in Bernstein–Bézier representation (see e.g.: [3], [8]), where

$$\beta_i(t) = \binom{2}{i}(1-t)^{2-i}t^i, \quad i = 0,1,2$$

are the quadratic Bernstein polynomials and

$$w_{ij} > 0, \quad 0 \leq i \leq 2, \ 0 \leq j \leq n-1$$

are weights. The control points b_i for $i = 0, 1, \ldots, 2n$ are to be constructed such that r_j interpolates the data $y_{2j}, y_{2j+1}, y_{2j+2}$ on $[0,1]$ and a geometrically C^2 continuous curve results, when $r_0, \ldots, r_{n-1}$ are joined together at the data points $y_2, y_4, \ldots, y_{2n-2}$. This immediately implies

$$b_{2j} = y_{2j}, \quad 0 \leq j \leq n$$

and leaves the construction of the "interior" control points b_{2j+1} for $0 \leq j \leq n-1$ open. In the sequel we avoid certain degeneracies by assuming that for each j the data points y_{2j}, y_{2j+1}, y_{2j+2} are mutually different. A somewhat stronger assumption will be made in section 3.

Because $y_{2j+1} \in I\!R^2$ can be interpolated at an arbitrary parameter value, there is only one additional scalar condition in each piece $r_j(t)$ to be satisfied. Geometric C^2 continuity imposes two scalar conditions at each junction. There are three degrees of freedom per interval, because $b_{2j+1} \in I\!R^2$ and (up to normalization) essentially one weight parameter can be chosen (see [3], [8]). This leaves two degrees of freedom for additional boundary conditions. These are imposed by prescribing tangents at y_0 and y_{2n}. We later comment on construction methods for these tangents, if they are not given right from the start.

Necessarily, our GC^2 interpolant, it it exists, must have a curvature which vanishes everywhere or nowhere, because each rational piece has this property. Thus our interpolation problem can be solvable only if the data either lie on a straight line or are "strictly convex" in the sense that there is some smooth and regular interpolant to the data with nonvanishing curvature. We shall assume the latter, and then our interpolation process will preserve "strict convexity", because the curvature of our interpolant will not vanish.

2 Local GC^1 Hermite interpolation

This section serves as a preparation for later investigations and treats the much simpler problem of local GC^1 two-point Hermite interpolation. Here, an additional tangent direction is prescribed at every even–numbered data point. It is a repeated instance of the special case $n = 1$ of the problem posed in the previous section, and can be interpreted as a geometric three-point Hermite-Birkhoff interpolation scheme.

To interpolate three data points y_0, y_1, y_2 in $I\!R^2$ and two tangent directions r_0 and r_2 at y_0 and y_2, respectively, by a nondegenerate conic, we again have to assume that the data can be interpolated by a smooth and regular curve with nonvanishing curvature. Thus y_1 should lie inside the triangle defined by y_0, y_2 and the tangents at y_0, y_2 with intersection point $Q = b_1$. Note that b_1 will be the interior control point of (1.1) for $n = 1$ (see Figure 1).

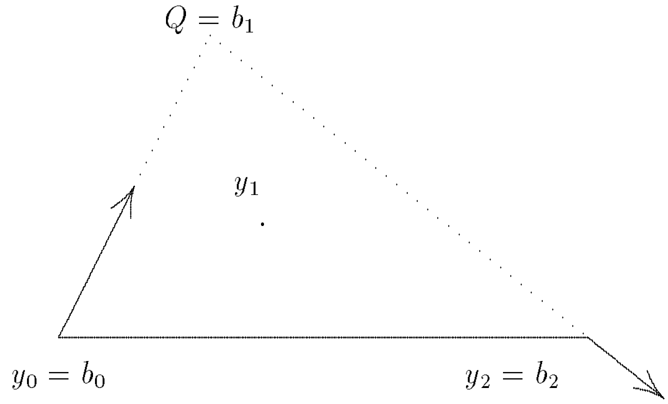

Figure 1: Basic triangle for GC^1 interpolation

It is well known from projective geometry that this problem is uniquely solvable: three points and two tangents uniquely define a conic. The classical construction of an additional point of the conic proceeds as follows:

Take an arbitrary line L emanating from y_0 and intersecting Qy_2. Let R be the intersection of lines L and y_1y_2, and construct S as the intersection of QR with y_0y_1. Then the intersection P of lines Sy_2 and L is the unique intersection point of L and the conic within the triangle.

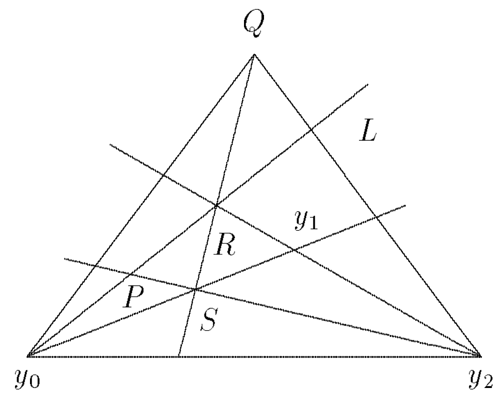

Figure 2: Interpolation for 3 points and two tangents

To derive a direct formula for the solution we observe that there are unique positive real numbers α_0, α_2 with

$$y_1 - b_1 = \alpha_0(b_0 - b_1) + \alpha_2(b_2 - b_1),$$

provided by barycentric coordinates of y_1 with respect to b_0, b_2 and b_1. From the standard representation by Bernstein polynomials we can write any value $r(t)$ of a rational quadratic as

$$r(t) - b_1 = \frac{\beta_0(t)(b_0 - b_1) + \beta_2(t)(b_2 - b_1)}{\beta_0(t) + \beta_1(t)w_1 + \beta_2(t)},$$

where we used $w_0 = w_2 = 1$ without loss of generality (see [8]). If r interpolates y_1 at the parameter t_1 we get the system

$$\begin{aligned} \alpha_0 &= \frac{\beta_0(t_1)}{\beta_0(t_1) + \beta_1(t_1)w_1 + \beta_2(t_1)}, \\ \alpha_2 &= \frac{\beta_2(t_1)}{\beta_0(t_1) + \beta_1(t_1)w_1 + \beta_2(t_1)} \end{aligned}$$

which is uniquely solved by

$$t_1 = (1 + \sqrt{\alpha_0/\alpha_2})^{-1} \tag{2.1}$$

and

$$w_1 = \frac{1}{2} \frac{1 - \alpha_0 - \alpha_2}{\sqrt{\alpha_0\alpha_2}}. \tag{2.2}$$

This yields a local GC^1 Hermite interpolation method for piecewise conics of arbitrary type which is convexity preserving and reproduces all conics. Its complexity grows linearly with the number of data points.

