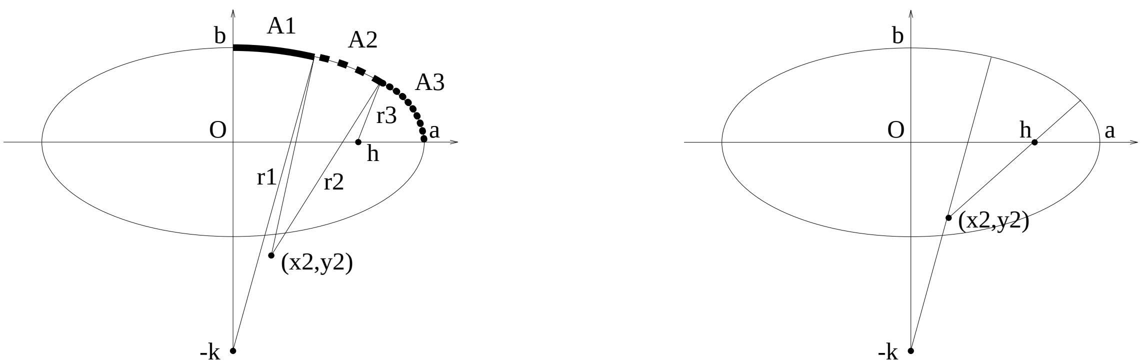
The Ellipse and the Five-centred Arch

Paul L. Rosin and Michael L.V. Pitteway
Department of Information Systems and Computing,
Brunel University,
Middlesex,
UK.
email: {Paul.Rosin,Mike.Pitteway}@brunel.ac.uk

1 Introduction

There has been a long history in the approximation of ellipses by circular arcs in order to simplify their construction and manipulation. This was of use for a wide variety of applications such as mathematics (generating figures), astronomy (analysing orbits), art (marking out large oval frames for ceiling painting), architecture (building masonry arches, floor plans, etc), and, more recently, the conversion of fonts from a general conic specification to circular arcs [6, 7]. Documented evidence goes as far back as the Italian Renaissance when various schemes were published by the architect Sebastiano Serlio in the sixteenth century [5]. More contentiously, it has been argued that fifteen centuries previously the Romans used such approximations when designing and building their ampitheatres [4]. Looking yet another fifteen centuries further back takes us to the construction of the megalithic stone "circles", which were in fact often elliptical or egg shaped. Many theories concerning megalithic man's knowledge of geometry have been propounded, some of which are based on piecewise circular approximations [14].



(a) unconstrained arc joins (b) C^1 arc joins

Figure 1: Geometry of the five-centred arch

More recently there have been many techniques published in the technical drawing literature (see for example Browning [2]), most of which concentrate on the three-centred arch (i.e. three arcs are used to construct the semi-ellipse, and so the full ellipse would be made up from four arcs), e.g. Rosin [13]. In this paper we look at the next step in improving accuracy: insertion of an additional arc between each of the previous arcs to form the five-centred arch. Figure 1a shows the basic geometry. The ellipse has semi-major and semi-minor axes of length a and b respectively. The three approximating arcs are drawn in one of the quadrants, and have parameters

arc A_1: centre $(0, -k)$, radius r_1
arc A_2: centre (x_2, y_2), radius r_2
arc A_3: centre $(h, 0)$, radius r_3.

The remaining arcs are determined by reflecting these three about the two axes. A constraint that is usually included so as to simplify the construction and to ensure a reasonable appearance is that A_1 and A_3 pass through the vertices of the ellipse: $(0, b)$ and $(a, 0)$. Another common constraint is to set the radii of A_1

and A_3 to match the ellipse's curvatures at its vertices. Since the curvature at $(0, b)$ and $(a, 0)$ is $\frac{b}{a^2}$ and $\frac{a}{b^2}$ this fixes the centres to

$$\begin{aligned} h &= \frac{a^2 - b^2}{a} \\ k &= \frac{a^2 - b^2}{b}. \end{aligned}$$

