The locus of centers of ellipses inscribed in quadrilaterals

Alan Horwitz
Penn State University
25 Yearsley Mill Rd.
Media, PA 19063
alh4@psu.edu
8/11/03

Introduction

Let $R$ be a four–sided convex polygon in the $xy$ plane. A problem often referred to in the literature as Newton’s problem, was to determine the locus of centers of ellipses inscribed in $R$. By inscribed we mean that the ellipse lies inside $R$ and is tangent to each side of $R$. Chakerian(1) gives a partial solution of Newton’s problem using orthogonal projection, which is the solution actually given by Newton, which we state as

**Theorem 1** Let $M_1$ and $M_2$ be the midpoints of the diagonals of $R$. Then if $E$ is an ellipse inscribed in $R$, the center of $E$ must lie on $Z$, the open line segment connecting $M_1$ and $M_2$.

However, Theorem 1 does not really give the precise locus of centers of ellipses inscribed in $R$. It is stated in (2, pp. 217–219) that the locus of centers of ellipses inscribed in $R$ actually equals $Z$, but Newton only proved that the center of $E$ must lie on $Z$, as is noted in (1). Indeed, it is not even clear that an ellipse exists which is inscribed in $R$, let alone whether every point of $Z$ is the center of such an ellipse. The main result of this note is that it is indeed the case that every point of $Z$ is the center of an ellipse inscribed in $R$. This result was actually proved by the author in (3, Theorem 11), but the approach given here is decidedly different and much shorter and more succinct. In addition, we are also able to prove that there is a unique ellipse of maximal area inscribed in $R$. While it is perhaps possible to prove these results using orthogonal projection, we use, instead, a theorem of Marden(4, Theorem 1) relating the foci of an ellipse tangent to the lines thru the sides of a triangle and the zeros of a partial fraction expansion. We state the part we shall use here.

**Theorem 2** (Marden): Let $F(z) = \frac{t_1}{z - z_1} + \frac{t_2}{z - z_2} + \frac{t_3}{z - z_3}$, $t_1 + t_2 + t_3 = 1$, and let $Z_1$ and $Z_2$ denote the zeros of $F(z)$. Let $L_1, L_2, L_3$ be the line segments
connecting \( z_2, z_3, z_1, z_3, \) and \( z_1, z_2, \) respectively. If \( t_1 t_2 t_3 > 0 \), then \( Z_1 \) and \( Z_2 \) are the foci of an ellipse, \( E \), which is tangent to \( L_1, L_2, \) and \( L_3 \) in the points \( \zeta_1, \zeta_2, \zeta_3, \) where

\[
\zeta_1 = \frac{t_2 z_3 + t_3 z_1}{t_2 + t_3}, \quad \zeta_2 = \frac{t_1 z_3 + t_3 z_1}{t_1 + t_3}, \quad \zeta_3 = \frac{t_1 z_2 + t_2 z_1}{t_1 + t_2},
\]

respectively.

**Main Result**

**Theorem 3** Let \( R \) be a four–sided convex polygon in the \( xy \) plane and let \( M_1 \) and \( M_2 \) be the midpoints of the diagonals of \( R \). Let \( Z \) be the open line segment connecting \( M_1 \) and \( M_2 \). If \( (h, k) \in Z \) then there is a unique ellipse with center \((h, k)\) inscribed in \( R \).

We shall now prove Theorem 3 for the case when no two sides of \( R \) are parallel. Such a quadrilateral is sometimes called a trapezium. Our methods extend easily to the case when exactly two sides of \( R \) are parallel, that is, when \( R \) is a trapezoid. Of course, if \( R \) is a parallelogram, then the midpoints of the diagonals coincide, and the line segment \( Z \) is just a point. Since ellipses, tangent lines to ellipses, and four–sided convex polygons are preserved under affine transformations, we may assume that the vertices of \( R \) are \((0, 0), (1, 0), (0, 1), \) and \((s, t)\) for some real numbers \( s \) and \( t \). Let \( I \) denote the open interval between \( \frac{1}{2} \) and \( \frac{1}{2} s \). Then \( M_1 = \left( \frac{1}{2}, \frac{1}{2} \right), M_2 = \left( \frac{1}{2} s, \frac{1}{2} t \right), \) and the equation of the line thru \( M_1 \) and \( M_2 \) is

\[
y = L(x) = \frac{1}{2} \frac{s - t + 2x(t - 1)}{s - 1}, \quad x \in I
\]

Since \( R \) is convex, four–sided and no two sides of \( R \) are parallel, it follows easily that

\[
s > 0, t > 0, s + t > 1, \text{ and } s \neq 1 \neq t
\]

We shall need the following lemmas.