3 Basic Equations for GC^2 Interpolation

For a general rational quadratic curve in Bernstein–Bézier representation with control points b_0, b_1, b_2 and weights w_0, w_1, w_2, we consider arbitrary partitions $\{i, j, k\}$ of $\{0, 1, 2\}$ and let $h_k = \|b_i - b_j\|_2$ be the distance between points b_i and b_j, while c_k is the distance of b_k to the line L_k through b_i and b_j. Then the curvature at b_0 is ([1], p.71)

$$\kappa_0 = \frac{1}{2} \frac{c_2}{h_2^2} \frac{w_0 w_2}{w_1^2}.$$

Let $r = r(t)$ be an arbitrary point on the conic. The areas A_i of the triangles r, b_j, b_k satisfy

$$\frac{w_0 w_2}{4w_1^2} = \frac{A_0 A_2}{A_1^2}$$

(see [8], p. 10). If q_i is the distance of r to the line L_i, then $A_i = \frac{1}{2}\, q_i \cdot h_i$ and

$$\frac{w_0 w_2}{4w_1^2} = \frac{q_0 h_0 q_2 h_2}{q_1^2 h_1^2}.$$

This gives

$$\kappa_0 = 2\, \frac{c_2}{h_2^2} \frac{q_0 h_0 q_2 h_2}{q_1^2 h_1^2}, \quad \kappa_2 = 2\, \frac{c_0}{h_0^2} \frac{q_0 h_0 q_2 h_2}{q_1^2 h_1^2}$$

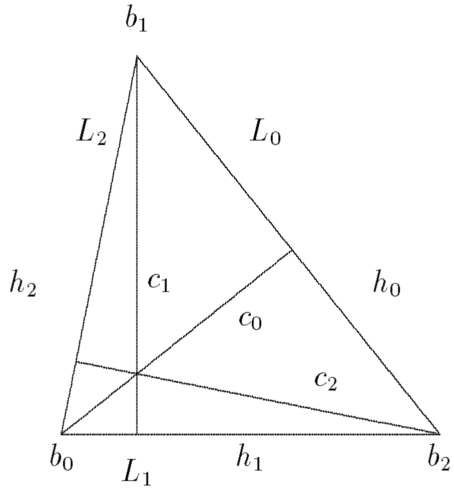

Figure 3: Data for curvature calculation

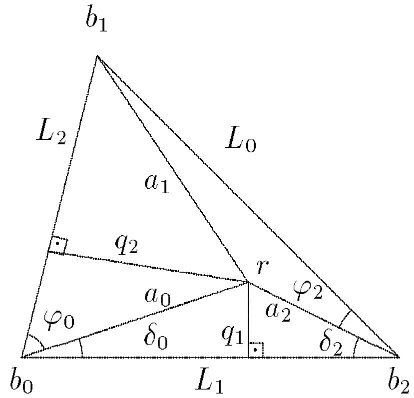

Figure 4: Local angles

for the curvatures κ_0 at b_0 and κ_2 at b_2. If b_0, r, and b_2 are fixed, everything can be expressed by the angles φ_0, δ_0 and φ_2, δ_2 in Figure 4. We introduce $a_i = \|r - b_i\|_2$ and find

$$
\begin{aligned}
\frac{q_2}{a_0} &= \sin\varphi_0 & \frac{q_0}{a_2} &= \sin\varphi_2 \\
\frac{q_1}{a_0} &= \sin\delta_0 & \frac{q_1}{a_2} &= \sin\delta_2 \\
\frac{c_1}{h_2} &= \sin(\varphi_0+\delta_0) = \frac{c_2}{h_1}, & \frac{c_1}{h_0} &= \sin(\varphi_2+\delta_2) = \frac{c_0}{h_1},
\end{aligned}
$$

the last two equations following from Figure 3. This implies

$$
\begin{aligned}
\kappa_0 &= \frac{2\sin(\varphi_0+\delta_0)}{h_2} \cdot \frac{h_0 \sin\varphi_2 \sin\varphi_0}{h_1 \sin\delta_0 \sin\delta_2} \\
&= \frac{2\sin(\varphi_0+\delta_0)\sin\varphi_2\sin\varphi_0}{\sin\delta_0\sin\delta_2} \cdot \frac{\sin(\varphi_0+\delta_0)}{h_1\sin(\varphi_2+\delta_2)} \qquad (3.1) \\
&= \frac{2\sin\varphi_0 \cdot \sin\varphi_2}{h_1 \sin\delta_0 \cdot \sin\delta_2} \, \frac{\sin^2(\varphi_0+\delta_0)}{\sin(\varphi_2+\delta_2)},
\end{aligned}
$$

and similarly

$$\kappa_2 = \frac{2}{h_1} \, \frac{\sin\varphi_0 \cdot \sin\varphi_2}{\sin\delta_0 \sin\delta_2} \, \frac{\sin^2(\varphi_2+\delta_2)}{\sin(\varphi_0+\delta_0)}. \tag{3.2}$$

Furthermore,

$$\kappa_2 = \kappa_0 \left(\frac{\sin(\varphi_2+\delta_2)}{\sin(\varphi_0+\delta_0)} \right)^3 . \tag{3.3}$$

Before we proceed, we remark that it is not generally possible to prescribe three points and two curvature values to define a conic, in spite of the analogous fact that three points with two tangent directions uniquely determine a conic. In fact, if κ_0 and κ_2 were prescribed together with b_0, r, and b_1, there are two equations (3.2) and (3.3) for the two unknowns φ_0 and φ_2. However, these equations are not necessarily solvable, as can be seen in the special symmetric case $\kappa = \kappa_0 = \kappa_2$, $\delta = \delta_0 = \delta_2$, where (3.3) implies $\varphi_0 = \varphi_2$ and (3.2) requires

$$\kappa = \frac{2}{h_1} \, \frac{\sin^2\varphi}{\sin^2\delta} \, \sin(\varphi+\delta),$$

which is an unsolvable equation in case

$$\kappa \, h_1 \sin^2\delta > 2,$$

for example. Note that this negative result corresponds to the similar one in [2].