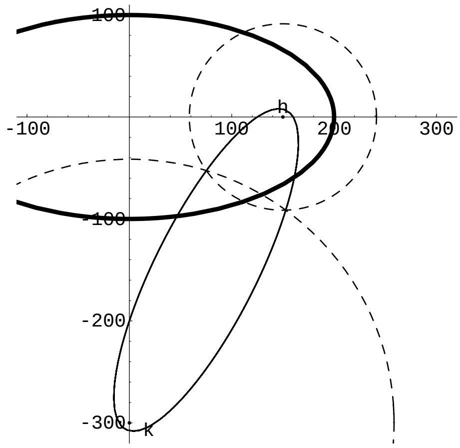


Figure 2: The ellipse being approximated ($a = 200$, $b = 100$) is drawn bold. The two circular loci of C^1 constraints for each end of A_2 are shown for $r_2 = 141$; their intersections lie on the elliptical locus of tangent constraints.

It is well known that the visual appearance of the approximation is improved if all the arcs join smoothly. This enables the arrangement to be simplified to the geometry shown in figure 1b. Given the previous constraint on A_1 and A_3 only one variable remains: the centre of A_2, and this can be specified by the length of r_2. To identify the locus of centres of A_2 we first note that if tangent continuity holds with A_3 then the locus is a circle with centre $(h, 0)$ and radius $r_2 - (a - h)$. Likewise, the locus of centres of A_2 with tangent continuity with A_1 is a circle with centre $(0, -k)$ and radius $(k + b) - r_2$. On eliminating r_2 from the equations of these two circles we obtain the locus of centres of A_2 which has continuous tangents with both its neighbouring arcs. Interestingly, as seen in figure 2, this takes the form of an ellipse (which we call the C^1 constraint ellipse) whose parameters (semi-major and semi-minor axes, orientation of major axis measured from the vertical, and centre) are

$$\begin{aligned} a' &= \frac{a^3 - b^3}{2ab} \\ b' &= \frac{a - b}{2} \\ \theta' &= \tan^{-1} \frac{b}{a} \\ xc' &= \frac{a^2 - b^2}{2a} = \frac{h}{2} \\ yc' &= -\frac{a^2 - b^2}{2b} = -\frac{k}{2}. \end{aligned}$$

In addition we note that its foci lie at $(h, 0)$ and $(0, -k)$. As shown in figure 3 the C^1 constraint ellipse lies on a plane in XYR space

$$r = A - By$$

where

$$\begin{aligned} A &= \frac{a^3 + b^3}{2ab} \\ B &= \frac{(a + b)\sqrt{a^2 + b^2}}{a^2 + ab + b^2} \end{aligned}$$

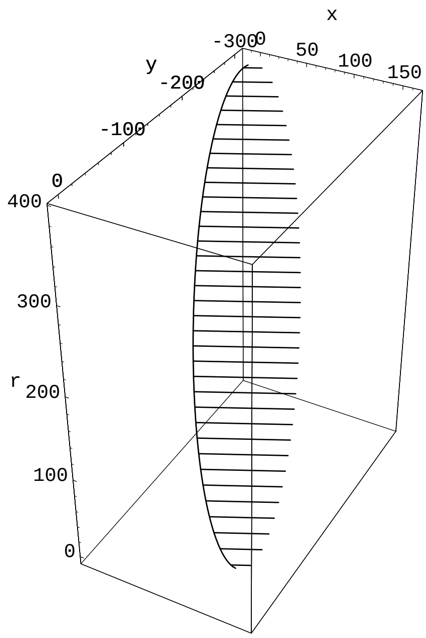


Figure 3: C^1 constraint ellipse plotted in XYR space

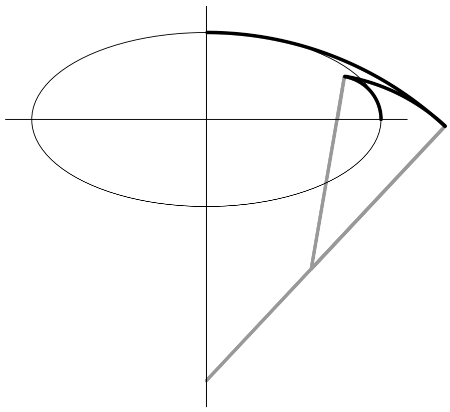


Figure 4: A "C^1" five-centred arch produced by a solution on the far side of the C^1 constraint ellipse plotted in figure 3; despite matching tangents at their joins the solution is invalid since the arcs double back.

