Computer Aided Geometric Design 8 (1991) 227–238
North-Holland

Approximation of circular arcs by cubic polynomials

Michael Goldapp

Technische Universität, Angewandte Geometrie und Computergrafik, W-3300 Braunschweig, Germany

Received December 1989

Abstract

Goldapp, M., Approximation of circular arcs by cubic polynomials, Computer Aided Geometric Design 8 (1991) 227–238.

This paper deals with the approximation of circular arcs using polynomials of degree 3. Different types of boundary conditions and target functions are considered. The approximation turns out to have sixth order accuracy, giving a very small error in the radius.

Keywords. Bézier curves, circles, arcs, conic sections, geometric modelling, parametric approximation.

1. Overview

This paper gives a systematic approach in the approximation of circular arcs by means of cubic polynomials. Different types of boundary conditions and target functions are considered. The curves generated are optimal in that they minimize the remaining error. Also, the accuracy is of order six, which is two magnitudes greater than which is normally achievable using cubic approximation (cf. [de Boor et al. '87]).

In geometric modelling applications, the need for a substitute for circles arises when conic sections or rational curves, respectively, are not available or are not recommended. On the other hand, CAD systems and the like always seem to offer some representation of degree 3 polynomials, no matter whether they use B-splines, Bézier or the plain canonical base for curves. For a survey of curves and surfaces in a CAD system, cf. [Bézier '86].

Only unit circles with center in the origin have to be taken into account, as arbitrary sizes are available after simple scaling.

From the characteristic shape of the circle, the error in the radius along the new curve

$$\rho(t) = \sqrt{x^2(t) + y^2(t)} - 1 \tag{1}$$

seems to be the natural criterion for approximation quality. Instead, the function

$$\epsilon(t) = x^2(t) + y^2(t) - 1 \tag{2}$$

is much easier to handle. Some simple calculation shows that both functions have their zero sets and extremal values in the same locations. In the following, the term 'radius' not only stands for the constant r that describes a circle's size, but also for the varying distance $\rho(t) + 1$ from its center to an approximative curve. The actual meaning can be determined from the context in which the term appears.

Blinn ['87] gives the simple polynomials

$$x(t) = 1 - 1.344\ t^2 + 0.344\ t^3, \qquad y(t) = 1.656\ t - 0.312\ t^2 - 0.344\ t^3 \tag{3}$$

as a sufficient approximation of a quarter circle. The corresponding curve interpolates the arc at both end points tangentially. Using formula (1), it can easily be shown that the radius varies as

$$1 - 4.2 \cdot 10^{-4} < \rho(t) + 1 < 1 + 4.2 \cdot 10^{-4}$$

for $t \in [0, 1]$. Arbitrary arcs can be obtained by computing the Bézier points and then subdividing the curve using the famous de Casteljau algorithm [Boehm et al. '83]. However, the approximation quality does not change by this procedure, although shorter arcs can be treated better.

The first approach to approximations of arcs was taken by Bézier, when he designed a CAD system solely based on polynomials in Bernstein representation [Bézier '77, '86]. The results are given without a complete derivation. Moreover, there seem to be some misprints in this chapter of his book [Bézier '86].

He only considers in depth the case of a curve that interpolates both ends and the middle of the arc, for which he gives a formula that defines the location of the second control point of the curve. For a quarter circle, the error is $2.7 \cdot 10^{-4}$, whatsoever criterion he might have used.

The 'same area' criterion Bézier also mentions can be replaced by the simple equation

$$\int_0^1 \left[x^2(t) + y^2(t) - 1\right] \mathrm{d}t = 0. \tag{4}$$

In fact, this is not the exact formula to guarantee that the new curve has the same sector area as the arc. It only forces the areas of the component curves drawn above the t-axis to be identical.

The exact sector area is

$$\tfrac{1}{2}\int_0^1 \left[x(t)\frac{\mathrm{d}y(t)}{\mathrm{d}t} - y(t)\frac{\mathrm{d}x(t)}{\mathrm{d}t}\right] \mathrm{d}t \tag{5}$$

which computes to $\alpha/2$ for an arc of angular width α. It should be noted that both formulae coincide in the case of $x(t) = \cos(\alpha t)$, $y(t) = \sin(\alpha t)$. However, other approximation criteria can be thought of, each more or less suitable for particular applications. A great many of curves can be defined by the different arc points they are required to interpolate. Also, small variations of curvature are desirable. As another example, a curve that retains the circle's convexity or its arc length also might be of interest.

A Hermite-type approximation is considered first. The resulting curve interpolates the arc tangentially at both ends, which is the strongest possible boundary condition here. There is still one other condition available for specification (Section 2). If the boundary conditions are loosened, rather better approximations can be found. Section 3 deals with the case of a curve that is only C^0-continuous to a given arc. There, the well-known Chebychev polynomials become the means of finding the result. The approximation error can be reduced further if the curve is allowed to miss the arc's end points (Section 4). Section 5 brings some examples using the results for the different approaches.