We now proceed to use (3.1) and (3.2) to express curvature continuity around a point y_{2i} for the GC^2 interpolation problem posed in the introduction. In order to let the geometric construction of Fig. 5 be well–defined, we assume the restrictions

$$\begin{array}{ccccccccl} 0 & < & \delta_i^+, & & 0 & < & \delta_i^-, & 1 \le i \le n, & \\ 0 & < & \delta_i^+ + \delta_i^- & < & \gamma_i & < & \pi, & 1 \le i \le n-1 & \\ 0 & < & \varphi_0^+ & < & \varphi_0^+ + \delta_0^+ & < & \pi, & & \\ 0 & < & \varphi_n^- & < & \varphi_n^- + \delta_n^- & < & \pi, & & \end{array} \tag{3.4}$$

on the given data, where γ_i is the angle between the chords $y_{2i-2}y_{2i}$ and $y_{2i}y_{2i+2}$. Furthermore, the varying angles φ_i^- and φ_i^+ should satisfy

$$\varphi_i^-, \varphi_i^+ > 0, \quad \varphi_i^+ + \varphi_i^- + \delta_i^- + \delta_i^+ = \gamma_i, \quad 1 \le i \le n-1.$$

Note that these restrictions are always satisfied if the data form a sufficiently dense sample from a smooth regular curve with nonvanishing curvature.

Curvature continuity $\kappa_i^- = \kappa_i^+$ at y_{2i} in Fig. 5 implies

$$\begin{aligned} \kappa_i^- &= \frac{2}{h_{i-1}} \, \frac{\sin\varphi_{i-1}^+ \sin\varphi_i^-}{\sin\delta_{i-1}^+ \sin\delta_i^-} \, \frac{\sin^2(\varphi_i^- + \delta_i^-)}{\sin(\varphi_{i-1}^+ + \delta_{i-1}^+)} \\ = \kappa_i^+ &= \frac{2}{h_i} \, \frac{\sin\varphi_i^+ \sin\varphi_{i+1}^-}{\sin\delta_i^+ \sin\delta_{i+1}^-} \, \frac{\sin^2(\varphi_i^+ + \delta_i^+)}{\sin(\varphi_{i+1}^- + \delta_{i+1}^-)}, \quad 1 \le i \le n-1 \end{aligned}$$

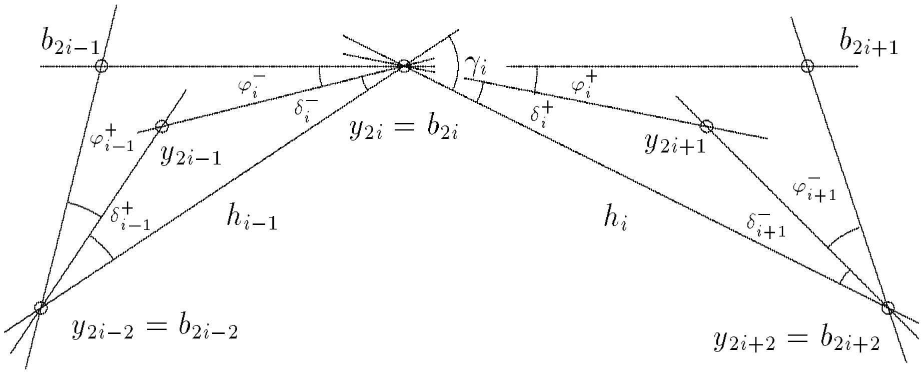

Figure 5: Local angles around y_{2i}

and, with $\varphi_i^+ + \delta_i^+ + \varphi_i^- + \delta_i^- = \gamma_i$, $\varphi_i := \varphi_i^+$ we get

$$\begin{aligned} &\frac{2}{h_{i-1}} \frac{\sin \varphi_{i-1} \sin(\gamma_i - \delta_i^- - \delta_i^+ - \varphi_i)}{\sin \delta_{i-1}^+ \sin \delta_i^-} \cdot \frac{\sin^2(\gamma_i - \varphi_i - \delta_i^+)}{\sin(\varphi_{i-1} + \delta_{i-1}^+)} \\ = &\frac{2}{h_i} \frac{\sin \varphi_i \sin(\gamma_{i+1} - \delta_{i+1}^- - \delta_{i+1}^+ - \varphi_{i+1})}{\sin \delta_i^+ \sin \delta_{i+1}^-} \cdot \frac{\sin^2(\varphi_i + \delta_i^+)}{\sin(\gamma_{i+1} - \varphi_{i+1} - \delta_{i+1}^+)}. \end{aligned} \tag{3.5}$$

This is a nonlinear system of equations in unknowns $\varphi_1, \ldots, \varphi_{n-1}$. Two boundary tangents must be prescribed by fixing φ_0^+ and φ_n^-, and to make the system (3.5) formally valid for $i = 1$ and $i = n - 1$, we define

$$\varphi_0^- = \delta_0^- = 0, \quad \varphi_n^+ = \delta_n^+ = 0, \quad \gamma_0 = \varphi_0^+ + \delta_0^+, \quad \gamma_n = \varphi_n^- + \delta_n^-.$$

The Jacobian of the system is tridiagonal.

4 Existence of a GC^2 interpolant

Our main result of this section will be

Theorem 4.1 *Under the conditions* (3.4) *there is a solution to the interpolation problem.*

For uniqueness we require in addition that

$$\gamma_i < \frac{\pi}{2}, \ 1 \le i \le n \tag{4.1}$$

to get

Theorem 4.2 *The conditions* (3.4) *and* (4.1) *imply uniqueness of the interpolant.*

The rest of this section is devoted to the proof of Theorems 4.1 and 4.2. As a preparation, we state the monotonicity of

$$\sin \varphi \cdot \sin^2(\varphi + \alpha)$$

in both variables φ, α within

$$0 < \varphi < \frac{\pi}{2}, \ 0 < \alpha < \frac{\pi}{2}, \quad 0 < \varphi + \alpha < \frac{\pi}{2}.$$

Furthermore,

$$f_\alpha(\varphi) = \frac{\sin\varphi}{\sin(\varphi+\alpha)} \quad \text{with} \quad f'_\alpha(\varphi) = \frac{\sin\alpha}{\sin^2(\varphi+\alpha)}$$

is monotonic in φ for all α with $0<\alpha<\pi$, and covers $(0,\infty)$ when φ varies over $(0,\pi-\alpha)$.