and y is measured along the major axis of the continuity ellipse. As we move away from the intended ellipse the central arc's radius that is required for tangent continuity increases linearly. For a given radius r_2 there are two centres on the C^1 constraint ellipse, but only the one closer to the origin provides the desired solution. Plotting out an example of the other class of solutions we see in figure 4 that the arcs double back on each other.

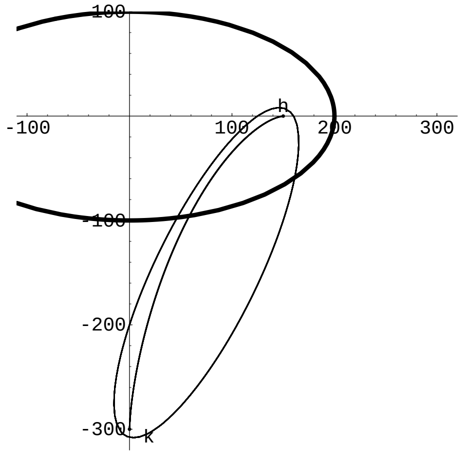


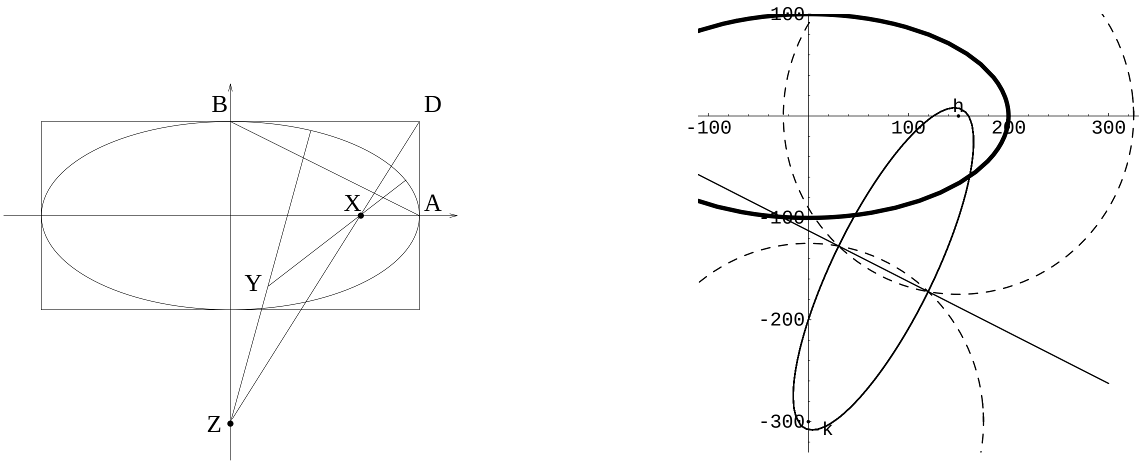
Figure 5: The ellipse being approximated is drawn bold. Both the elliptical locus of tangent continuity constraints and the evolute specifying the constraint of matching A_2's curvature to the ellipse's at the point of contact are shown.

An alternative constraint is that A_2 should touch the ellipse and match its curvature at that point just as arcs A_1 and A_3 do. Another way of stating this is that the centre of A_2 lies on the evolute of the ellipse which is described by the astroid-like Lamé curve

$$(ax)^{\frac{2}{3}} + (by)^{\frac{2}{3}} = (a^2 - b^2)^{\frac{2}{3}},$$

which is plotted in one quadrant in figure 5. It can be seen that the two constraints are incompatible.

