Interpolation with piecewise quadratic visually C^2 Bézier polynomials

Robert Schaback

Abstract. For data satisfying certain generalized convexity conditions, the existence and uniqueness of a piecewise quadratic curvature continuous parametric spline interpolant in $\mathbb{R}^2$ is proved. Convexity is preserved by the interpolation, and the numerical construction of the interpolant can be carried out efficiently. The interpolation is 4–th order accurate, and a number of examples shows both the applicability and the limitations of this interpolation scheme.

1 Introduction

For given data points

$$f_1, \ldots, f_n, n \in \mathbb{R}^2,\ n \geq 2,\ f_i \neq f_{i+1} \text{ for } 1 \leq i \leq n-1 \tag{1}$$

we want to construct a continuous interpolating curve

$$s : I \to \mathbb{R}^2,\ I \subset \mathbb{R} \text{ a closed interval}$$

whose pieces are of class C^2 and whose tangent direction and curvature vary continuously. This property is called "visual" or "geometric" C^2 continuity and is abbreviated by GC^2 (see e.g. [Boehm et. al., '84] and [Farin, '82]). The interpolant should be convexity preserving in order to avoid unwanted oscillations. Since nonparametric strictly convex or concave C^2 functions have no inflection points and nonvanishing curvature, we introduce a natural common generalization of convexity/concavity to the parametric situation:

Definition 1.1 *A GC^2 curve s in $\mathbb{R}^2$ is (strictly)* **noninflecting**, *if the curvature of s nowhere vanishes.* ■

If the cross–product for planar vectors is defined as

$$x \times y = -x_1 y_2 + x_2 y_1 \text{ for } x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},\ y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

the (signed) curvature of a planar GC^2 curve f is $f'(t) \times f''(t)/\|f'(t)\|^3$ at $f(t)$. The sign of the curvature indicates whether the osculating circle lies to the "left" or "right" of the curve in the sense of a moving observer looking into the direction of the tangential vector. Reversing the parametrization will change the sign of this form of curvature.

The following theorem shows that parametric quadratic splines are appropriate tools for constructing noninflecting curves:

Theorem 1.1 *A visual C^2 parametric quadratic spline curve is strictly noninflecting or a straight line.*

Proof. It is easy to verify that a quadratic polynomial in Bézier form either has identically zero curvature (in case it is a straight line) or has nonvanishing curvature. ■

The sign of the curvature of a quadratic polynomial $s(t)$ in Bézier form with control points b_0, b_1, b_2 in $\mathbb{R}^2$ and

$$s(t) = b_0(1-t)^2 + 2b_1t(1-t) + b_2t^2, \qquad t \in [0,1]$$

is determined by the position of b_1 relative to b_0 and b_2. If b_1 is collinear to b_0 and b_2, the curve is linear. Otherwise, b_1 lies in the "left" or "right" half–space defined by the oriented line $L(b_0, b_2)$, thus determining the sign of the curvature.

2 Solvability conditions

In the following we shall construct the interpolating curve by patching together $n-1$ quadratic polynomials in Bézier form with control points $f_i = b_{2i}, b_{2i+1}, b_{2i+2} = f_{i+1}$ for $1 \le i \le n-1$. The interior control points b_{2i+1} will be determined by continuity requirements only. As a result of the above discussion, we can assume for the rest of this paper that the curve is always "turning to the right". For "left–turning" data the indexing can be reversed to get "right–turning" data (positive curvature). Then the points b_{2i+1} have to be constructed on the "left" side of the oriented lines $L(b_{2i}, b_{2i+2}) = L(f_i, f_{i+1})$, except for the trivial case that all data points f_i are collinear, which we exclude from now on.

Since each quadratic piece has two degrees of freedom and visual C^2 continuity imposes two conditions on each breakpoint $f_2, f_3, \ldots, f_{n-1}$, we have to fix two additional parameters via boundary conditions. This can be done by adding two points $b_0 \neq b_2$ and $b_{2n+2} \neq b_{2n}$ defining tangent directions at b_2 and b_{2n}, respectively. The complete interpolation problem then will be as follows:

> Construct a visually C^2 curve consisting piecewise of quadratic polynomials in Bézier form with control points $f_i = b_{2i}, b_{2i+1}, b_{2i+2} = f_{i+1}$ $1 \le i \le n-1$, and satisfying boundary conditions
>
> $$\begin{aligned} b_3 - b_2 &= \rho_1(b_2 - b_0), && \rho_1 > 0 \text{ arbitrary}, \\ b_{2n} - b_{2n-1} &= \rho_n(b_{2n+2} - b_{2n}), && \rho_n > 0 \text{ arbitrary}, \end{aligned}$$
>
> by finding suitable control points b_{2i+1}, $1 \le i \le n-1$.

Since curvature continuity is a nonlinear condition and no additional degrees of freedom are available, the problem is equivalent to a nonlinear system of equations (see (2)). It is not always solvable (see Figure 1), even if the data form a piecewise linear curve that is a limit of noninflecting curves. We therefore must look for a weak sufficient condition for solvability. Furthermore, we shall see that unicity of the interpolant requires an even stronger condition than the one that will be used for the existence proof.

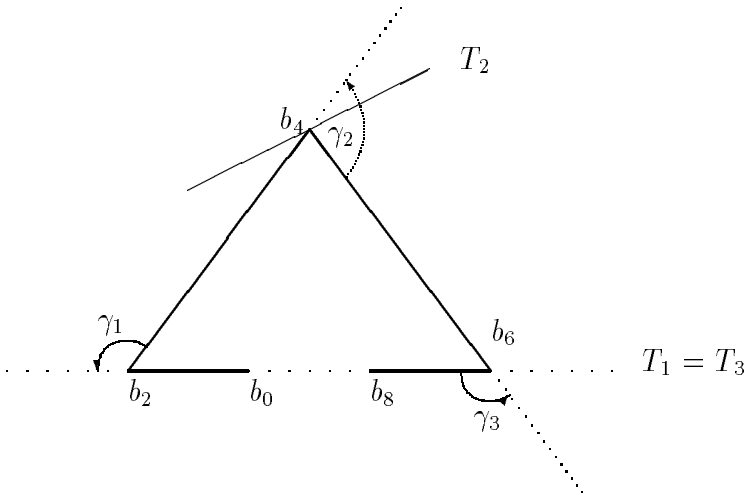

Figure 1: Unsolvable problem, $n = 3$

To discuss solvability conditions we introduce tangents T_i at interpolation points $f_i = b_{2i}$, $1 \le i \le n$, where the boundary conditions force T_1 and T_n to coincide with the lines $L(b_0, b_2)$ and $L(b_{2n}, b_{2n+2})$. respectively. Then b_{2i+1} is the intercept of T_i with T_{i+1}. The problem posed in Figure 1 is unsolvable, because T_2 has to intersect both T_1 and T_3 in finite points b_3 and b_5 which must lie on the "left" of the lines $L(b_2, b_4)$ and $L(b_4, b_6)$, respectively.