The simultaneous proof of Theorems 4.1 and 4.2 proceeds by inductively checking the range of the nonlinear mapping $(\varphi_{i-1},\varphi_i) \mapsto \varphi_{i+1}$ given by (3.5) when interpreted as an implicit definition of φ_{i+1} for given values of φ_i and φ_{i-1}. Uniqueness will be a consequence of strict monotonicity.

We start with a fixed value of φ_0^+ (given as a boundary condition) and satisfying the assumptions (3.4). The equation $\kappa_1^- = \kappa_1^+$ is

$$\begin{aligned}
\kappa_1^- &= \frac{2}{h_0}\,\frac{\sin\varphi_0^+}{\sin\delta_0^+}\,\frac{\sin\varphi_1^-}{\sin\delta_1^-}\,\frac{\sin^2(\varphi_1^-+\delta_1^-)}{\sin(\varphi_0^++\delta_0^+)}\\
&= \frac{2}{h_1}\,\frac{\sin\varphi_1^+}{\sin\delta_1^+}\,\frac{\sin\varphi_2^-}{\sin\delta_2^-}\,\frac{\sin^2(\varphi_1^++\delta_1^+)}{\sin(\varphi_2^-+\delta_2^-)} = \kappa_1^+,
\end{aligned}$$

and κ_1^- varies in an interval of the form $(0,\hat\kappa_1)$ when φ_1^- varies in $(0,\gamma_1-\delta_1^--\delta_1^+)$. Because of $\varphi_1^-+\varphi_1^+ = \gamma_1-\delta_1^--\delta_1^+$, the variable φ_1^+ varies also in $(0,\gamma_1-\delta_1^+-\delta_1^-)$, but in the opposite direction to φ_1^-. As a function of φ_1^- and $\varphi_1^+ = \gamma_1-\delta_1^+-\delta_1^--\varphi_1^-$, the left–hand–side of the expression

$$\frac{h_1}{h_0}\cdot\frac{\sin\delta_1^+\sin\delta_2^-}{\sin\delta_0^+\sin\delta_1^-}\cdot\frac{\sin\varphi_1^+}{\sin(\varphi_1^++\delta_1^+)}\cdot\frac{\sin\varphi_1^-}{\sin\varphi_1^+}\cdot\frac{\sin^2(\varphi_1^-+\delta_1^-)}{\sin^2(\varphi_1^++\delta_1^+)} = \frac{\sin\varphi_2^-}{\sin(\varphi_2^-+\delta_2^-)} = f_{\delta_2^-}(\varphi_2^-) \tag{4.2}$$

covers all of $(0,\infty)$. Thus, (4.2) is solvable for φ_2^-, and φ_2^- covers $(0,\pi-\delta_2^-)$ when φ_1^- varies in $(0,\gamma_1-\delta_1^+-\delta_1^-)$. Note that all variations are strictly monotonic under the additional assumption (4.1).

Now we have to discard all values of φ_2^- in the interval $[\gamma_2-\delta_2^--\delta_2^+,\pi-\delta_2^-)$, but surely there is an interval of the form $(0,\varepsilon_1)$ with $0<\varepsilon_1<\gamma_1-\delta_1^+-\delta_1^-$ such that φ_2^- varies over all of $(0,\gamma_2-\delta_2^+-\delta_2^-)$ when φ_1^- varies over $(0,\varepsilon_1)$. Furthermore, $\varphi_2^- = \mathcal{O}(\varphi_1^-)$ for $\varphi_1^-\to 0$.

We now proceed inductively for $i\ge 2$ and assume the existence of an interval $(0,\varepsilon_{i-1})$ such that when φ_1^- varies in $(0,\varepsilon_{i-1})$, the variable φ_i^- varies over all of $(0,\gamma_i-\delta_i^+-\delta_i^-)$ and all of the φ_j^- vary in subintervals $(0,\rho_j)$ with $\rho_j<\gamma_j-\delta_j^+-\delta_j^-$, while all equations $\kappa_j^-=\kappa_j^+$ for $1\le j\le i-1$ are satisfied. Furthermore, all φ_j^- for $2\le j\le i$ are supposed to satisfy $\varphi_j^- = \mathcal{O}(\varphi_1^-)$ for $\varphi_1^-\to 0$.

The equation $\kappa_i^+=\kappa_i^-$ is equivalent to

$$\frac{h_i}{h_{i-1}}\cdot\frac{\sin\delta_i^+\sin\delta_{i+1}^-}{\sin\delta_{i-1}^+\sin\delta_i^-}\cdot\frac{\sin\varphi_{i-1}^+}{\sin(\varphi_{i-1}^++\delta_{i-1}^+)}\cdot\frac{\sin\varphi_i^-}{\sin\varphi_i^+}\cdot\frac{\sin^2(\varphi_i^-+\delta_i^-)}{\sin^2(\varphi_i^++\delta_i^+)} = \frac{\sin\varphi_{i+1}^-}{\sin(\varphi_{i+1}^-+\delta_{i+1}^-)} = f_{\delta_{i+1}^-}(\varphi_{i+1}^-), \tag{4.3}$$

and again the left–hand side will cover all of $(0,\infty)$ when φ_i^- and $\varphi_i^+ = \gamma_i-\delta_i^+-\delta_i^--\varphi_i^-$ vary over $(0,\gamma_i-\delta_i^+-\delta_i^-)$ in opposite direction. This allows φ_{i+1}^- to solve (4.3) while varying in $(0,\pi-\delta_{i+1}^-)$, and we have to discard the values of φ_{i+1}^- in $[\gamma_{i+1}-\delta_{i+1}^+-\delta_{i+1}^-,\pi-\delta_{i+1}^-)$ by restricting φ_i^- to some interval $(0, \rho_i)$ with $0 < \rho_i < \gamma_i - \delta_i^+ - \delta_i^-$. If we denote the left–hand side of (4.3) by $g_i(\varphi_i^-)$, we can take the value