2 Some Previous Constructions



(a) construction method

(b) the constraints

Figure 6: Lockwood's five-centred arch

We will now examine several proposed five-centred arches described in earlier issues of the *Mathematical Gazette*. Lockwood [10] shows a graphical construction that provides tangent continuity (see figure 6). The centres of A_1 and A_3 are determined by drawing the perpendicular to AB through D. The intersections with the axes at X and Z are equivalent to h and $-k$ in our notation. Smooth joins are implemented by keeping $AX + XY + YZ = BZ$, which in combination with the values for h and k provide another means of determining the elliptical locus of tangent continuity. To eliminate the final variable Lockwood suggests setting $XY = YZ$.[1] Since $r_1 = AX + XY$, $r_2 = BZ = 2XY + AX$, and $r_3 = AX$ this effectively makes $r_2 = \frac{r_1+r_3}{2}$. Assuming the geometry for tangent continuity this leads to the linear constraint

$$y = -\frac{b}{a}x - \frac{(a^2-b^2)^2}{2a^2b}$$

which is the extension of the C^1 constraint ellipse's minor axis. This enables the centre of A_2 to be found rather more simply by finding its intersection with one of the circular loci of constraints (see figure 6b), yielding:

$$\frac{a-b}{2}\left[\frac{a+b}{a} - \frac{a}{\sqrt{a^2+b^2}}, -\frac{a+b}{b} + \frac{b}{\sqrt{a^2+b^2}}\right].$$

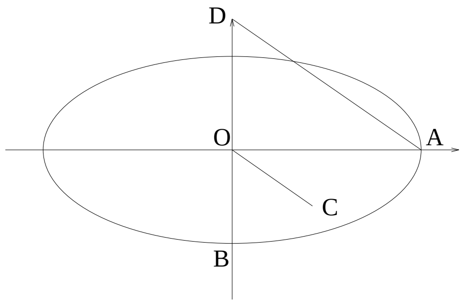


Figure 7: Walker's five-centred arch

Walker [15] states that the most appropriate radius for A_2 is the geometric mean $r_2 = \sqrt{r_1 r_3} = \sqrt{ab}$ which he matches to the ellipse's radius of curvature at the point of contact. He constructs $OD = \sqrt{ab}$ and then draws parallel to DA the line OC of length $a - b$, which locates the centre of A_2 at

$$\left(\frac{a-b}{\sqrt{1+\frac{b}{a}}}, \frac{a-b}{\sqrt{1+\frac{a}{b}}}\right).$$

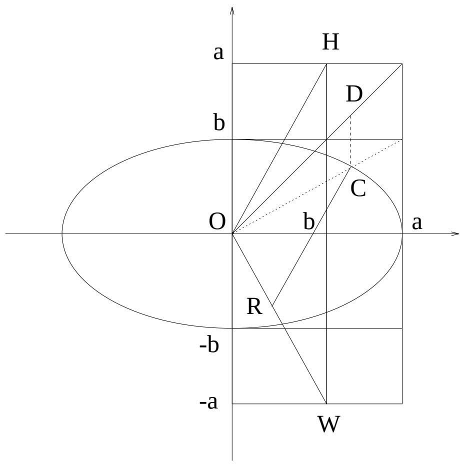


Figure 8: Chaplin's five-centred arch construction

Chaplin [3] describes another method for ensuring that A_2 matches the ellipse's radius of curvature. OD equal to a is drawn at $45°$ and a vertical is dropped from D to the ellipse at C (figure 8). The line CR

[1] In fact, Lockwood originally described his approach considerably earlier [8]. His construction was later much simplified by Lodge, who also provided the $XY = YZ$ constraint [11].

is drawn parallel to OH and the intersection with OW locates R, the centre of A_2, at

$$\left(\frac{a^2-b^2}{2\sqrt{2}a}, \frac{a^2-b^2}{2\sqrt{2}b}\right) = \frac{1}{2\sqrt{2}}(h,k).$$