When the underlying work for this article had been completed, the Oberwolfach talk by Dokken et al. ['90] was given, anticipating some of the results. They only considered the 'error-in-radius' criterion, obtaining curves of lowest possible errors with sixth order accuracy.

2. Hermite-type approximation

Throughout this paper, the class of functions used is the set of cubic polynomials in Bernstein–Bézier representation. The circle being a planar parametric curve, the approximation is of the form

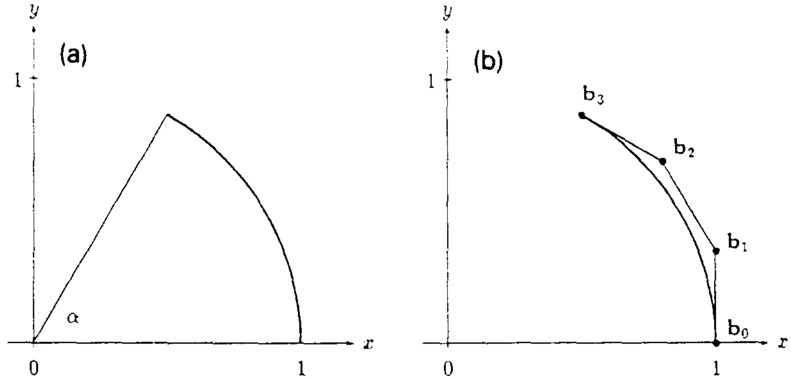

Fig. 1. (a) The given arc. (b) The approximation curve and its control points.

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \sum_{i=0}^{3} b_i B_i^n(t) \tag{6}$$

with the control points b_i and the Bernstein polynomials $B_i^n(t)$ given by

$$b_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix}, \qquad B_i^n(t) = \binom{3}{i}(1-t)^{3-i}t^i, \quad t \in [0, 1]. \tag{7}$$

Further details on Bézier curves and other sources of information can be found in the survey article by Boehm et al. ['83] and also in the recent book by Farin ['88].

Let the circular arc to be approximated be given by its angular width α, starting in the point $[{}^1_0]$ on the positive x-axis. Only arcs up to the half circle can be represented by one single curve segment. Thus, $0 < \alpha < \pi/2$ is pre-requisite. Fig. 1 illustrates the relationship stated above.

In the following, a cubic curve is constructed that passes through the end points and there has a collinear tangent with the arc. Since a Bézier curve passes through b_0 and b_3 and its boundary tangents are collinear to the vectors $(b_1 - b_0)$ at the starting point and $(b_3 - b_2)$ at the end point, the control points of the new curve must be

$$b_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad b_1 = \begin{bmatrix} 1 \\ h \end{bmatrix}, \qquad b_2 = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix} + h \cdot \begin{bmatrix} \sin\alpha \\ -\cos\alpha \end{bmatrix}, \qquad b_3 = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}, \tag{8}$$

with some adjustable parameter h that not only controls the length of the tangents but also the shape of the curve in the middle (see Fig. 2).

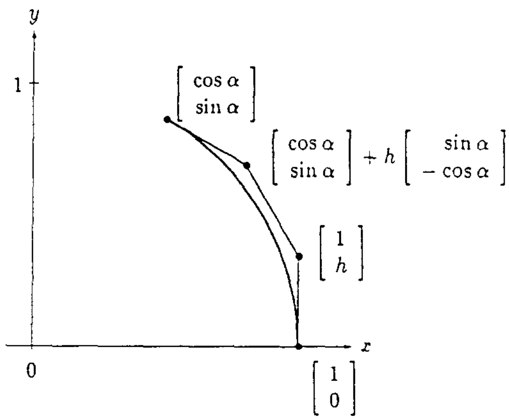

Fig. 2. Control points for tangential approximation.

Due to the assumptions as above, the error function $\epsilon(t)$ as of (2) has double zeros on the boundary and is of the form

$$\epsilon(t) = t^2(t-1)^2 P_2(t) \tag{9}$$

where $P_2(t)$ denotes a degree 2 polynomial of t (and of h), which is

$$\begin{aligned} P_2(t) &= (t^2 - t)\cdot(18h^2(1+\cos\alpha) - 24h\sin\alpha + 8(1-\cos\alpha)) \\ &\qquad + (9h^2 + 6h\sin\alpha - 6(1-\cos\alpha)) \qquad (10a) \\ &= h^2\cdot 9(2t(t-1)(1+\cos\alpha)+1) \\ &\qquad - h\cdot 6(4t(t-1)-1)\sin\alpha + 2(4t(t-1)-3)(1-\cos\alpha). \qquad (10b) \end{aligned}$$

Note that for curve symmetry, the parameter t only occurs as $t(t-1)$ which does not change if t is replaced by $1-t$.