$$\rho_i := \inf\{\rho \mid 0 < \rho < \gamma_i - \delta_i^+ - \delta_i^-, \quad g_i(\rho) = f_{\delta_{i+1}^-}(\gamma_{i+1} - \delta_{i+1}^+ - \delta_{i+1}^-)\}$$

for this purpose, making use of $g_i(\varphi_i^-) = \mathcal{O}(\varphi_i^-) = \mathcal{O}(\varphi_1^-)$ for $\varphi_1^- \to 0$ and $f_{\delta_{i+1}^-}(\varphi_{i+1}^-) = \mathcal{O}(\varphi_{i+1}^-)$ for $\varphi_{i+1}^- \to 0$. This, in turn, will require some positive $\varepsilon_i < \varepsilon_{i-1}$ to restrict φ_1^- to an interval $(0, \varepsilon_i)$. Furthermore, $\varphi_{i+1}^- = \mathcal{O}(\varphi_i^-) = \mathcal{O}(\varphi_1^-)$ for $\varphi_1^- \to 0$, and the induction step is complete.

For $i + 1 = n$ we find that φ_n^- covers all of $(0, \pi - \delta_n^-)$ when φ_1^- varies over $(0, \varepsilon_{n-1})$. Therefore, any boundary tangent at the other end of the data set can be prescribed, and there always is a solution.

The assumptions (4.1) guarantee strict monotonicity in each step of the proof. This proves uniqueness of the solution. $\square$

5 Boundary Conditions

As is well known from projective geometry, five given points in general position in a plane will lie on a unique conic. This immediately yields simple and efficient strategies for determining end tangents, which, in addition, will preserve reproduction of conics.

For completeness we include the simple recipe here. If points $y_0, y_1, \ldots, y_4 \in I\!R^2$ are given, we suggest to program a function

$$F : I\!R^2 \times I\!R^2 \times I\!R^2 \times I\!R^2 \to I\!R^2$$

such that $F(a, b, c, d)$ is the intersection point of the lines through a, b, and c, d, respectively. Then the function calls

$$\begin{aligned} a &:= F(y_0, y_2, y_1, y_4) \\ b &:= F(y_0, y_1, y_2, y_4) \\ c &:= F(y_0, y_2, y_3, y_4) \\ d &:= F(y_0, y_3, y_2, y_4) \\ s &:= F(a, b, c, d) \end{aligned}$$

generate a point s such that y_0, s, y_4 are the control points of the interpolating conic in Bernstein-Bézier representation. This immediately yields the tangents to the conic at y_0 and y_4. If the weight w_1 or the parameters of y_1, y_2, y_3 are required, use the GC^1 two–point Hermite construction of section 2. Note that the above construction, when naively programmed in $I\!R^2$ as described here, will fail whenever the given points are not in general position. Going over to projective geometry will remedy the situation somewhat and will produce tangents at y_0 and y_4, even if some of the lines are parallel. However, for sufficiently dense data samples from smooth planar curves with nonvanishing curvature there will be no problems.

This basic construction can be easily adapted to generate convexity-preserving boundary tangents at y_0 and y_{2n} from $y_0, \ldots, y_4$ and $y_{2n-4}, \ldots, y_{2n}$. Then we have a GC^2 Lagrange interpolation method which preserves convexity and all conics. If we use the local five–point construction for estimation of tangents in the GC^1 case, we end up with the same properties.

A given boundary tangent at y_0 can be viewed as a twofold data point; thus it seems reasonable to look for a method that replaces tangent data at y_0 by an additional data point y_{-1}. This strategy is similar to the "not-a-knot" boundary condition for nonparametric spline interpolation. For all conics through y_{-1}, y_0, y_1, y_2 the control point P of the rational Bézier representation in terms of y_{-1}, P, and y_2 must lie on the upper part of the line AB of Figure 6 (compare with Figure 2).

This can be used to relate the angles φ_2^- and φ_{-1}^+, which in turn makes it possible to write the curvature κ_2^- at y_2 as a complicated function of φ_2^- alone. But then the system (3.5) does not require a boundary condition for φ_0^+. An existence proof for the solution can be given by the topological techniques of [9], but we omit the details here, because it is much simpler to apply the previously proposed five-point method.

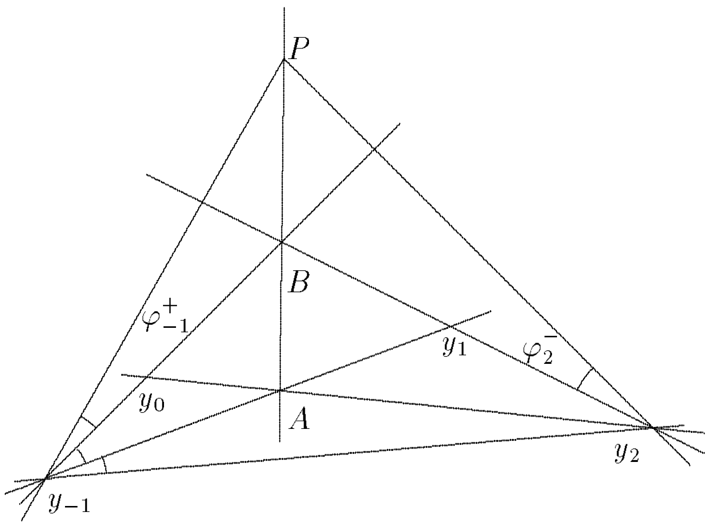

Figure 6: not-a-knot condition

We conclude this section by pointing out that the five–point method can also be used to generate an additional data point that fits in with conic precision, making it possible to assume that the number of data points always is odd, provided that it is at least 5.