The eccentric angle of C is $\frac{\pi}{4}$, making it straightforward to calculate the radius of curvature as

$$r_2 = \frac{\left(a^2+b^2\right)^{\frac{3}{2}}}{2\sqrt{2}ab} = \frac{\left(r_1^{\frac{2}{3}} + r_3^{\frac{2}{3}}\right)^{\frac{3}{2}}}{2\sqrt{2}}.$$

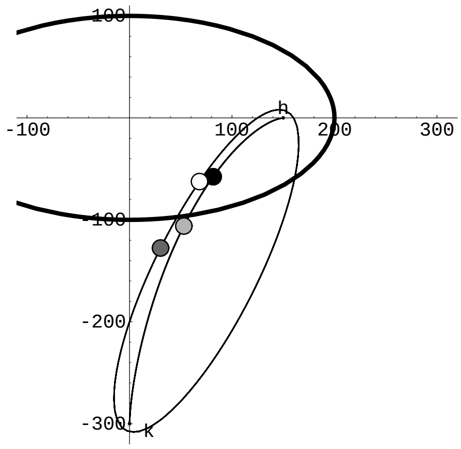


Figure 9: Different A_2 centres on the ellipse's evolute and C^1 constraint ellipse: Walker = black; Lockwood = dark gray; Chaplin = light gray; LS fit = white.

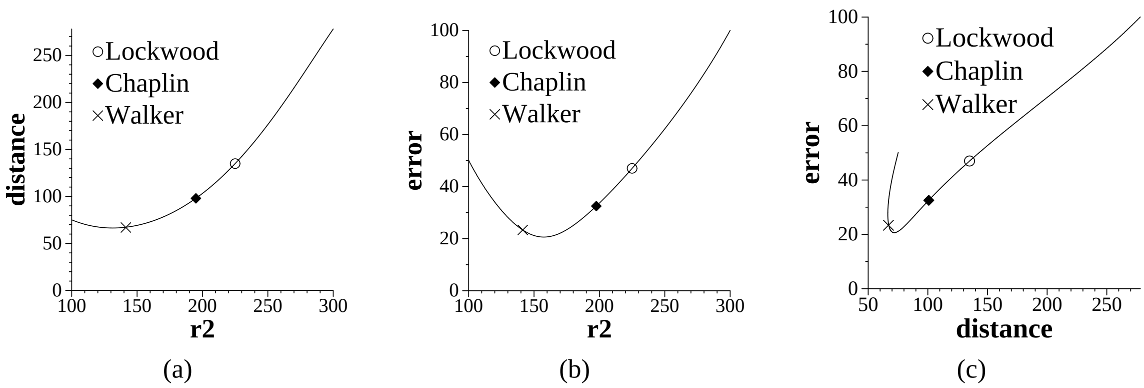


Figure 10: (a) Distance between the C^1 constraint ellipse and the ellipse's evolute parameterised by r_2. (b) Approximation error of tangent continuous arcs. The C^1 arcs with radii equal to Chaplin and Walker's constructions are also included. (c) Replotting error against distance shows how Walker's construction is a good compromise to minimising error and discontinuity.

Both Walker and Chaplin intended that the three specified arcs should be joined using additional arcs (e.g. using French curves). On examination the reason becomes clear; if the arcs both touch the ellipse and match its curvature at these points then they will not touch or cross each other since moving along the evolute in one quadrant, e.g. from $(0,-k)$ to $(h,0)$, all the osculating circles lie inside each other.[2]

Figure 9 shows the position along either the evolute or the C^1 constraint ellipse of the centres of the alternative choices for A_2. In order to determine the significance of the various positions we now look

[2]For the same reason three-centred arches cannot match the curvatures at the ellipse's vertices – unlike all the five-centred arches considered so far. However, Rosin [13] describes a construction which effectively matches curvatures and rescales h and k by $\frac{1}{2(2-\sqrt{2})}$. This decreases and increases r_1 and r_3 respectively, providing a valid (but not C^1) join as well as improving the similarity between the mean circular arc and ellipse curvatures.