Another zero a of the error function can be selected to determine h. Substituting t by a in (10b), with the abbreviation $\bar{a} = 2a(a-1)$, gives

$$h^2\cdot 9((1+\cos\alpha)\bar{a}+1) - h\cdot 6\sin\alpha\,(2\bar{a}-1) + 2(1-\cos\alpha)(2\bar{a}-3) = 0, \tag{11}$$

which is a quadratic equation with two real solutions, the positive one being the sought value of h:

$$h = \frac{\sin\alpha\,(2\bar{a}-1) + \sqrt{2\bar{a}(2\cos\alpha - 1) - \cos^2\alpha\,(1+2\bar{a}) + 7 - 6\cos\alpha}}{3 + 3\bar{a}(1+\cos\alpha)} \tag{12a}$$

$$= \frac{2}{3}\,\frac{(2\bar{a}-1)\sin\tfrac{1}{2}\alpha\cos\tfrac{1}{2}\alpha + \sin\tfrac{1}{2}\alpha\sqrt{4-(1+2\bar{a})\sin^2\tfrac{1}{2}\alpha}}{1+2\bar{a}\cos^2\tfrac{1}{2}\alpha}. \tag{12b}$$

Evaluating (9) and (10b) for $t = \frac{1}{2}$ gives the error in the middle as

$$\epsilon(\tfrac{1}{2}) = (\tfrac{3}{4}h\sin\tfrac{1}{2}\alpha + \cos\tfrac{1}{2}\alpha)^2 - 1. \tag{13}$$

Substituting h as of (12b) into (13) results in

$$\epsilon(\tfrac{1}{2}) = \left(\frac{\sin^2\tfrac{1}{2}\alpha\left((2\bar{a}-1)\cos\tfrac{1}{2}\alpha + \sqrt{4-(2\bar{a}+1)\sin^2\tfrac{1}{2}\alpha}\right)}{2+4\bar{a}\cos^2\tfrac{1}{2}\alpha} + \cos\tfrac{1}{2}\alpha\right)^2 - 1, \tag{14}$$

for which the Taylor expansion at $\alpha_0 = 0$ is

$$|\epsilon(\tfrac{1}{2})| = |\epsilon(\alpha_0+\alpha)| = \left|\frac{1+2\bar{a}}{4096}\alpha^6 + \cdots\right| \tag{15a}$$

while the Euclidian distance is

$$|\rho(\tfrac{1}{2})| = |\rho(\alpha_0+\alpha)| = \left|\frac{1+2\bar{a}}{8192}\alpha^6 + \cdots\right|. \tag{15b}$$

Together with the fact that at $\frac{1}{2}$ the error functions have their maximum values this means that the approximation error converges with the sixth order of the interval length $[0, \alpha]$, which is two magnitudes faster than in the general case of cubic approximation (e.g. [de Boor et al. '87]). Dokken et al. ['90] failed to derive these expansions due to their different approach.

For obvious symmetry, $a = \frac{1}{2}$ has been choosen in previous works (cf. [Bézier '77, '86], [Dokken et al. '90]). Setting $a = \frac{1}{2}$ and $\bar{a} = -\frac{1}{2}$ in (11)— or simply forcing $\epsilon(\frac{1}{2})$ of (13) to zero —gives

$$h_1 = \tfrac{4}{3}\sin\tfrac{1}{2}\alpha/(1+\cos\tfrac{1}{2}\alpha) = \tfrac{4}{3}\tan\tfrac{1}{4}\alpha$$

(cf. [Bézier '77]). It is easy to verify that $\epsilon(\frac{1}{2}) \leqslant 0$ while the equality only holds for $a = \frac{1}{2}$.

Substituting h_1 into $P_2(t)$ as in (10a) and then $P_2(t)$ into (9) gives

$$\left|\max_{t\in[0,1]} \epsilon(t)\right| = \frac{4}{27}\frac{\sin^6\frac{1}{4}\alpha}{\cos^2\frac{1}{4}\alpha} \tag{16}$$

(cf. [Dokken et al. '90]). Computing the Taylor expansion of the right-hand side of (16) at $\alpha_0 = 0$ gives

$$\epsilon(\alpha_0+\alpha) = \tfrac{4}{27}\left(\alpha^6 + \tfrac{1}{1920}\alpha^{10} + \cdots\right)/4096 \tag{17}$$

while the Taylor expansion of $\rho(\alpha) = \sqrt{1+\epsilon(\alpha)} - 1$ is

$$\rho(\alpha_0+\alpha) = \tfrac{2}{27}\left(\alpha^6 + \tfrac{1}{1920}\alpha^{10} + \cdots\right)/4096. \tag{18}$$

With $a = \frac{1}{2}$, formula (9) becomes $\epsilon(t) = C\,t^2(t-1)^2(t-\frac{1}{2})^2$ which is non-negative for $t \in [0, 1]$. This means that the approximative curve is always farther away from the origin than the arc itself (cf. [Dokken et al. '90]). This suggests that there is an approximation with less error in the radius for smaller a.