**Lemma 4** If \( h \in I \) and \( s + t > 1 \), then \( s + 2h(t - 1) > 0 \)

**Proof.** If \( t > 1 \), then \( s, h, \) and \( t - 1 \) are all positive. If \( t \leq 1 \) and \( s \geq 1 \), then

\[
I = \left( \frac{1}{2} s, \frac{1}{2} \right) \Rightarrow s + 2h(t - 1) \geq s - 2h > 0. \quad \text{Finally, if } t \leq 1 \text{ and } s \leq 1 \text{, then}
\]

\[
I = \left( \frac{1}{2} s, \frac{1}{2} \right) \Rightarrow s + 2h(t - 1) > s + t - 1 > 0. \quad \Box
\]

We leave the proof of the next lemma to the reader.

**Lemma 5** Let \( E_1 \) and \( E_2 \) be ellipses with the same foci. Suppose also that \( E_1 \) and \( E_2 \) pass through a common point, \( z_0 \). Then \( E_1 = E_2 \).

**Proof of Theorem 3** Let \( L_1: y = 0, L_2: x = 0, L_3: y = \frac{t}{s - 1}(x - 1), \) and \( L_4: y = 1 + \frac{t - 1}{s}x \) denote the lines which make up the boundary of...
R. $L_1, L_2,$ and $L_3$ form a triangle, $T_1$, whose vertices are the complex points $z_1 = 0$, $z_2 = 1$, and $z_3 = -\frac{t}{s} - i$. $L_1, L_2,$ and $L_4$ form a triangle, $T_2$, whose vertices are the complex points $w_1 = 0$, $w_2 = i$, and $w_3 = -\frac{s}{t}$. First, we want to find ellipses $E_1$ and $E_2$ tangent to $L_1, L_2,$ and $L_3$, and to $L_1, L_2,$ and $L_4$, respectively. We shall use Theorem 2 so that $E_1$ has foci $Z_1$ and $Z_2$, which are the zeros of $F(z) = \frac{t_1}{z} + \frac{t_2}{z-1} + \frac{t_3}{z + \frac{1}{t} i}$, and $E_2$ has foci $W_1$ and $W_2$, which are the zeros of $G(z) = \frac{s_1}{z} + \frac{s_2}{z-i} + \frac{1-s_1-s_2}{z + \frac{1}{t} i}$. To guarantee that $E_1$ and $E_2$ are ellipses, we require, by Theorem 2 that $s_1 s_2 s_3 > 0$ and $t_1 t_2 t_3 > 0$, where $s_3 = 1 - s_1 - s_2$ and $t_3 = 1 - t_1 - t_2$. For example, let $s = 3$, $t = 2$, $t_1 = -\frac{1}{4}$, $t_2 = \frac{3}{2}$, $s_1 = \frac{1}{3}$, and $s_2 = \frac{1}{2}$. Then $t_1 t_2 t_3 = \frac{3}{32} > 0$ and $s_1 s_2 s_3 = \frac{3}{36} > 0$. The foci of $E_1$ are approximately $Z_1 = -0.1957 - 0.0496 i$ and $Z_2 = -0.3043 - 1.2004 i$. Note that $E_1$ is not inscribed in $T_1$ since not all of the $t_j$’s are positive (see Figure 1). The foci of $E_2$ are approximately $W_1 = -0.0159 + 0.0191 i$ and $W_2 = -0.2484 + 0.981 i$. Note that $E_2$ is inscribed in $T_2$ since all of the $s_j$’s are positive (see Figure 2). Assume now that $(h, k) \in Z$, or equivalently, that $k = L(h), h \in I$. We want $E_1$ and $E_2$ each to have center $(h, k)$. The center, $C_1$, of $E_1$ is $\frac{1}{2}(Z_1 + Z_2)$. A simple computation shows that $C_1 = \frac{1}{2(s-1)}(it(t_1 + t_2) + (s-1)(t_2 - 1))$, which, upon taking real and imaginary parts yields $C_1 = \left(\frac{1}{2} - \frac{1}{2} t_2, -\frac{1}{2} t_1 + t_2\right)$. Similarly, the center of $E_2$ is $C_2 = \left(-\frac{1}{2} s_1 + s_2, \frac{1}{2}(s_2 - 1)\right)$. We actually do not require these explicit formulas for $C_1$ and $C_2$. However, solving $(h, k) = \left(\frac{1}{2} - \frac{1}{2} t_2, -\frac{1}{2} t_1 + t_2\right)$ for $t_1$ and $t_2$ shows that the center of $E_1$ is $(h, k)$ if and only if

$$t_1 = 2h - 1 - 2k \frac{s-1}{t}, \quad t_2 = 1 - 2h. \quad (1)$$

Similarly, solving $(h, k) = \left(-\frac{1}{2} s_1 + s_2, \frac{1}{2}(s_2 - 1)\right)$ for $s_1$ and $s_2$ shows that the center of $E_2$ is $(h, k)$ if and only if

$$s_1 = 2k - 1 - 2h \frac{t-1}{s}, \quad s_2 = 1 - 2k \quad (2)$$

So given $(h, k) \in Z$, let $s_1, s_2, t_1, t_2$ be defined by (1) and (2). Substituting $k = L(h)$ into (1) and (2) yields $t_1 t_2 t_3 = (s + 2h(t-1)) \frac{(s-2h)^2}{t^3} > 0$ since $h \in I$ and by Lemma 3 similarly, $s_1 s_2 s_3 = (s + 2h(t-1))(2h-1)(s-2h) \frac{(t-1)^2}{s^2(s-1)^2}$.
that case, area(ABC) Chakerian’s result assumes that the point P lies in a four–sided polygon, with three vertices A, B, and C, while our result assumes that P lies outside ABC. In that case, area(ABC) = area(CPA) + area(APB) − area(BPC). The details of the proof are similar.