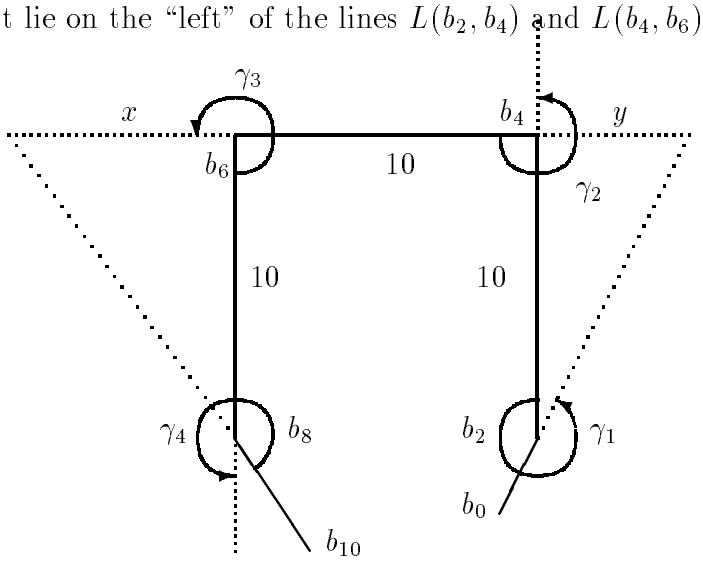

Figure 2: Partially unsolvable problem

Another example is given by Figure 2. This problem is unsolvable for $\sqrt{x} + \sqrt{y} < 2\sqrt{10}$. The piecewise linear interpolant of $b_0, b_2, \ldots, b_{10}$ takes two successive 90° turns and fails to satisfy the second part of

Definition 2.1 *A data set $b_0, b_2, \ldots, b_{2n+2}$ is called* **noninflecting**, *if all angles*

$$\gamma_i := \angle\,(b_{2i+2} - b_{2i},\ b_{2i} - b_{2i-2}), \qquad (1 \le i \le n),$$

measured in counterclockwise direction, are in $(0, 180°)$ or in $(180, 360°)$ (i.e. the data are always "turning right" or always "turning left"). If the sum of two consecutive angles γ_i always is in $(0, 180°)$ or in $(180, 360°)$, the data set is called **nonreflecting**.

Note that any nonreflecting and noninflecting data set admits $n-1$ "accompanying triangles" formed by b_{2i}, b_{2i+2} and the intercept $\widetilde{b_{2i+1}}$ of the lines $L(b_{2i-2}, b_{2i})$ and $L(b_{2i+2}, b_{2i+4})$ for $1 \le i \le n-1$ (see Figure 4). Each accompanying triangle must contain a control point b_{2i+1} of the solution, if the latter exists. The main existence result of this paper will be

Theorem 2.1 *For any noninflecting and nonreflecting data set there is at least one visually C^2 interpolant by a parametric quadratic spline.*

The proof is contained in the next two sections.

Note that Figure 2 can be made into a solvable problem by small perturbations of b_2 and b_8. However, as Figure 3 shows, a perturbed problem with noninflecting and nonreflecting data may have multiple solutions. To get uniqueness, we therefore need additional conditions.

Definition 2.2 *A nonreflecting and noninflecting data set is called* **nondegenerate**, *if all accompanying triangles have obtuse angles at the points $\widetilde{b_{2i+1}}$.*

An equivalent form of this condition is

Either all γ_i satisfy $\gamma_i \in (0, 90°)$ and $\gamma_i + \gamma_{i+1} < 90°$

or all γ_i satisfy $\gamma_i \in (270°, 360°)$ and $720° - \gamma_i - \gamma_{i+1} < 90°$,

and in short: A nondegenerate data set

- always turns to the left or always turns to the right,
- the direction changes only by acute angles,
- two successive direction changes sum up to an acute angle.

We remark that this form of nondegeneracy is a natural notion:

Theorem 2.2 *An ordered and sufficiently dense sample of sufficiently many points from a noninflecting C^2 curve will form a nondegenerate data set.* ∎

Our main result concerning uniqueness will be

Theorem 2.3 *A nondegenerate data set has a unique visually C^2 piecewise quadratic interpolant.*

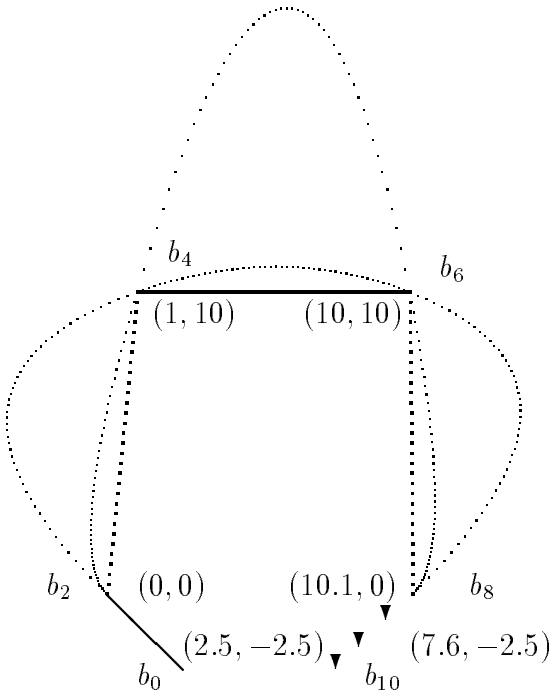

Figure 3: Data with two solutions (a third solution is suppressed)

3 Principles of construction

Consider part of a noninflecting and nonreflecting data set as in Figure 5. We introduce tangents T_i at b_{2i} forming angles α_i with $\widetilde{b_{2i+2} - b_{2i}}$. The intercept b_{2i+1} of T_i and T_{i+1} must lie in the accompanying triangle $b_{2i}b_{2i+2}b_{2i+1}$ to avoid unwanted inflection points. Curvature at b_{2i} is (up to a constant factor) equal to

$$\kappa_i^- = \frac{\text{dist } (b_{2i-2}, T_i)}{\|b_{2i-1} - b_{2i}\|_2^2} \quad \text{from the left and}$$

$$\kappa_i^+ = \frac{\text{dist } (b_{2i+2}, T_i)}{\|b_{2i+1} - b_{2i}\|_2^2} \quad \text{from the right}$$

(see e.g. [Boehm et. al., '84, p. 11]).