Next, the optimal values for both a and h under the 'error-in-radius' criterion will be derived. $\epsilon(t)$ can also be written as

$$\epsilon(t) = C\,t^2(t-1)^2(t-a)(t-1+a) \tag{19}$$

with $a \in [0, 1]$ as the zero, and a constant C that does not depend on t. Some simple calculations give

$$t_1 = 0, \quad t_2 = 1, \quad t_3 = \tfrac{1}{2} \quad \text{and} \quad t_{4,5} = \tfrac{1}{2} \pm \tfrac{1}{6}\sqrt{24a^2 - 24a + 9}$$

as the locations of the extremal values of $\epsilon(t)$. These values can be computed, giving

$$\epsilon(t_1) = \epsilon(t_2) = 0, \qquad \epsilon(t_3) = -\tfrac{1}{64}C(2a-1)^2 \quad \text{and}$$

$$\epsilon(t_4) = \epsilon(t_5) = \tfrac{4}{27}Ca^3(1-a)^3.$$

If $\epsilon(t)$ is forced to have the smallest possible variation on the interval [0, 1], it seems to be sufficient for consecutive extremal values to equioscillate (see Fig. 3). Hence,

$$\epsilon(t_3) - \epsilon(t_4) = \tfrac{1}{64}C(2a-1)^2 + \tfrac{4}{27}Ca^3(1-a)^3 = 0 \tag{20}$$

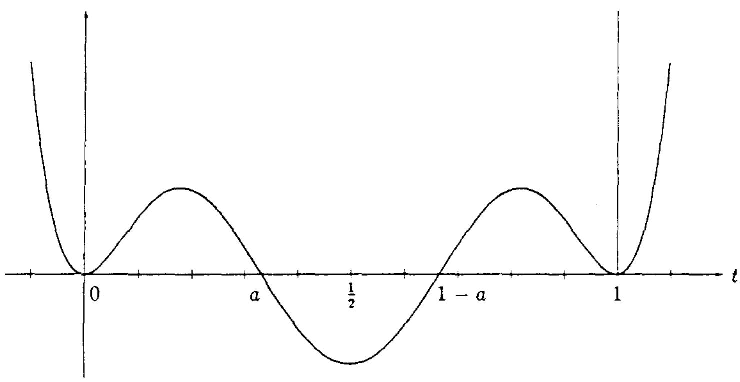

Fig. 3. The equioscillating error function.

holds for a. This is a degree 6 equation that can be transformed into a cubic equation with one real and two complex solutions. The corresponding root of (19) is

$$a = \tfrac{1}{2}\left(1 - \sqrt{1 + \tfrac{3}{2}\left(\sqrt[3]{\sqrt{2} - 1} - \sqrt[3]{\sqrt{2} + 1}\right)}\right) \tag{21}$$

for which $a \approx 0.3373$, while Fig. 3 yet suggests that $a \approx \frac{1}{3}$. The corresponding value of h_2 is obtained using (11) with

$$\bar{a} = \tfrac{3}{4}\sqrt[3]{\sqrt{2} - 1} - \tfrac{3}{4}\sqrt[3]{\sqrt{2} + 1}\,.$$

This leaves to show that the equioscillation of $\epsilon(t)$ really gives the smallest extremal values. This is not necessarily true as C actually depends on h and thus, on a.

Equation (10a) is actually

$$(a^2 - a)C + 9h^2 + 6h \sin\alpha - 6(1 - \cos\alpha) = 0$$

with

$$C = 18h^2(1 + \cos\alpha) - 24h \sin\alpha + 8(1 - \cos\alpha),$$

which leads to

$$(a^2 - a)C(1 + \cos\alpha) + \tfrac{1}{2}C + 6h \sin\alpha\,(\cos\alpha + 3) - 6\sin^2\alpha - 4(1 - \cos\alpha) = 0.$$

Substituting h by (12a) and solving for C gives

$$\begin{aligned} C = \Big(&8\sin^2\alpha\,(\cos\alpha - 1)(a^2 - a) + 4\sin^2\alpha\,(\cos\alpha + 6) + 8 \\ &-4\sin\alpha\,(\cos\alpha + 3)\sqrt{16 - (3 + \cos\alpha)^2 - 4(a^2 - a)(1 - \cos\alpha)^2}\Big) \\ &/\big(1 + 2(1 + \cos\alpha)(a^2 - a)\big)^2. \end{aligned} \tag{22}$$

It has to be shown that $C(a)$ is monotonic for $a \in [0, \frac{1}{2}]$. Let $f(x)$ and $g(x)$ be non-zero, monotonic functions over an interval, then $1/g(x)$ and $f(x) \cdot g(x)$ are also monotonic. Thus, it suffices to show that both the numerator and the denominator of (22) are non-zero and monotonic for $a \in [0, \frac{1}{2}]$. Elementary transformations give the result that both have zeros only for $\frac{1}{2}\,(1 \pm \mathrm{i}\tan\frac{1}{2}\alpha)$, while the derivatives have an additional zero at $\frac{1}{2}$.