more closely at the relationship between the evolute, the C^1 constraint ellipse, and the quality of the arc approximation. For this task the parametric form of the evolute is useful

$$\begin{aligned} x(t) &= \frac{a^2 - b^2}{a} \cos^3 t \\ y(t) &= -\frac{a^2 - b^2}{b} \sin^3 t \end{aligned}$$

where $t \in [\frac{3}{2}\pi, 2\pi]$ for the first quadrant. Stepping along the C^1 constraint ellipse and the evolute the shortest distance between them is numerically estimated, as shown in figure 10a. It can be seen that Walker's method is almost (but not quite) on the position of the evolute that is closest to the elliptical locus. Thus it could be considered as a good compromise between the two conflicting constraints: tangent continuity and curvature matching. Moreover, on retaining r_2 but shifting A_2 to provide C^1 continuity the approximation errors (figure 10b) show that Walker's method is better than the other two in this respect too. Replotting the error against distance in figure 10c further shows that Walker's construction is a good compromise to minimising both the approximation error and tangent discontinuity since it lies close to both optima.

3 Improved Approximations

(a) values of r_2 (b) approximation errors

Figure 11: The optimal and other values of r_2 when tangent continuity is enforced at the joints and curvature matching at the ellipse's vertices. Also shown are the incurred errors.

So far we have considered several possibilities from the literature for selecting the radius for A_2. Alternatively, in an attempt to improve the approximation, we can numerically estimate the optimal value for r_2. Its solution over a range of ellipse eccentricities is shown in figure 11a, and its least squares cubic fit is

$$\frac{r_2}{b} = 0.519763 + 0.407721 \frac{a}{b} + 0.055386 \left(\frac{a}{b}\right)^2 + 0.000002 \left(\frac{a}{b}\right)^3$$

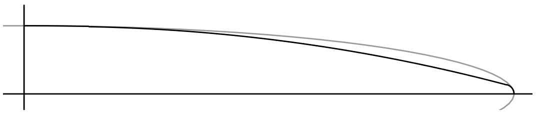
implying that the quadratic is a sufficient model. Re-expressed in terms of the radii of the outer two arcs and somewhat simplified gives

$$r_2 = \frac{1}{2} r_1^{\frac{2}{3}} r_3^{\frac{1}{3}} + \frac{2}{5} r_1^{\frac{1}{3}} r_3^{\frac{2}{3}} + \frac{19}{333} r_3$$

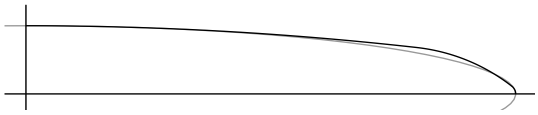
which can be seen from figure 11a to provide a good fit to the optimum. The actual errors in the approximations produced by the different schemes are shown in figure 11b.[3] Lockwood's proposal (the arithmetic mean) for r_2 does poorly for elongated ellipses. For ellipses with $\frac{b}{a} < \frac{1}{2}$ Walker's proposal (the geometric mean) is almost as good as the optimal solution. The optimum and the LS fit are clearly superior for elongated ellipses. For close to round ellipses the LS approximation to the optimum is inadequate, and the other methods do better; nevertheless, all produce low errors. An example of the five-centred arch approximations of ellipses is shown in figure 12.

Future investigations will consider making further improvements by relaxing some of the constraints. For instance, to simplify matters all the previous examples and our analysis has fixed the outer two arcs

[3] The errors with respect to the ellipse with $b = 100$ were calculated by sampling the three circular arcs into 1000 equal arclength sections. The error at each point was calculated using an accurate estimate of the length of the perpendicular to the ellipse developed by Rosin [12].