6 Convergence

Numerical experiments suggest $\mathcal{O}(h^5)$ accuracy of the interpolation process as described above, if the data $y_i = f(t_i)$, $0 \le i \le 2n$ are sampled from a smooth and regular curve $f : [a, b] \subset \mathbb{R} \to \mathbb{R}^2$ with nonvanishing curvature, where

$$h := \min_{1 \le i \le 2n} (t_i - t_{i-1}), \quad a \le t_0 < t_1 < \ldots < t_{2n} \le b$$

tends to zero. The technique of the $\mathcal{O}(h^4)$ convergence proof of [9] for piecewise quadratic polynomial GC^2 interpolation can be carried over to this situation. It requires a thorough analysis of a scaled version of the underlying nonlinear system and applies a Newton-Kantorovitch type theorem with the following main ingredients:

- The system must behave like $\mathcal{O}(h^3)$ for exact data from f to get $\mathcal{O}(h^5)$ accuracy of the interpolant.
- The original variables should be scaled by a factor with behavior $\mathcal{O}(h^{-1})$.
- The scaled system must have a uniform bound for the inverse of its Jacobian, and
- its second derivatives must be uniformly bounded for $h \to 0$.

The final result will be

Theorem 6.1 *For sufficiently dense samples of data $y_i = f(t_i)$ from a regular and sufficiently smooth curve f with nonvanishing curvature, and for uniformly bounded ratios*

$$0 < \gamma \le \frac{\|f(t_{2i}) - f(t_{2i\pm1})\|}{\|f(t_{2i}) - f(t_{2i\pm2})\|} \le 1-\gamma,\ 1 \le i \le n-1 \tag{6.1}$$

with a fixed value of $\gamma \in (0,1)$, the interpolation process of the preceding sections has $\mathcal{O}(h^5)$ accuracy.

Proof: As in [9], there are some complicated expansions to be calculated. So let $f : [a,b] \longrightarrow I\!R^2$ be a smooth planar curve with positive curvature, parametrized by arclength. We can follow [2] and [9] to assume a local two–dimensional representation of f around a single point by

$$f(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix},\quad f'(t) = \begin{pmatrix} \cos\theta(t) \\ \sin\theta(t) \end{pmatrix}$$

$$\theta(t) = \sum_{i=1}^{k-1} \theta_i t^i + \mathcal{O}(t^k) \tag{6.2}$$

without loss of generality. Here $\theta(t)$ is the angle between tangents of f at $f(0)$ and $f(t)$, while the curvature at $f(t)$ is

$$\kappa(t) = \theta'(t) = \sum_{i=1}^{k-1} i\theta_i t^{i-1} + \mathcal{O}(t^{k-1}).$$

The angle $\alpha(t)$ between the chord $f(t) - f(0)$ and the tangent to f at $f(0)$ can be expanded as

$$\alpha(t) = \frac{1}{2}\theta_1 t + \frac{1}{3}\theta_2 t^2 + \frac{1}{4}\theta_3 t^3 + \frac{72\theta_4 - \theta_1^2\theta_2}{360} t^4 + \frac{120\theta_5 - 4\theta_1\theta_2^2 - 3\theta_1^2\theta_3}{720} t^5 + \mathcal{O}(t^6) \tag{6.3}$$

for $t \to 0$ by a straightforward REDUCE program based on the expansion (6.2) of $\theta(t)$. Here and in the sequel the expansions will be truncated for simplicity of typesetting and reading; it is no major problem to extend them by appropriate calculations with any software system for symbolic computation.

Further analysis of the system (3.5) along the lines of [9] requires the angle

$$\begin{aligned}
\rho(t,u) \quad = \quad & \tfrac{1}{2}\theta_1(t+u) + \tfrac{1}{3}\theta_2(t^2+ut+u^2) + \tfrac{1}{4}\theta_3(t^3+t^2u+tu^2+u^3) + \\
& \frac{72\theta_4 - \theta_1^2\theta_2}{360}(t^4+u^4) + \frac{18\theta_4 + \theta_1^2\theta_2}{90}(t^3u+tu^3) + \frac{12\theta_4 - \theta_1^2\theta_2}{60}t^2u^2 + \\
& \frac{120\theta_5 - 4\theta_1\theta_2^2 - 3\theta_1^2\theta_3}{720}(t^5+u^5) + \frac{40\theta_5 + 4\theta_1\theta_2^2 + 3\theta_1^2\theta_3}{240}(t^4u+tu^4) + \\
& \frac{60\theta_5 - 4\theta_1\theta_2^2 - 3\theta_1^2\theta_3}{360}(t^3u^2+t^2u^3) + \mathcal{O}(t^6) + \mathcal{O}(u^6)
\end{aligned}$$

between the tangent at $f(0)$ and the chord through $f(t)$ and $f(u)$. Note that

$$\rho(t,u) = \left\{ \begin{array}{cc} \alpha(t) & u=0 \\ \alpha(u) & t=0 \\ \theta(t) & u=t \end{array} \right\}, \tag{6.4}$$

and we shall assume $0 < u < t$ throughout. The second part of the system (3.5) can now be rewritten in a simpler form: the correspondences are

$$\begin{array}{rclrclrcl} y_{2i} &=& f(0), & y_{2i+1} &=& f(u), & y_{2i+2} &=& f(t), \\ \varphi_i^+ &=& \alpha(u), & \delta_i^+ &=& \alpha(t)-\alpha(u), & \varphi_i^+ + \delta_i^+ &=& \alpha(t), \\ \varphi_{i+1}^- &=& \theta(t)-\rho(t,u), & \delta_{i+1}^- &=& \rho(t,u)-\alpha(t), & \varphi_{i+1}^- + \delta_{i+1}^- &=& \theta(t)-\alpha(t), \\ h_i &=& c(t), & & & & \|y_{2i+1}-y_{2i}\|_2 &=& c(u), \end{array}$$

where

$$\begin{array}{rcl} c(t) &=& t - \dfrac{\theta_1^2}{24}t^3 - \dfrac{\theta_1\theta_2}{12}t^4 + \dfrac{3\theta_1^4 - 432\theta_1\theta_3 - 256\theta_2^2}{5760}t^5 \\ & & + \dfrac{\theta_1^3\theta_2 - 32\theta_1\theta_4 - 40\theta_2\theta_3}{480}t^6 + \mathcal{O}(t^7) \end{array}$$

is the chordlength $\|f(t)-f(0)\|_2$ as a function of arclength t. After inserting the expansions of the angles in Taylor series of sines, and after expansion of denominators by application of Neumann's series there is a representation

$$\begin{array}{rcl} \kappa^+(t,u) &=& \theta_1 + t^2(\kappa_{2,1}u + \mathcal{O}(u^2)) + u\mathcal{O}(t^3) + \mathcal{O}(t^4) \\ \kappa_{2,1} &=& (18\theta_1^4\theta_2 + 54\theta_1^2\theta_4 - 135\theta_1\theta_2\theta_3 + 80\theta_2^3)/(135\theta_1^2) \end{array} \tag{6.5}$$