In the following, other approximation goals will be considered. Using (5), after substituting $x(t)$ and $y(t)$ by the Bézier polynomials as of (6) and (8), the sector area $A(\alpha)$ of the approximative curve can be computed as

$$A(\alpha) = -\tfrac{3}{20}h^2 \sin\alpha + \tfrac{3}{5}h(1 - \cos\alpha) + \tfrac{1}{2}\sin\alpha. \tag{23}$$

For given h, $|A(\alpha) - \frac{1}{2}\alpha|$ is the difference between the sector area of the arc and the one of the new curve. On the other hand, solving the equation $A(\alpha) = \frac{1}{2}\alpha$ for h results in a curve that not only coincides with the arc in both end points and tangents, but also in the area.

Performing these steps for formula (4) gives the equation

$$h^2(12 - 9\cos\alpha) + 26h \sin\alpha + 18(\cos\alpha - 1) = 0, \tag{24}$$

which can be taken to generate a very good approximation to the exact area.

3. Least error approximation with given end points

When higher accuracy is essential while the tangential continuity can be discarded, a more general approximative curve can be constructed.

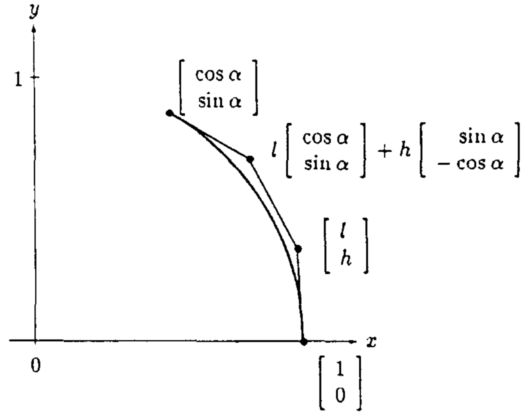

Fig. 4. Control points for C^0 approximation.

According to the list under (8), for a curve that does not necessarily have the same tangents as the arc, the control points must be of the form

$$b_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad b_1 = \begin{bmatrix} 1 \\ h \end{bmatrix}, \quad b_2 = l \cdot \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix} + h \cdot \begin{bmatrix} \sin\alpha \\ -\cos\alpha \end{bmatrix}, \quad b_3 = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}, \tag{25}$$

with some adjustable parameters h and l (see Fig. 4).

From the boundary conditions, the error function is

$$\epsilon(t) = Ct(t-1)(t-a)(t-1+a)(t-b)(t-1+b) \tag{26}$$

with a constant C, that does not depend on t. The remaining zeros a and b can be freely selected from [0, 1], making the curve comply to two arbitrary extra conditions.

In the next step the values of both a and b are to be adjusted for the optimal curve (which has a minimal maximum error). According to a well-known result from approximation theory, of all monic degree n polynomials $P_n(u)$, $n \geqslant 1$, the polynomial $T_n(u)/2^{n-1}$ minimizes

$$\max_{t\in[-1,1]} |P_n(u)| \tag{27}$$

with $T_n(u)$ being the *Chebychev* polynomials with $T_n(u) = \cos(n \arccos u)$.

$T_n(u)$ has its roots for

$$u_i = \cos\left(\frac{2i+1}{n} \cdot \frac{\pi}{2}\right), \quad i = 0, 1, \ldots, n-1,$$

and extremal values for

$$u_j = \cos\left(\frac{2j}{n} \cdot \frac{\pi}{2}\right), \quad j = 0, 1, \ldots, n$$

with $|T_n(u_j)| = 2^{n-1}$. Hence, to make the polynomial (26) equioscillating, $T_6(u)/32$ is the right selection, transforming the roots by

$$t_i = \tfrac{1}{2}\left(1 + \frac{u_i}{u_0}\right) \tag{28}$$

to appear in the interval [0, 1] (see Fig. 5).

By this procedure, the roots

$$u_0 = \tfrac{1}{2}\sqrt{2+\sqrt{3}}, \quad u_1 = \tfrac{1}{2}\sqrt{2}, \quad u_2 = \tfrac{1}{2}\sqrt{2-\sqrt{3}},$$
$$u_3 = -u_2, \quad u_4 = -u_1 \quad \text{and} \quad u_5 = -u_0$$

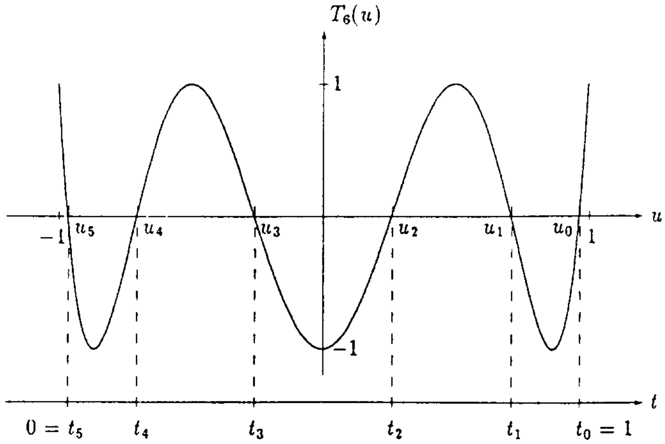

Fig. 5. Roots of the error function.

are mapped to

$$t_0 = 1, \qquad t_1 = \tfrac{1}{2}\sqrt{3}, \qquad t_2 = \tfrac{1}{2}(3 - \sqrt{3}),$$
$$t_3 = \tfrac{1}{2}(-1 + \sqrt{3}), \qquad t_4 = 1 - \tfrac{1}{2}\sqrt{3} \quad \text{and} \quad t_5 = 0.$$

There is an extremal value for $u = 0$, which is $t = \frac{1}{2}$.