**Lemma 6** Given a triangle ABC and a point P \( \notin \partial (ABC) \), let \( \alpha = \text{area}(BPC) \), \( \beta = \text{area}(CPA) \), and \( \gamma = \text{area}(APB) \). Let \( L_1 \), \( L_2 \), and \( L_3 \) be the three lines thru the sides of ABC, and let E be an ellipse with center P which is tangent to \( L_1 \), \( L_2 \), and \( L_3 \). If \( \sigma = \frac{1}{2} (\alpha + \beta + \gamma) \), then \( \text{area}(E) = \frac{4\pi}{\text{area}(ABC)} \sqrt{\sigma (\sigma - \alpha)(\sigma - \beta)(\sigma - \gamma)} \)

Now let \( A_E = \text{area of an ellipse E inscribed in R} \). We want to maximize and/or minimize \( A_E \) as a function of \( h \), where \( (h, L(h)) \) denotes the center of E. We discuss the case when no two sides of R are parallel. Let \( A = (0, 0), B = (1, 0), C = \left(0, \frac{t}{s-1}\right)\), which are the vertices of the triangle we
earlier called \( T_1 \). Then area(\( ABC \)) = \( \frac{1}{2} \frac{t}{|s-1|} \), and since \( E \) is inscribed in \( ABC \), we can apply Lemma \([6]\) with \( P = (h, k) \). Substituting \( k = L(h) \) yields
\[
\sigma(\sigma - \alpha)(\sigma - \beta)(\sigma - \gamma) = \frac{1}{256} t^2 (-1 + 2h) (s + 2ht - 2h) \frac{s - 2h}{(s - 1)^3}.
\]
By Lemma \([6]\)
\[
A_E = \frac{\pi}{2} \frac{\sqrt{2h - 1}}{\sqrt{(2h - 1) (s + 2h(t - 1)) (s - 2h)}}.
\]
Thus we want to optimize
\[
A(h) = (s - 2h) (2h - 1) (s + 2h(t - 1)), \quad h \in I.
\]
Now \( A(1/2) = A(s/2) = 0 \), and \( A(h) \geq 0 \) for \( h \in I \) by Lemma \([6]\). Hence \( A'(h_0) = 0 \) for some \( h_0 \in I \) with \( A(h_0) \) a local maximum, and \( A(h) \) does not attain its global minimum on \( I \). Also, \( A(h_0) \) must be the only local maximum of \( A(h) \) on \( I \), else \( A'(h) \) would have three zeros in \( I \). Thus \( A(h_0) \) is the global maximum of \( A(h) \) on \( I \). Since ratios of areas of ellipses are preserved under affine transformations, we have proven

**Theorem 7** Let \( R \) be any given four-sided convex polygon in the xy plane. Then there is a unique ellipse of maximal area inscribed in \( R \). There is no ellipse of minimal area inscribed in \( R \).

**Example:** Take \( s = 4, t = 2 \), so that \( R \) has vertices \((0,0), (1,0), (0,1), (4,2)\). Then the maximal area ellipse has center \( \left( \frac{4}{3}, \frac{7}{9} \right) \).

**Hyperbolas**

Using our earlier notation, let \( X \) be the open line segment which is the part of \( L \) lying inside \( R \), where \( L \) is the line thru the midpoints of the diagonals. If \((h, k) \in X - Z - M_1 - M_2 \), it is natural to think that there should be a hyperbola, \( H \), with center \((h, k)\), which is tangent to each line making up the boundary of \( R \). This is actually correct, but only if one considers an asymptote of \( H \) to be tangent to \( H \) (at infinity, of course). \(^1\) This is not hard to prove using the methods of this paper. An asymptote of \( H \) can arise when employing Theorem \([2]\) since it is possible for one of \( t_i + t_j, j \neq i \), to be 0.

**References**

[1] G. D. Chakerian, A Distorted View of Geometry, MAA, Mathematical Plums, Washington, DC, 1979, 130-150.

[2] Heinrich Dörrie: 100 Great Problems of Elementary Mathematics, Dover, New York, 1965.

[3] Alan Horwitz, “Finding ellipses and hyperbolas tangent to two, three, or four given lines”, Southwest Journal of Pure and Applied Mathematics 1(2002), 6-32.

[4] Morris Marden, "A note on the zeros of the sections of a partial fraction, Bulletin of the AMS 51 (1945), 935–940.

\(^1\) This was also proven in \([3]\), but the statement there is not quite correct since this author omitted the case where the "tangent line" is an asymptote.