Using the obvious identities

$$\frac{\text{dist } (b_{2i+2}, T_i)}{h_i} = \sin\ \alpha_i$$

$$\frac{\text{dist } (b_{2i-2}, T_i)}{h_{i-1}} = \sin\ (\gamma_i - \alpha_i)$$

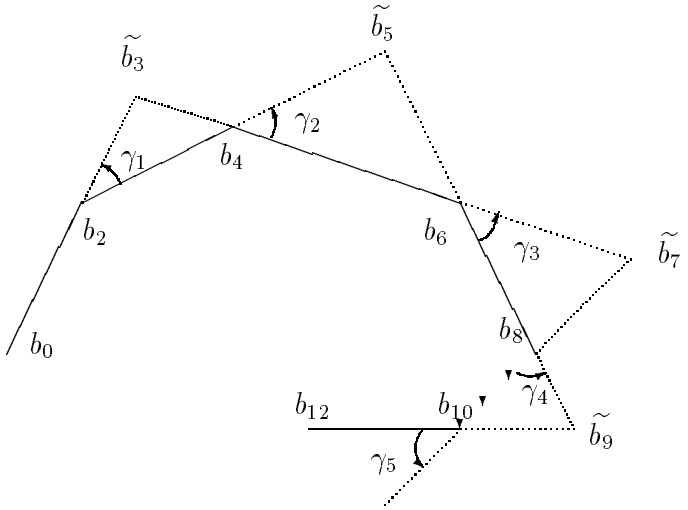

Figure 4: Accompanying triangles

$$\begin{aligned} \frac{\text{dist } (b_{2i-2}, T_i)}{\|b_{2i-2} - b_{2i-1}\|_2} &= \sin\ \delta_{i-1} = \sin(\gamma_i - \alpha_i + \alpha_{i-1}) \\ \frac{\text{dist } (b_{2i}, T_{i-1})}{\|b_{2i-1} - b_{2i}\|_2} &= \sin\ \delta_{i-1} \end{aligned}$$

we can express curvature continuity at b_{2i}, $2 \le i \le n-1$, by the equation

$$\frac{1}{h_i}\ \sin\ \alpha_i\ \frac{\sin^2(\gamma_{i+1} - \alpha_{i+1} + \alpha_i)}{\sin^2(\gamma_{i+1} - \alpha_{i+1})} = \frac{1}{h_{i-1}}\ \sin(\gamma_i - \alpha_i)\ \frac{\sin^2(\gamma_i - \alpha_i + \alpha_{i-1})}{\sin^2 \alpha_{i-1}} \tag{2}$$

involving only constants h_j, γ_j and the variables α_j. This gives $n-1$ nonlinear equations for the unknowns $\alpha_2, \ldots, \alpha_{n-1}$; we use $\alpha_1 = \gamma_1$ and $\alpha_n = 0$ to satisfy the boundary conditions.

The approach via the system (2) is used later for efficient numerical treatment of the problem. It does not yield much insight for the proof of existence and uniqueness of a solution.

For the latter we theoretically investigate a numerically hazardous "shooting strategy":

Start. Choose an arbitrary point b_3 between b_2 and $\widetilde{b}_3$ on T_1.

Iteration. For $2 \le i \le n-1$, if b_{2i-1} and T_i are given, construct b_{2i+1} on T_i by using curvature continuity at b_{2i} in the form

$$\|b_{2i+1} - b_{2i}\|_2^2 = \frac{\text{dist } (b_{2i+2}, T_i)}{\text{dist } (b_{2i-2}, T_i)}\ \|b_{2i-1} - b_{2i}\|_2^2.$$

If this quantity comes out to be too large, b_{2i+1} moves out of the accompanying triangle formed by $b_{2i}, b_{2i+2}, \widetilde{b_{2i+1}}$, and the process can not be repeated. Otherwise the tangent T_{i+1} is fixed by b_{2i+1} and b_{2i+2}, and the next value of i can be used.

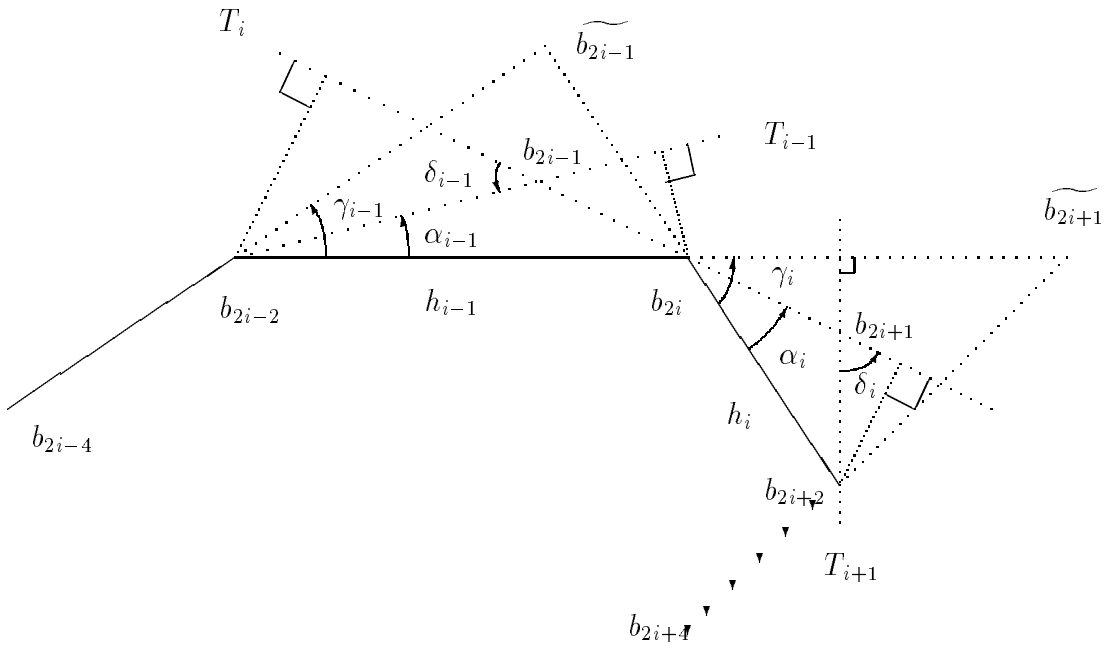

Figure 5: Local construction

This process tries to get through to the triangle $b_{2n-2}\ b_{2n}\ \widetilde{b_{2n-1}}$ and to hit the line $L(b_{2n}, \widetilde{b_{2n-1}})$ to produce a curvature continuous solution. It is comparable to the "single shooting" method for solving a boundary value problem in ordinary differential equations. Its (theoretical) success is equivalent to the solvability of the problem, but experience shows that it is numerically extremely sensitive. There are simple nondegenerate data sets for which 32–bit precision does not suffice to find some b_3 to shoot through to the last triangle. Replacement by a "multiple shooting" method is equivalent to solving the system (2), which is recommended for practical purposes (see section 6).