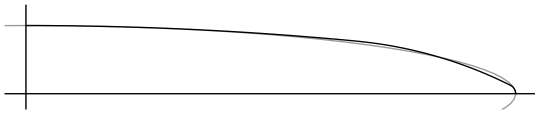


(a) Lockwood



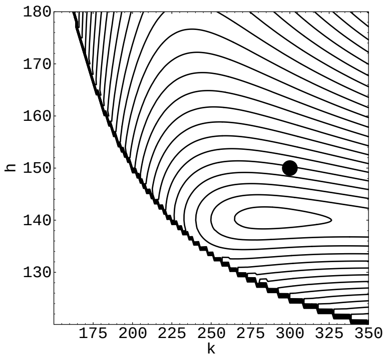
(b) Walker

(c) Chaplin

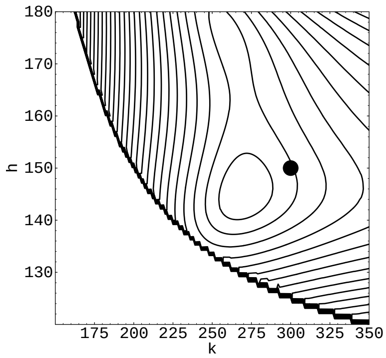


(d) LS fit

Figure 12: Five-centred arch approximations of an ellipse with $\frac{b}{a} = \frac{1}{7}$



(a) Lockwood's r_2 = arithmetic mean



(b) Walker's r_2 = geometric mean

Figure 13: Iso-contour lines of error plots obtained by modifying A_1 and A_3. Empty areas correspond to invalid solutions. The circle indicates the standard values of h and k to match the ellipse's curvature.

(passing them thorough the ellipse's vertices and matching its curvature there too) and considered different choices for the central arc. However, there is no reason to suppose that the values of r_1 and r_3 are particularly suitable in terms of minimising the distance (whether mean or maximum) between the ellipse and the circular arcs. This can be illustrated by fixing r_2 to a function of r_1 and r_3 while the latter two are varied under the constraint that the arcs are C^1. The error plots in figure 13 show that for all the choices of r_2 that we have considered in this paper (Chaplin's and the LS fit radii lie within the range of Lockwood and Walker's radii) more accurate approximations can be obtained by reducing h and k (i.e. decreasing r_1 and increasing r_3).

4 Perimeter Estimates

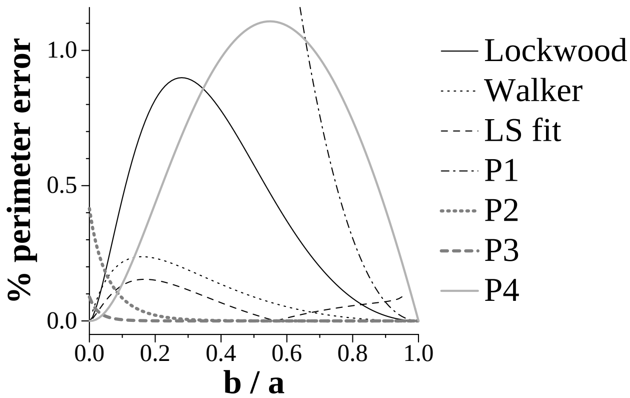


Figure 14: Percentage error in perimeter estimate for ellipse with $a = 1$

As a final analysis, the five-centred arches' perimeters are calculated based on the subtended angle θ_i of each arc A_i in a quadrant

$$4\left(\theta_1 r_1 + \theta_2 r_2 + \theta_3 r_3\right)$$

and their accuracies are compared against the true perimeter and also three approximations, the first being relatively crude, the second is an often used and effective one by Ramanujan, while the third is the best from Almkkvist and Berndt's extensive list of approximations [1]:

$$\begin{aligned} P_1 &= \sqrt{2}\sqrt{a^2+b^2}\pi \\ P_2 &= \left(3\left(a+b\right) - \sqrt{\left(3a+b\right)\left(a+3b\right)}\right)\pi \\ P_3 &= \frac{256 - 48\lambda^2 - 21\lambda^4}{256 - 112\lambda^2 + 3\lambda^4}\pi(a+b) \end{aligned}$$

where $\lambda = \frac{a-b}{a+b}$. In addition a new perimeter estimator has been devised with particular attention to elongated ellipses:

$$P_4 = 4a - \frac{b^2\left(2\pi - 4\right)\ln(\frac{b}{a})}{a-b}.$$

Both the general and particular solutions to the arches' perimeters are rather cumbersome and so are not included here with the exception of Lockwood's [9]