The error function now is a degree 6 polynomial with zeros at t_i, $i = 0, 1, \ldots, 5$. Using the abbreviation $\bar{t} = t(t-1)$,

$$\begin{aligned}\epsilon(t) = &\big(9(2\bar{t}^2(1 + \cos\tfrac{1}{2}\alpha) + \bar{t}\big) \cdot h^2 \\ &+ 9\big(2\bar{t}^2(1 - \cos\tfrac{1}{2}\alpha) + \bar{t}\big) \cdot l^2 \\ &- 36\bar{t}^2 \sin\tfrac{1}{2}\alpha \cdot hl + 6(2\bar{t}^2 + \bar{t}) \sin\tfrac{1}{2}\alpha \cdot h \\ &- 6\big(2\bar{t}^2(1 - \cos\tfrac{1}{2}\alpha) + (4 - \cos\tfrac{1}{2}\alpha)\bar{t} + 1\big) \cdot l \\ &+ 2\bar{t}^2(1 - \cos\tfrac{1}{2}\alpha) + 9\bar{t} + 6\big)\bar{t}. \end{aligned} \tag{29}$$

Substituting $\bar{t} = t_1(t_1 - 1)$ and $\bar{t} = t_2(t_2 - 1)$ in (29) gives two simultaneous quadratic equations in both h and l. For a numerical solution, excellent initial values are available as it is known that $l \approx 1$ and $h \approx h_1$ (cf. equation (15)). Note that the complexity is reduced slightly as $2t_1(t_1 - 1) = t_2(t_2 - 1)$ holds.

The error in the middle of the arc (which also is one of the extremal values) is

$$\epsilon(\tfrac{1}{2}) = \tfrac{1}{16}\big((3l + 1)\cos\tfrac{1}{2}\alpha + 3h \sin\tfrac{1}{2}\alpha\big)^2 - 1 \tag{30}$$

For $l = 1$ the formula above is identical to (13).

4. Non-C^0 approximation

If even the curve's end points do not have to coincide with those of the arc, the control points are

$$b_0 = \begin{bmatrix} p \\ 0 \end{bmatrix}, \qquad b_1 = \begin{bmatrix} l \\ h \end{bmatrix}, \qquad b_2 = l \cdot \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix} + h \cdot \begin{bmatrix} \sin\alpha \\ -\cos\alpha \end{bmatrix}, \qquad b_3 = p \cdot \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}. \tag{31}$$

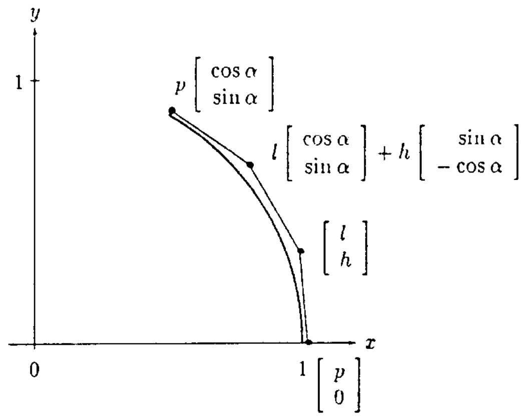

Fig. 6. Control points for non-C^0 approximation.

with adjustable parameters h, l and p (see Fig. 6). To maintain simple continuity between adjacent segments of approximation both the beginning and the end points of the curve may vary only on a line towards the circle's center.

Then the function (2) becomes a degree 6 polynomial that can be written as

$$\epsilon(t)=C(t-t_0)(t-t_1)(t-t_2)(t-t_3)(t-t_4)(t-t_5). \tag{32}$$

with a constant C and the 6 zeros t_i that can be freely selected from the interval [0, 1] to optimize the approximation.

Hence, to make the polynomial (32) equioscillating, the roots t_i must be those of $T_6(u)/32$ transformed by $t_i=\frac{1}{2}(u_i+1)$ to appear in the interval [0, 1] (see Fig. 7).