4 Existence proof

We now proceed to study the topological properties of a single step of the shooting method defined in the last section.

Let two lines L_1 and L_2 intersect in some point S with an angle α of less than 180 degrees. This angle forms two open cones C_1 and C_2; let $A_1 \neq S$ and $A_2 \neq S$ be points on L_1 and L_2, respectively, but on different sides of S (see Figure 6).

Definition 4.1 *A continuous curve $z : [0,1] \to \mathbb{R}^2$ is "transversal in C_1", if (see Figure 6)*

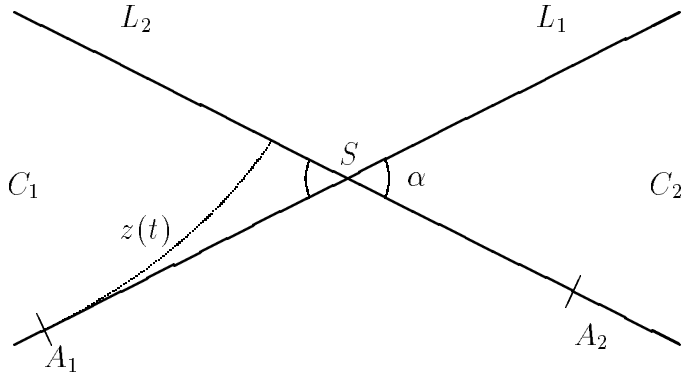

Figure 6: Transversal curve

a) $z((0,1)) \subset C_1$,

b) $z(0) = A_1$,

c) $z(1) \in L_2$,

d) $z(1) \neq S$.

Definition 4.2 *On the configuration of Figure 6 we define a mapping* $F : C_1 \to C_2$ *by*

a) $F(x) \in C_2$ *lies on the line* $L(x)$ *through* x *and* S

b) $$\frac{\text{dist } (A_1, L(x))}{\|x - S\|_2^2} = \frac{\text{dist } (A_2, L(x))}{\|F(x) - S\|_2^2}$$

This definition makes sure that the piecewise quadratic polynomial defined in Bézier form by control points A_1, x, S and $S, F(x), A_2$ is curvature continuous at S. Therefore F corresponds to a single step of the shooting method.

If we introduce polar coordinates as in Figure 7, we get for the image $F(r,\varphi) = (R(r,\varphi),\varphi)$ of (r,φ) the expression

$$R^2(r,\varphi) = \frac{h_2}{h_1} \frac{sin\ \varphi}{sin\ (\alpha - \varphi)} r^2 \tag{3}$$

which shows that F is a smooth function.

This follows from

$$\frac{h_1}{r^2} = \frac{h_2}{R^2}, \quad h_1 = h_1\ sin(\alpha - \varphi), \quad h_2 = h_2\ sin\ \varphi.$$

Lemma 4.1 *Let* z *be a transversal curve in* C_1. *Then:*

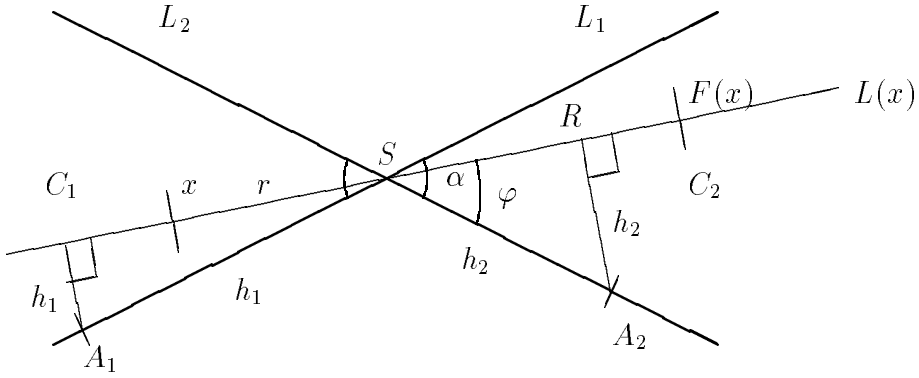

Figure 7: Construction of $F(x)$ in polar coordinates

a) $u(t) := F(z(t))$ *is a continuous curve on* $(0,1)$ *in* C_2.

b) $u(1) := S$ *makes* u *continuous on* $(0,1]$.

c) $\lim_{t\to 0} \|u(t)\|_2 = \infty$.

d) $\lim_{t\to 0} \operatorname{dist}(u(t), L_1) = 0$.

Proof. Everything follows easily from elementary continuity arguments using the properties of z and F described in the polar coordinate system of (3) and Figure 7. ■

We now observe that in the configuration of Figure 8 two instances of Figure 7 are intertwined. A transversal curve z_{i-1} in the cone C_i^- will map by Lemma 4.1 into a curve z_i in the cone C_i^+. This curve starts in b_{2i} and has the line $L(b_{2i-2}, b_{2i})$ as an asymptote. Now the intersection of this curve with the accompanying triangle $b_{2i}\ b_{2i+2}\ \widetilde{b_{2i+1}}$ is transversal in the cone C_{i+1}^-. Applying Lemma 4.1 again we get a curve z_{i+1} in the cone C_{i+1}^+ and so on. The shooting process starts with z_1 being part of T_1.

To prove Theorem 2.1 we now use induction on the statement

> "There is an open nonvoid interval on the line $L(b_2, \widetilde{b_3})$ that maps uniquely onto a curve z_i in C_i^+ that starts from b_{2i} and has the line $L(b_{2i-2}, b_{2i})$ as an asymptote."

This statement follows from the previous lemma for $i = 2$, and the discussion illustrated by Figure 8 gives the induction step.

The resulting curve z_{n-1} in C_{n-1}^+ necessarily intersects the line T_n given by the boundary condition at b_{2n} at least once. This proves existence of a solution. ■

Uniqueness is not guaranteed, because z_{n-1} might intersect T_n in several points. This can be actually observed by plotting the curves z_i, and this is how the solutions of Figure 3 were found.