for the right-hand part of the curvature in (3.5). Therefore (3.5) vanishes like $\mathcal{O}(h^3)$ for data from a smooth and regular curve. Note that the corresponding system in [9] had a system that vanished like $\mathcal{O}(h^2)$, giving overall $\mathcal{O}(h^4)$ accuracy of the interpolation. No dramatic effect of u being near to 0 or to t arises at this point, because the ratios

$$\begin{array}{rclcl} \dfrac{\varphi_i^+}{\delta_{i+1}^-} &=& \dfrac{\alpha(u)}{\rho(t,u)-\alpha(t)} &=& \dfrac{u\theta_1/2 + u^2\theta_2/3 + \mathcal{O}(u^3)}{u\theta_1/2 + u(t+u)\theta_2/3 + u\mathcal{O}(t^2)} \\ \dfrac{\varphi_{i+1}^-}{\delta_i^+} &=& \dfrac{\theta(t)-\rho(t,u)}{\alpha(t)-\alpha(u)} &=& \\ & & \multicolumn{3}{c}{\dfrac{(t-u)\theta_1/2 + (t-u)(2t+u)\theta_2/3 + (t-u)\mathcal{O}(t^2)}{(t-u)\theta_1/2 + (t-u)(t+u)\theta_2/3 + (t-u)\mathcal{O}(t^2)}} \end{array}$$

tend to 1 even if the denominators vanish (see (6.4) for the other cases). The remaining terms of $\kappa^+(t,u)$ are functions of t alone.

We now proceed to analyze partial derivatives of a properly scaled version of the system (3.5). Expansion (6.5) shows that (3.5) will vanish cubically with $t \to 0$, yielding degenerate Jacobians in the limit. To care for this, we scale the variables in a suitable way, using $\psi_i := \varphi_i/u_i$ instead of $\varphi_i = \varphi_i^+$, where

$$u_i^- := \|y_{2i-1} - y_{2i}\|,\ u_i^+ := \|y_{2i} - y_{2i+1}\|,\ u_i := \min(u_i^+, u_i^-).$$

Note that these numbers are fixed within a single interpolation problem; they can be treated as constants when taking derivatives. Then we have

$$\frac{\partial \kappa^+}{\partial \psi_i} = \frac{u_i}{u}\, u\, \frac{\partial \kappa^+}{\partial \varphi_i}$$

and

$$\frac{\partial \kappa^+}{\partial \psi_{i+1}} = \frac{u_{i+1}}{t-u}\,(t-u)\,\frac{\partial \kappa^+}{\partial \varphi_{i+1}} = -\frac{u_{i+1}}{t-u}\,(t-u)\,\frac{\partial \kappa^+}{\partial \varphi_{i+1}^-}.$$

The latter derivatives can be expanded into

$$\begin{aligned} u\,\frac{\partial \kappa^+}{\partial \varphi_i} &= 2 + 4\frac{u}{t} + \mathcal{O}(u) + \mathcal{O}(t^4) \\ (t-u)\,\frac{\partial \kappa^+}{\partial \varphi_{i+1}^-} &= 2\frac{u}{t} + \mathcal{O}(u). \end{aligned}$$

Due to (6.1) and

$$\begin{aligned} \frac{u_i}{u} &\leq 1 + \mathcal{O}(t), & \frac{u_{i+1}}{t-u} &\leq 1 + \mathcal{O}(t), \\ 2\,\frac{u}{t} &\leq 2 - 2\gamma + \varepsilon, & 2 + 4\gamma - \varepsilon &\leq 2 + 4\,\frac{u}{t} \end{aligned}$$

for any given $\varepsilon > 0$ and sufficiently small $t > u > 0$, the Jacobian J of (3.5) will be strictly diagonally dominant and satisfy

$$\|J\|_\infty \leq 16, \;\; \|J^{-1}\|_\infty \leq \frac{1}{12\gamma}$$

for $t \to 0$. The second derivatives can be shown to be bounded for $t \to 0$, using a similar expansion together with the proposed scaling. The rest of the proof, applying a Kantorovich type convergence theorem for Newton's method, precisely follows the lines of [9]. □

7 Numerical treatment

As in [9], the asymptotic considerations of the convergence analysis suggest a modified Newton–Raphson method with stepsize control for solving the nonlinear system (3.5). The same paper provides convexity-preserving starting values of order $\mathcal{O}(h^3)$, and the technique of [12] yields derivative estimates of arbitrarily high order. The latter are not necessarily convexity–preserving in all cases, but due to their high order they will preserve convexity for sufficiently small h. So there will be no major problems when solving (3.5) with a safeguarded Newton-Raphson technique.

The $- \; + \; -$ sign distribution and the strict diagonal dominance of the tridiagonal Jacobian makes a simple Gauss-Seidel iteration applicable around the solution. This worked rather well in practice, especially when combined with starting values from tangents which were estimated using the five-point method of section 5. The difference between this GC^1 interpolant and the final GC^2 interpolant often is graphically invisible, but of course the curvature plots show the difference. In anticipation of possible numerical problems with the various sines occurring in (3.5), a simple univariate zero–finding procedure with linear convergence was employed to solve (3.5) for φ_i. Repeating this process cyclically over i proved to be efficient enough.

8 Examples

To illustrate the $\mathcal{O}(h^5)$ convergence, we sampled $2n + 1 = 2^{k+1} + 1$ points (plus two boundary tangents) of a 90 degree arc of a logarithmic spiral f with exponent -0.25 at equidistant (non-arclength) parameter values. The rational interpolants to these data were sampled at 513 points at equidistant parameter values of their piecewise rational Bernstein–Bézier representation, giving $2^{8-k} - 1$ interior points in each of the $n - 1 = 2^k - 1$ pieces. To each of these points of the interpolant, the nearest point of f was calculated by a simple minimization routine, yielding the discrete L_∞ errors given in Table 1. In spite of the rather crude technique of error measurement, the results clearly show an error reduction of about 1/32 when doubling the number of data points.

n	L_∞ error
2	$2.212_{10} - 5$
4	$8.739_{10} - 7$
8	$2.689_{10} - 8$
16	$8.494_{10} - 10$
32	$2.660_{10} - 11$

Table 1: Errors of interpolation to logarithmic spiral

Due to the high approximation order, plots of interpolants practically always reproduce the interpolated curve within plot precision, if the data are sampled from a smooth curve. In such cases, curvature plots give much more information; a number of examples of this type can be found in [13], comparing this method with other rational interpolants.