By this procedure, the roots

$$u_0=\tfrac{1}{2}\sqrt{2+\sqrt{3}}\,, \qquad u_1=\tfrac{1}{2}\sqrt{2}\,, \qquad u_2=\tfrac{1}{2}\sqrt{2-\sqrt{3}}\,,$$

$$u_3=-u_2, \qquad u_4=-u_1 \quad \text{and} \quad u_5=-u_0$$

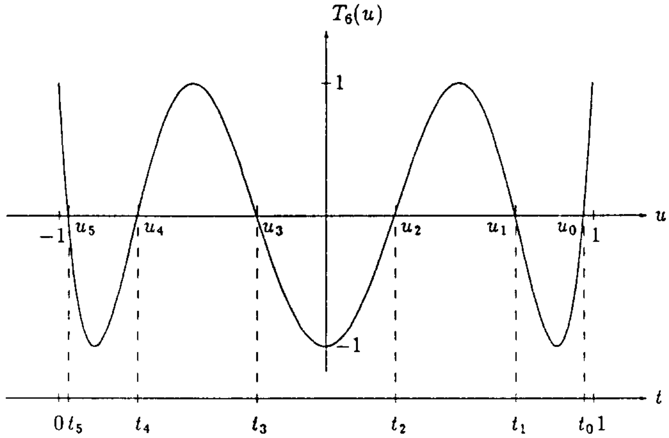

Fig. 7. Roots of the error function.

are mapped to

$$t_0 = \tfrac{1}{2} + \tfrac{1}{4}\sqrt{2+\sqrt{3}}\,, \qquad t_1 = \tfrac{1}{2} + \tfrac{1}{4}\sqrt{2}\,, \qquad t_2 = \tfrac{1}{2} + \tfrac{1}{4}\sqrt{2-\sqrt{3}}\,,$$

$$t_3 = \tfrac{1}{2} - \tfrac{1}{4}\sqrt{2-\sqrt{3}}\,, \qquad t_4 = \tfrac{1}{2} - \tfrac{1}{4}\sqrt{2} \quad \text{and} \quad t_5 = \tfrac{1}{2} - \tfrac{1}{4}\sqrt{2+\sqrt{3}}\,.$$

There is an extremal value for $u = 0$, which is $t = \frac{1}{2}$.

Writing (32) for t_1, t_2 and t_3 in a similar way as (29) yields a system of three nonlinear equations for the variables h, l and p. This set of equations can only be solved numerically, again with the excellent initial values $h \approx h_1$ (cf. (15)), $l \approx 1$ and $p \approx 1$.

5. Examples

The results obtained above will now be used to specify approximations of a quarter circle for each of the three types of boundary conditions.

For a quarter circle ($\alpha = \frac{1}{2}\pi$), equation (11) simplifies to

$$h^2 \cdot 9(\bar{a} + 1) - h \cdot 6(2\bar{a} - 1) + 2(2\bar{a} - 3) = 0, \tag{33}$$

giving $h_2 \approx 0.55191496$, while (15) leads to $h_1 = \frac{4}{3}(\sqrt{2} - 1) \approx 0.55228475$. Using these values, the sector areas of the curves compute as $\frac{3}{20}(-h^2 + 4h + \frac{10}{3})$.

The curve with the same area as the quarter circle described by (23) is

$$h^2 - 4h + \tfrac{5}{3}(\pi - 2) = 0 \tag{34}$$

with the solution $h_3 = 2 - \frac{1}{3}\sqrt{66 - 15\pi} \approx 0.55177848$, while the area approximation (23) yields

$$h^2 + \tfrac{13}{6}h - \tfrac{3}{2} = 0 \tag{35}$$

with the solution $h_4 = \frac{1}{12}(\sqrt{385} - 13)$, which is $h_4 \approx 0.55178474$.

The fifth curve was computed by numerical methods, using the approach given by Hoschek ['87]. He presented an approximation scheme of high-order Bézier curves by those of degree 3 by computing sample points on the curve given and then approximating them. This approach allows the generalization in that any type of curve can be used as input, e.g. a circular arc. Performing these steps gives $h_5 \approx 0.55188037$. The quality achieved here is quite remarkable compared to the analytical methods presented.

Table 1 given below summarizes the results with regard to several criteria. The terms ϵ_{max} and ρ_{max} denote the maximum error in the radius encountered as specified by (2) and (1). Then

Table 1
Results for Hermite-type approximation

	$\lvert\epsilon_{min}\rvert$	$\lvert\rho_{min}\rvert$	$A(\alpha) - \pi/4$	$\kappa(0)$	κ_{min}	κ_{max}
h_1	$\approx 545 \cdot 10^{-6}$	$\approx 273 \cdot 10^{-6}$	$\approx -220 \cdot 10^{-6}$	≈ 0.9786	≈ 0.9938	≈ 1.0080
h_2	$\approx 392 \cdot 10^{-6}$	$\approx 196 \cdot 10^{-6}$	$\approx -59 \cdot 10^{-6}$	≈ 0.9806	≈ 0.9926	≈ 1.0084
h_3	$\approx 537 \cdot 10^{-6}$	$\approx 268 \cdot 10^{-6}$	0	≈ 0.9815	≈ 0.9922	≈ 1.0086
h_4	$\approx 530 \cdot 10^{-6}$	$\approx 265 \cdot 10^{-6}$	$\approx -3 \cdot 10^{-6}$	≈ 0.9814	≈ 0.9922	≈ 1.0085
h_5	$\approx 429 \cdot 10^{-6}$	$\approx 214 \cdot 10^{-6}$	$\approx 44 \cdot 10^{-6}$	≈ 0.9809	≈ 0.9925	≈ 1.0085

The first entry in column 3 ($273 \cdot 10^{-6}$) also appears in [Bézier '77, '86], while the number below that one was also given in [Dokken et al. '90].