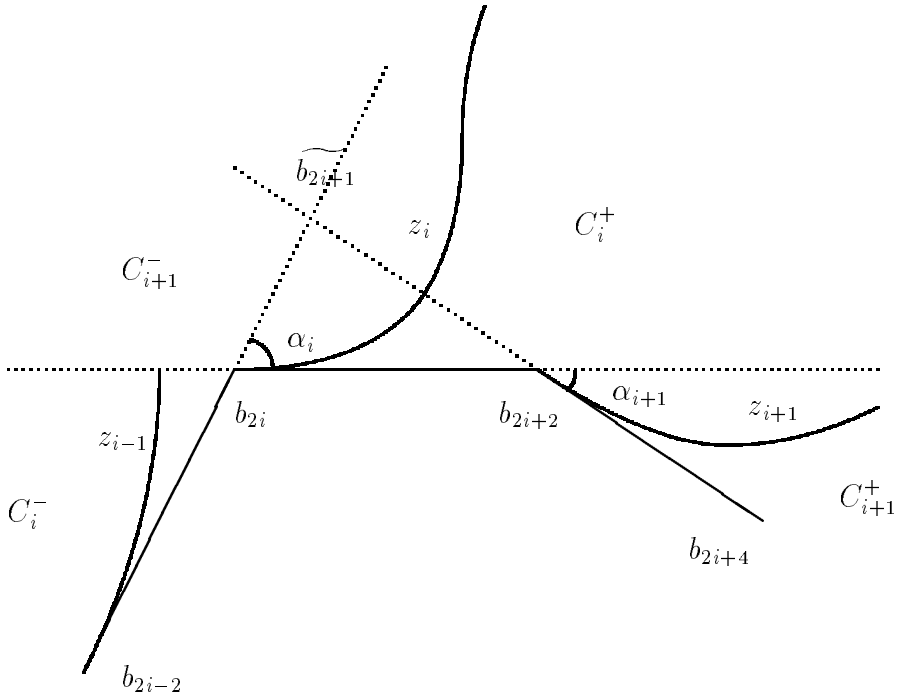

Figure 8: Shooting method

5 Uniqueness

Definition 5.1 *A transversal curve z in C_1 is strictly* **monotonic**, *if in the polar coordinates of Figure 7, the radius $r(\varphi)$ is strictly monotonic.*

Lemma 5.1

a) For the image $(R(r(\varphi),\varphi),\varphi)$ of a monotonic transversal curve $z = (r(\varphi),\varphi)$ in C_1 the radius $R(r(\varphi),\varphi)$ is strictly monotonic.

b) If $B_2 \neq S$ is a point on the line L_1 and the boundary of C_2 such that the angle $\gamma = \angle\,(A_2B_2S)$ is obtuse, the image curve $F(z)$ meets the line $L(B_2,A_2)$ in exactly one point (see Figure 9).

Proof. The first assertion is a direct consequence of (3). If $F(z)$ meets $\overline{B_2A_2}$ in two points P_1, P_2 with $\varphi_1 < \varphi_2$ (see Figure 10), then $R(r(\varphi),\varphi)$ is no monotonic function of φ. ■

For a nondegenerate data set we can apply Lemma 5.1 in each step of the shooting method, because all angles at $\widetilde{b_{2i+1}}$ are obtuse. Then the curves constructed in the previous section have unique intersections with the sides $\widetilde{b_{2i+1}}\ b_{2i+2}$ of the triangles $b_{2i}\ \widetilde{b_{2i+1}}\ b_{2i+2}$. This also holds for $i = n-1$, proving uniqueness. ■

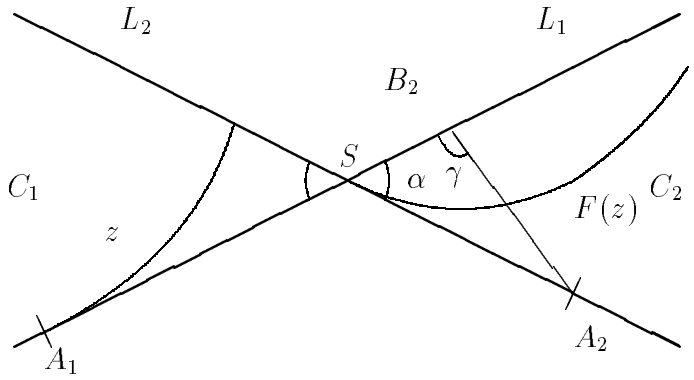

Figure 9: Monotonicity

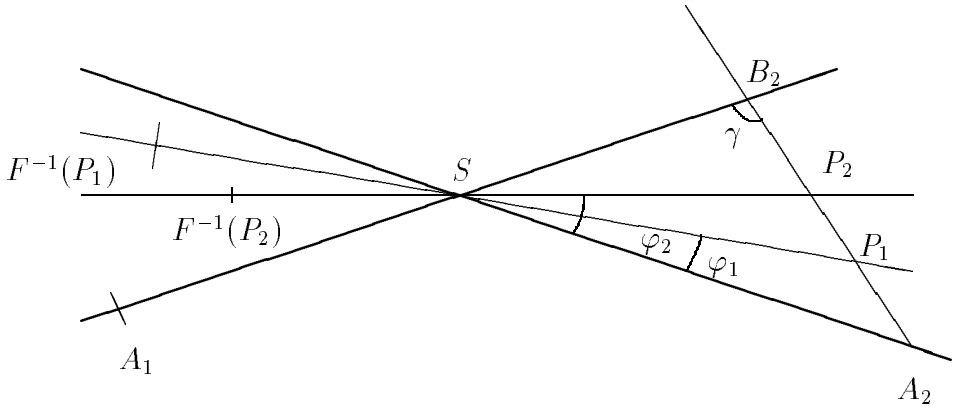

Figure 10: Uniqueness

6 Convergence

Let $f : [a, b] \longrightarrow I\!R^2$ be a smooth planar curve with positive curvature, parametrized by arclength. This section analyzes the convergence of the interpolation process for data from f, i.e. for data density tending to zero. We can follow [deBoor et. al., '87] to assume a local two–dimensional representation of f around a single point by

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

$$\theta(t) = \theta_1 t + \theta_2 t^2 + \mathcal{O}(t^3)$$

without loss of generality. Curvature at $f(t)$ is

$$\kappa(t) = \theta'(t) = \theta_1 + \frac{1}{2}\theta_2 t + \mathcal{O}(t^2).$$

The angle $\alpha(t)$ between the tangent at $f(0)$ and the chord

$$f(t) - f(0) = \int_0^t f'(\sigma)\mathrm{d}\sigma = \frac{1}{6}\begin{pmatrix} 6t - \theta_1^2 t^3 \\ 3\theta_1 t^2 + 2\theta_2 t^3 \end{pmatrix} + \mathcal{O}(t^4) \tag{4}$$

can be expanded via

$$\tan\alpha(t) = \frac{3\theta_1 t^2 + 2\theta_2 t^3 + \mathcal{O}(t^4)}{6t - \theta_1^2 t^3 + \mathcal{O}(t^4)}$$

as

$$\alpha(t) = \tan\alpha(t) + \mathcal{O}(t^3) = \frac{1}{2}\theta_1 t + \frac{1}{3}\theta_2 t^2 + \mathcal{O}(t^3) \tag{5}$$

for $t \to 0$. The chord length

$$h(t) = \|f(t) - f(0)\|_2 = t - \frac{1}{24}\theta_1^2 t^3 + \mathcal{O}(t^4)$$

can replace arc length t up to terms of order $\mathcal{O}(t^3)$. This allows expansions like (5) to be written in terms of chord length h instead of arc length:

Lemma 6.1 *The angle $\alpha = \alpha(f, t_0, t_1)$ between the tangent to a smooth planar curve f at $f(t_0)$ and a chord $f(t_1) - f(t_0)$ of length $h = \|f(t_1) - f(t_0)\|$ is*

$$\alpha = \frac{1}{2}\kappa(t_0)h + \frac{1}{3}\kappa'(t_0)h^2 + \mathcal{O}(h^3) \tag{6}$$

for $t_1 > t_0, t_1 \to t_0$, where κ is the curvature of the curve.

The expansion (6) is independent of the parametrization except for the orientation hidden in the correspondence of κ' and $t_1 > t_0$. For $t_1 < t_0$, the analogous expansion is

$$\alpha = -\frac{1}{2}\kappa(t_0)h + \frac{1}{3}\kappa'(t_0)h^2 + \mathcal{O}(h^3), \tag{7}$$

obtained from Lemma 6.1 by reversing the parametrization and the sign of curvature. Note that in both cases the angle has to be taken as oriented from the tangent to the chord.

Similarly, one easily verifies expansions of angles between successive chords in terms of the curvature at the central or a non–central point:

Lemma 6.2 *The angle $\gamma = \gamma(f, t_0, t_1, t_2)$ between two chords $f(t_{i+1}) - f(t_i), i = 0, 1$, of a planar curve f with chord lengths $h_i = \|f(t_{i+1}) - f(t_i)\|$ is*

$$\begin{array}{rcl} \gamma & = & \frac{1}{2}\kappa(t_1)(h_0 + h_1) + \frac{1}{3}\kappa'(t_1)(h_1^2 - h_0^2) + \mathcal{O}(\max(h_0^3, h_1^3)) \\ & = & \frac{1}{2}\kappa(t_0)(h_0 + h_1) + \frac{1}{3}\kappa'(t_0)(2h_0^2 + 3h_0h_1 + h_1^2) + \mathcal{O}(\max(h_0^3, h_1^3)) \end{array} \tag{8}$$

for $t_2 > t_1 > t_0, t_2 - t_0 \to 0$, where κ is the curvature of the curve. The angle between the tangent at $f(t_1)$ and the chord $f(t_1) - f(t_0)$ is

$$\alpha = \frac{1}{2}\kappa(t_0)h_0 + \frac{2}{3}\kappa'(t_0)h_0^2 + \mathcal{O}(h_0^3) \tag{9}$$

in terms of the curvature at t_0.

Proof: Expansion (9) simply follows from (8) for $t_2 = t_1, h_1 = 0$, and the first representation of γ in (6) can be written as the difference between (6) and (7). To obtain the non–central representation, the angle between the chord $f(t_2)-f(t_1)$ and the tangent at $f(t_0)$ is evaluated as

$$\frac{1}{2}\theta_1(h_0 + h_2) + \frac{1}{3}\theta_2(h_0^2 + h_0h_2 + h_2^2) + \mathcal{O}(h_2^3)$$

via an expansion of the chord like (4), using $h_2 = \|f(t_2) - f(t_0)\|$. The rest follows easily, using $h_2 = h_0 + h_1 + \mathcal{O}(\max(h_0^3, h_1^3))$. ■

We now go over to the interpolation of f at points $f_i = b_{2i} = f(t_i)$ for $1 \le i \le n$ with boundary conditions using the directions of $f'(t_1)$ and $f'(t_n)$. The density of knots is measured by

$$\begin{aligned} h &:= \max_{1\le i\le n-1} h_i, \\ h_i &:= \|f(t_{i+1}) - f(t_i)\|,\ 1 \le i \le n-1,\ \ h_0 := h_n := 0, \end{aligned}$$

and we consider the asymptotic situation $h \to 0, n \to \infty$. The data–defined chord angles γ_i and the unknown exact tangent angles α_i^* are representable as

$$\begin{aligned} \gamma_i &= \tfrac{1}{2}\kappa_i(h_i + h_{i-1}) + \tfrac{1}{3}\kappa_i'(h_i^2 - h_{i-1}^2) + \mathcal{O}(\max(h_{i-1}^3, h_i^3)), \\ \alpha_i^* &= \tfrac{1}{2}\kappa_i h_i + \tfrac{1}{3}\kappa_i' h_i^2 + \mathcal{O}(h_i^3), \end{aligned}$$

using Lemmas 6.1 and 6.2 and setting $\kappa_i := \kappa(t_i),\ \ \kappa_i' := \kappa'(t_i)$. As a first application we eliminate the unknown curvature κ_i and get numerically available tangent angles

$$\begin{aligned} \widehat{\alpha}_i &:= \gamma_i h_i/(h_{i-1} + h_i) \\ &= \tfrac{1}{2}\kappa_i h_i + \tfrac{1}{3}\kappa_i' h_i(h_i - h_{i-1}) + \mathcal{O}(h^3) \\ &= \alpha_i^* + \mathcal{O}(h^2) \end{aligned} \qquad (10)$$

as good approximations of α_i^*. We therefore recommend (10) for local methods and for starting Newton's method to solve the system (2).
The goal of this section is to prove

Theorem 6.1 *Global GC^2 piecewise quadratic interpolation of data from a smooth planar curve with nonvanishing curvature has an error of $\mathcal{O}(h^4)$ for data density $h \to 0$.*

The proof will be reduced to an application of

Theorem 6.2 *If local Hermite interpolation of a smooth planar curve near a point with nonvanishing curvature is done by quadratic polynomials using* **exact** *positions and* **approximate** *tangent directions with errors $\mathcal{O}(h^k)$ for $k = 2$ or 3, the error will be $\mathcal{O}(h^{k+1})$ in terms of the distance h between interpolation points.*

The proof of this result can be omitted because it is just a slight variation of the argument given by deBoor, Höllig, and Sabin in [deBoor et. al., '87] to prove their $\mathcal{O}(h^6)$ result for GC^2 cubics.