Here we provide results for some coarse data sets which were manually entered without smoothing. Boundary tangents were estimated by the five-point formula of section 5. Of course, data sets with four or less points cannot be handled by our method, and data sets with five points will automatically yield the unique conic that fits through these five points. Thus, at least seven points are needed for nontrivial examples.

Figure 7 contains a plot of the GC^2 interpolant to seven data points in convex position, while the corresponding curvature plot is given in Figure 8. Note that the interpolant consists of only three pieces, as shown in the curvature plot.

The last example, given in Figures 9 and 10, has 19 points in the shape of an @ letter. Here, the curve takes a left turn, making the right part of the curvature plot correspond to the lower part of the curve.

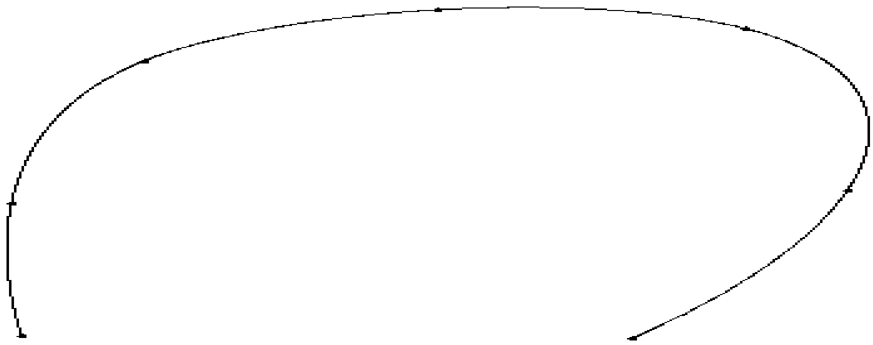

Figure 7: Seven data points with GC^2 interpolant

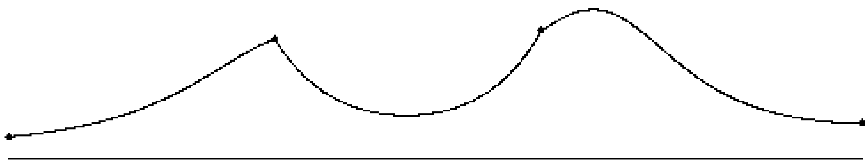

Figure 8: Curvature of GC^2 interpolant

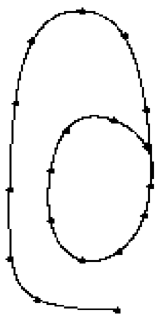

Figure 9: Nineteen data points with GC^2 interpolant

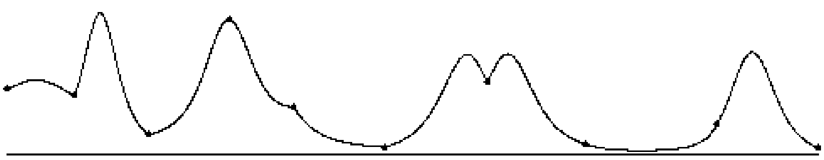

Figure 10: Curvature of GC^2 interpolant

References

[1] Boehm, W., Rational geometric splines, Comput. Aided Geom. Design **4** (1987) 67-78

[2] deBoor, C., Höllig, K., and Sabin, M., High Accuracy Geometric Hermite Interpolation, Computer Aided Geometric Design 4 (1987) 269–278

[3] Farin, G.E. Rational curves and surfaces, *Mathematical Methods in Computer Aided Geometric Design*, Tom Lyche and L. L. Schumaker (eds.), Academic Press, New York, 1989, 215-238

[4] Goodman, T. N. T., Shape preserving interpolation by parametric rational cubic splines, *Numerical Mathematics Singapore 1988*, R. P. Agarwal, Y. M. Chow, and S. J. Wilson (eds.), International Series of Numerical Mathematics, Vol. 86, Birkhäuser, Basel, 1988, 149–158

[5] Goodman, T. N. T., B. H. Ong, K. Unsworth, Constrained interpolation using rational cubic splines, Univ. of Dundee, Computer Sciences Report CS90/01, 1990

[6] Goodman, T. N. T., K. Unsworth, An algorithm for generating shape preserving parametric interpolating curves using rational cubic splines, Univ. of Dundee, Computer Sciences Report CS89/01, 1989

[7] Höllig, K., Algorithms for Rational Spline Curves, ARO-Report 88–1, 1988, 287–300

[8] Lee, E. T., The rational Bézier representation fo conics in Geometric Modeling, *Geometric Modeling: Algorithms and New Trends*, G. Farin (ed.), SIAM, Philadelphia, 1987, 3-19

[9] Schaback, R., Interpolation in $I\!R^2$ by piecewise quadratic visually C^2 Bezier polynomials, Computer Aided Geometric Design 6 (1989) 219–233

[10] Schaback, R., On Global GC^2 Convexity Preserving Interpolation of Planar Curves by Piecewise Bezier Polynomials, in T. Lyche and L.L. Schumaker (eds.): "Mathematical Methods in Computer Aided Geometric Design", Academic Press 1989, 539–547

[11] Schaback, R., Convergence of Planar Curve Interpolation Schemes, in C.K. Chui, L.L. Schumaker and J.D. Ward (eds): "Approximation Theory VI", 1989, 581-584

[12] Schaback. R., Geometric Differentiation and High–Accuracy Curve Interpolation, to appear in S.N. Singh (ed.): Approximation Theory, Splines, and Applications, Proceedings of a NATO-ASI conference, Maratea 1991

[13] Schaback. R., Rational Curve Interpolation to appear in T. Lyche and L.L. Schumaker (eds.): "Mathematical Methods in Computer Aided Geometric Design II", Academic Press 1992

Author's address:

Prof. Dr. R. Schaback
Institut für Numerische und Angewandte Mathematik
Lotzestraße 16–18

D–3400 Göttingen