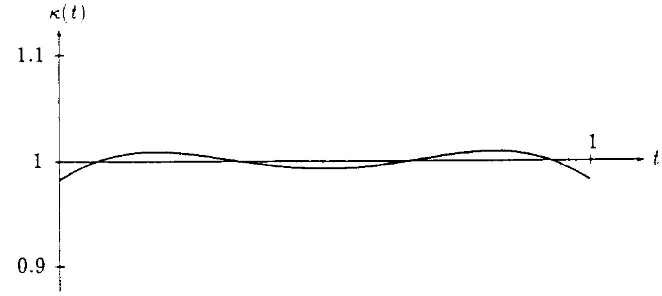

Fig. 8. Curvature of the approximative curve.

the difference in the sector area follows, while the last two columns have the variation of the curvature, which was calculated by the formula

$$\kappa(t)=\left\|\begin{bmatrix}x'(t)\\y'(t)\end{bmatrix}\times\begin{bmatrix}x''(t)\\y''(t)\end{bmatrix}\right\| \Big/ \left\|\begin{bmatrix}x'(t)\\y'(t)\end{bmatrix}\right\|^{3/2} \tag{36}$$

Fig. 8 shows the curvature $\kappa(t)$ plotted over the parameter t. Using (36), the curvature can be evaluated at particular locations, e.g. $\kappa(0)=\frac{2}{3}(1-h)/h^2$ (for the global minimum) and $\kappa(\frac{1}{2})=\frac{8}{3}\sqrt{2}\,h/(2-h)^2$ (for the local minimum). As the maximum curvature can not be obtained algebraically, the value of κ_{max} was computed numerically instead. The numerator of (36) is $18(t(1-t)(3h-2)h+(1-h)h)$ for a quarter circle, having real roots for t only if $\frac{2}{3}\leqslant h\leqslant 2$. Hence, the approximation curve remarkably does not have any inflection points. Further computations give the more general condition $\frac{2}{3}\tan\frac{1}{2}\alpha\leqslant h\leqslant 2\tan\frac{1}{2}\alpha$ for the approximative curve of angle α to have inflection points.

For the case of given end points only, the numerical result is $h_6\approx 0.553177370$ and $l_6\approx 0.998978326$. Now the error in the radius is about 3 times smaller than for the Hermite-type case, while also the area has come very close to the exact value. See Table 2.

For the non-C^0 approximation the values of the curve parameters computed numerically are $h_7\approx 0.553429256$, $l_7\approx 0.998733275$ and $p_7\approx 1.000055077$. This time the error in the radius is only slightly smaller than before, while the error in the area has halved again. See Table 3.

Note that there is hardly any improvement in the curvature at all, no matter what type of approximation is used.

Remark. The series expansions of (15a) and (15b) were obtained using the *Mathematica* system by Wolfram Research, Inc.

Table 2
Results for C^0-approximation

	$\lvert\epsilon_{min}\rvert$	$\lvert\rho_{min}\rvert$	$A(\alpha)-\pi/4$	$\kappa(0)$	κ_{min}	κ_{max}
h_6, l_6	$\approx 137\cdot 10^{-6}$	$\approx 68\cdot 10^{-6}$	$\approx 5.4\cdot 10^{-6}$	≈ 0.9694	≈ 0.9962	≈ 1.0052

Table 3
Results for non-C^0-approximation

	$\lvert\epsilon_{min}\rvert$	$\lvert\rho_{min}\rvert$	$A(\alpha)-\pi/4$	$\kappa(0)$	κ_{min}	κ_{max}
h_7, l_7, p_7	$\approx 110\cdot 10^{-6}$	$\approx 55\cdot 10^{-6}$	$\approx 2.7\cdot 10^{-6}$	≈ 0.9669	≈ 0.9967	≈ 1.0045

References

Bézier, P. (1977), Essay de définition numérique des curbes et des surfaces expérimentals (in French), Ph.D. thesis, Paris, p. 89.

Bézier, P. (1986), *The mathematical basis of the UNISURF CAD system*, Butterworths, London, ISBN 0-406-22175-5, p. 13.

Blinn, J. (1987), How many ways can you draw a circle?, IEEE Computer Graphics Appl. 8, 39.

Boehm, W., Farin, G. and Kahmann, J. (1983), A survey of curve and surface methods in CAGD, Computer Aided Geometric Design 1, 1–61.

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

Dokken, T., Daehlen, M., Lyche, T. and Mørken, K. (1990), Good approximation of circles by curvature-continuous Bézier curves, Computer Aided Geometric Design 7, 33–41.

Farin, G. (1988), *Curves and Surfaces for Computer Aided Geometric Design*, Academic Press, New York, ISBN 0-12-249050-9.

Hoschek, J. (1987), Approximate conversion of spline curves, Computer Aided Geometric Design 4, 59–66.