Convergence of order $\mathcal{O}(h^3)$ for the convexity preserving local Hermite interpolation by GC^1 piecewise quadratics using approximate tangent angles $\widehat{\alpha}_i$ defined in (10) follows immediately from Theorem 6.2. To prove $\mathcal{O}(h^4)$ convergence for the global GC^2 method of this paper, we have to show

$$\widetilde{\alpha}_i = \alpha_i^* + \mathcal{O}(h^3),\ \ h \to 0 \qquad (11)$$

uniformly with respect to i for the solution $\widetilde{\alpha}_i$ of the system (2). This will be done via Kantorovitch's convergence theorem for Newton's method (see e.g. [Ortega–Rheinboldt, '70, p. 421]) applied to a scaled version of the system (2).

To evaluate (2) at the exact tangent angles α^* we fix an index i and apply lemmas 6.1 and 6.2 to get the expansions

$$\begin{aligned}
\alpha^*_{i-1} &= \tfrac{1}{2}\kappa(t_i)h_{i-1} - \tfrac{2}{3}\kappa'(t_i)h_{i-1}^2 + \mathcal{O}(h^3)\\
\alpha^*_{i+1} &= \tfrac{1}{2}\kappa(t_i)h_i - \tfrac{2}{3}\kappa'(t_i)h_i + \mathcal{O}(h^3)\\
\gamma_{i+1} &= \tfrac{1}{2}\kappa(t_i)(h_i + h_{i+1}) + \tfrac{1}{3}\kappa'(t_i)(2h_i^2 + 3h_ih_{i+1} + h_{i+1}^2) + \mathcal{O}(h^3)\\
\gamma_i - \alpha^*_i &= \tfrac{1}{2}\kappa(t_i)h_{i-1} - \tfrac{1}{3}\kappa'(t_i)h_{i-1}^2 + \mathcal{O}(h^3)\\
\gamma_{i+1} - \alpha^*_{i+1} &= \tfrac{1}{2}\kappa(t_i)h_i + \tfrac{2}{3}\kappa'(t_i)h_i^2 + \mathcal{O}(h^3)\\
\gamma_i - \alpha^*_i + \alpha^*_{i-1} &= \kappa(t_i)h_{i-1} - \kappa'(t_i)h_{i-1}^2 + \mathcal{O}(h^3)\\
\gamma_{i+1} - \alpha^*_{i+1} + \alpha^*_i &= \kappa(t_i)h_i + \kappa'(t_i)h_i^2 + \mathcal{O}(h^3).
\end{aligned}$$

Putting these into (2) yields an $\mathcal{O}(h^2)$ expression after some calculations. Similar expansions can be derived for the angles $\widehat{\alpha}_i$ replacing α^*_i, and (2) comes out to be of order $\mathcal{O}(h)$ at $\widehat{\alpha}$. This is not surprising, because the approximate angles defined in (10) are exact up to $\mathcal{O}(h^3)$ for circular arcs.

To evaluate derivatives, we rescale (2) from the form $F(\alpha) = 0$ into $G(\beta) = 0$ by setting $\beta := T(\alpha)$, $G(\beta) := F(T^{-1}\beta)$ with T defined by

$$\beta_j := T_j(\alpha) := \frac{\alpha_j}{\min(h_{j-1}, h_j)},\ 2 \le j \le n-1.$$

Using $\widehat{\beta} := T(\widehat{\alpha})$ and $\beta^* := T(\alpha^*)$ we have

$$F(\widehat{\alpha}) = G(\widehat{\beta}) = \mathcal{O}(h), \quad F(\alpha^*) = G(\beta^*) = \mathcal{O}(h^2)$$

for the scaled version. Now we examine G' at β^* and get

$$\begin{aligned}
\frac{\partial G_i}{\partial \beta_{i-1}} &= 2\frac{\sin^2(\gamma_i - \alpha_i)\sin(\gamma_i - \alpha_i + \alpha_{i-1})}{\sin^3\alpha_{i-1}}\frac{\min(h_{i-1}, h_{i-2})}{h_{i-1}}\\
&= (4 + \mathcal{O}(h_{i-1}))\frac{\min(h_{i-1}, h_{i-2})}{h_{i-1}} = 4 + \mathcal{O}(h),\\
\frac{\partial G_i}{\partial \beta_{i+1}} &= 2\frac{\sin^2(\alpha_i)\sin(\gamma_{i+1} - \alpha_{i+1} + \alpha_i)}{\sin^3(\gamma_{i+1} - \alpha_{i+1})}\frac{\min(h_{i-1}, h_i)}{h_i}\\
&= (4 + \mathcal{O}(h_i))\frac{\min(h_{i-1}, h_i)}{h_i} = 4 + \mathcal{O}(h),\\
\frac{\partial G_i}{\partial \beta_i} &= 2\frac{\sin(\gamma_i - \alpha_i + \alpha_{i-1})}{\sin^2\alpha_{i-1}}\frac{\min(h_{i-1}, h_{i-2})}{h_{i-1}}\\
&\quad\cdot(\sin(\alpha_{i-1}) + 3\sin(\gamma_i - \alpha_i)\cos(\gamma_i - \alpha_i + \alpha_{i-1}))\\
&+ 2\frac{\sin(\gamma_{i+1} - \alpha_{i+1} + \alpha_i)}{\sin^2(\gamma_{i+1} - \alpha_{i+1})}\frac{\min(h_{i-1}, h_i)}{h_i}\\
&\quad\cdot(\sin(\gamma_{i+1} - \alpha_{i+1}) + 3\sin(\alpha_i)\cos(\gamma_{i+1} - \alpha_{i+1} + \alpha_i))\\
&= (8 + \mathcal{O}(h_{i-1}))\frac{\min(h_{i-1}, h_{i-2})}{h_{i-1}} + (8 + \mathcal{O}(h_i))\frac{\min(h_{i-1}, h_i)}{h_i} = 16 + \mathcal{O}(h)\\
&= 2\frac{\partial G_i}{\partial \beta_{i-1}} + 2\frac{\partial G_i}{\partial \beta_{i+1}} + \mathcal{O}(h).
\end{aligned}$$

The row–sum norm of G' at β^* then is bounded by $32 + \mathcal{O}(h)$, while the inverse has a norm bound of $\frac{1}{4} + \mathcal{O}(h)$. A similar calculation proves that all second derivatives of G_i at β^* are bounded with respect to $h \to 0$. This is where the special form of the scaling pays off.