$$\frac{b^2}{a}\tan^{-1}\frac{a}{b} + \frac{a^2}{b}\tan^{-1}\frac{b}{a}.$$

The perimeter estimate errors with respect to the true ellipse perimeter are plotted in figure 14 for a range of eccentricities. P_1 is very much worse than the arches' estimates, P_2 is better for moderate eccentricities, while P_3 is better yet and gives relatively good results even for highly elongated ellipses. Nevertheless, the percentage errors for all the above monotonically increase with increasing eccentricity. In contrast, the arches' percentage errors peak at around $\frac{b}{a}$ equal to 0.2 or 0.3 but their perimeter estimates converge to the correct values of $4a$ and $2\pi a$ at $\frac{b}{a} = 0$ and $\frac{b}{a} = 1$ respectively. While P_1, P_2, and P_3 provide correct solutions for circles, for $\frac{b}{a} = 0$ they yield $\sqrt{2}\pi a \approx 4.443a$, $(3-\sqrt{3})\pi a \approx 3.983a$, and $\frac{187}{147}\pi a \approx 3.996a$

respectively. Although P_4 performs very well for highly elongated ellipses its errors elsewhere are rather high. The ranking of the arches' estimates is the least squares fit, followed by Walker's and then Lockwood's methods. Note however that for almost round ellipses the least squares fit does poorly as noted previously.

5 Concluding Remarks

While the approximation of an ellipse by circular arcs appears to be a simple problem, investigation uncovers a wealth of possibilities. This paper shows that, in combination with some additional common constraints, the two most common constraints – matching all the circular arcs to the ellipse's curvature at the point of contact, and maintaining tangent continuity at the joins – are incompatible. However, it was shown that Walker's proposal (setting the inner arc's radius equal to the geometric mean of the radii of the two outer arcs) provides a good tradeoff between the two constraints. Further improvements were made using numerical estimation, enabling extremely good approximations.

References

[1] G. Almkkvist and B. Berndt. Gauss, Landen, Ramanujan, the arithmetic mean, ellipses, π, and the *Ladies Diary*. *The American Mathematical Monthly*, 95:585–608, 1988.

[2] H.C. Browning. *The Principles of Architectural Drawing*. Watson-Guptill Publications, 1996.

[3] T.K. Chaplin. Methods of constructing ellipses and parabolae. *Mathematical Gazette*, 29:12–14, 1945.

[4] J.C. Golvin. *L'Amphitheatre Romain*. Boccard, 1988.

[5] V. Hart and P. Hicks, editors. *Sebastiano Serlio on Architecture : Books I-V of "Tutte L'Opere D'Architettura et Prospetiva"*. Yale Univ. Press, 1996.

[6] F. Hussain and M. L. V. Pitteway. Rasterizing the outlines of fonts. In *Third International Conference on Raster Imaging and Digital Typography*, pages 171–182, 1993.

[7] P. Karow. *Digital Formats for Typefaces*. URW Verlag, 1987.

[8] E.H. Lockwood. Note 891. *Mathematical Gazette*, 14:136–137, 1928.

[9] E.H. Lockwood. Note 1045. Length of ellipse. *Mathematical Gazette*, 16:269–270, 1932.

[10] E.H. Lockwood. *A Book of Curves*. Cambridge University Press, 1961.

[11] A. Lodge. Note 912. *Mathematical Gazette*, 14:270, 1928.

[12] P.L. Rosin. Analysing error of fit functions for ellipses. *Pattern Recognition Letters*, 17(14):1461–1470, 1996.

[13] P.L. Rosin. A survey and comparison of traditional piecewise circular approximations to the ellipse. *Computer Aided Geometric Design*, 16(4):269–286, 1999.

[14] A. Thom. *Megalithic Sites in Britain*. Clarendon Press, 1967.

[15] A.W. Walker. Note 2402. approximating to an ellipse by circular arcs. *Mathematical Gazette*, 38:123, 1954.