Kantorovitch's theorem now implies that there exist $r > 0$ and $h_0 > 0$ such that there is a unique solution $\widetilde{\beta}$ of $G(\beta) = 0$ for each interpolation problem for f with data density h less than h_0, satisfying $\|\widetilde{\beta} - \beta^*\| < r$ in the sup–norm. Furthermore, the solution is reached with quadratic convergence from every starting point β with $\|\beta - \beta^*\| < r$. Because of

$$\begin{aligned} |\widehat{\beta}_i - \beta_i^*| &= |\widehat{\alpha}_i - \alpha_i^*| / \min(h_i, h_{i-1}) \\ &= (\tfrac{1}{3}|\kappa'(t_i)| h_i h_{i-1} + \mathcal{O}(\max(h_i^3, h_{i-1}^3)) / \min(h_i, h_{i-1}) \end{aligned}$$

the starting value $\widehat{\beta}$ will be successful for $h \to 0$, if the ratio

$$\max(h_i, h_{i-1}) / \min(h_i, h_{i-1}) \tag{12}$$

is uniformly bounded.

Since $G(\beta^*) = \mathcal{O}(h^2)$, the radius of the existence domain can be (theoretically) chosen as an $\mathcal{O}(h^2)$ function, and therefore the exact solution $\widetilde{\beta}$ satisfies $\|\widetilde{\beta} - \beta^*\| = \mathcal{O}(h^2)$. But then

$$\begin{aligned} \widetilde{\alpha}_i &= \widetilde{\beta}_i (h_i + h_{i-1}) = \beta_i^* (h_i + h_{i-1}) + \mathcal{O}(h^3) \\ &= \alpha_i^* + \mathcal{O}(h^3), \end{aligned}$$

as required for $\mathcal{O}(h^4)$ convergence of the interpolation process. ∎

7 Numerical treatment

The asymptotic considerations of the preceding section suggest a modified Newton–Raphson method for solving the nonlinear system (2). The Jacobians will be tridiagonal, and each iteration needs only $\mathcal{O}(n)$ operations. If (2) is written as $F(\alpha) = 0$ with $F : \mathbb{R}^{n-2} \to \mathbb{R}^{n-2}$, then a simple stepsize control is recommended to make sure that

a) $\|F(\alpha)\|_2$ decreases at each step

b) α stays in the admissible region (e.g. $0 < \alpha_i < \gamma_i$ for "right–turning" data or $\gamma_i < \alpha_i < 2\pi$ for "left–turning" data).

(see [Schaback, '88] for a comparable situation). Good results are obtained for starting values (10), because these satisfy (2) up to $\mathcal{O}(h)$. A Gauss–Seidel iteration of (2) may be employed for an existence proof via Brouwer's fixed–point theorem, but there are examples where the unique solution of (2) is a repelling fixed point of the Gauss–Seidel iteration.

Newton's method also works well in cases with multiple solutions and in cases of reflecting data with existing solutions. In cases of nonexisting solutions the method creeps to the boundary of the admissible domain, using stepsizes tending to zero. There are nondegenerate data sets where the matrix F' is not diagonally dominant at the solution, but no nondegenerate case is known where Newton's method failed to converge. For sufficiently dense samples of data with density $h \to 0$ and a bounded ratio (12), Newton's method will converge quadratically to the solution when started at (10), as the convergence results of the preceding section show.

8 Examples and Concluding Remarks

Figures 11 and 12 show two examples.

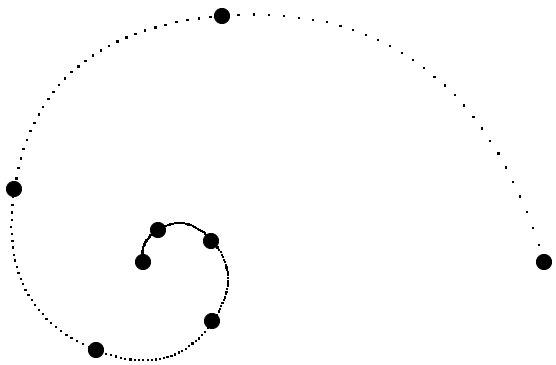

Figure 11: Logarithmic spiral, 8 points

For large numbers of data points, the solutions look very nice and closely approximate functions that supplied the data. If data points are far apart, the solution looks a little "too straight" at interior knots (see Figure 11). This is due to the fact that the interpolating parabolic pieces normally attain their curvature maximum in the interior of their domain.

Existence and uniqueness results extend to the case of piecewise rational quadratic curves with prescribed weights. This modifies the system (2) by some extra factors, but leaves the geometric proof for existence and uniqueness of solutions unchanged.

In a forthcoming paper we show how to place cubic polynomial pieces at positions where the data require an inflection point or a GC^2 continuation into a straight line. Furthermore, the tangential boundary conditions can be replaced by "not–a–knot–" conditions.

References

[1] Boehm, W., Farin, G.F., Kahmann, J. (1984), A survey of curve and surface methods in CAGD, Computer Aided Geometric Design 1, 1 – 60

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

[3] Farin, G.E. (1982), Visually C^2 cubic splines, Computer Aided Design 14, 137 – 139

[4] Ortega, J.M., and W.C. Rheinboldt (1970), Iterative Solution of Nonlinear Equations in Several Variables, Academic Press

[5] SCHABACK. R. (1988), Adaptive Rational Splines, NAM–Bericht 60, to appear in Constructive Approximation 1989.

Author's address:

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

D–3400 Göttingen

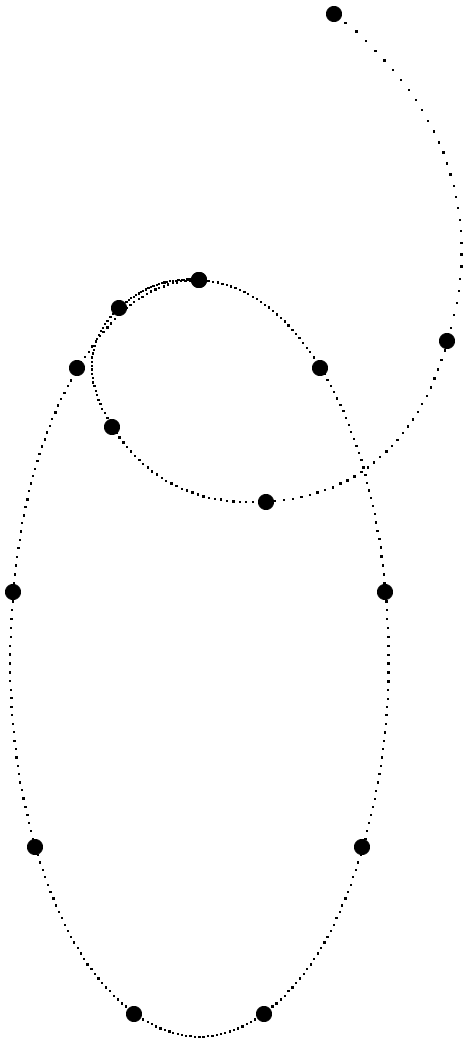

Figure 12: Letter "$\mathcal{O}$" formed by 15 points