Extremal area ellipses of a convex quadrilateral

JOHN R. SILVESTER

1. *Introduction*

In this paper we are going to generalise some standard results about the ellipses of least or greatest area, respectively circumscribing or inscribed in a triangle, to the corresponding ellipses for a convex quadrilateral.

Given ΔABC let D, E, F be the midpoints of BC, CA, AB respectively; then the medians AD, BE, CF concur at the centroid G of ΔABC. There is an ellipse with centre G, touching the sides of ΔABC at their respective midpoints. It is known as the *midpoint ellipse* or *Steiner inellipse* of ΔABC, and it is the ellipse of greatest area inscribed in ΔABC. Since $AG : GD = 2 : 1$, and similarly for the other medians, a central dilation (an enlargement) with centre G and scale factor -2 (or, equivalently, 2) applied to the Steiner inellipse produces a similar ellipse, also with centre G, and passing through A, B and C. This is the *Steiner circumellipse* of ΔABC, and it is the ellipse of least area circumscribing ΔABC. See Figure 1. For more detail, see [1, 2] and the references given there.

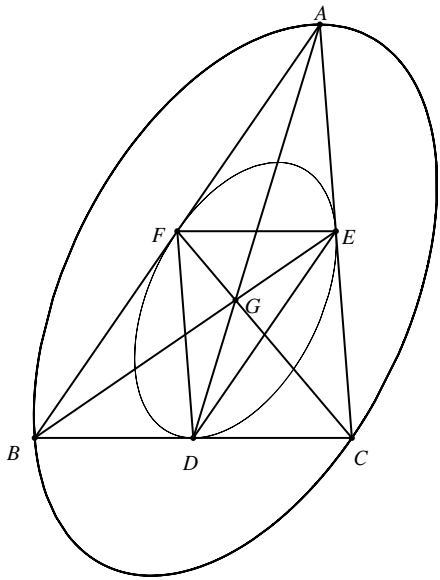


FIGURE 1

Notice that the Steiner inellipse of ΔABC is the Steiner circumellipse of ΔDEF, so that it is simultaneously the ellipse of least area circumscribing ΔDEF and the ellipse of greatest area inscribed in ΔABC.

In this paper we shall attempt to generalise the last remark to convex quadrilaterals; see the theorem in section 3. First we must learn how to

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press

recognise the relevant ellipses. Given a convex quadrilateral, the problem of finding the circumscribed ellipse of least area was discussed in [3], where, however, the main thrust of the article was to find a circumscribed ellipse of least eccentricity (the 'most nearly circular' circumscribed ellipse), a problem that goes back to Steiner. In [3] it is proved that a convex quadrilateral always has a unique circumscribed ellipse of least area, but it is remarked that 'there does not appear to be a nice characterisation' of this ellipse. Here, we shall attempt to find such a characterisation. It turns out to be relatively straightforward, given an ellipse Γ, to find all the (convex) quadrilaterals inscribed in Γ for which Γ is the circumscribed ellipse of least area. It is somewhat harder, given a convex quadrilateral, to find its circumscribed ellipse of least area.

In passing, we note the following curious fact. Taking ΔABC in the complex plane, let A, B, C be represented by the complex numbers a, b, c, and put $f(z) = (z - a)(z - b)(z - c)$. Then $f'(z)$ is a quadratic, so vanishes at two points. Where are they? It turns out that they are the foci of the Steiner inellipse of ΔABC. This is set as an exercise in Hardy [4], who gives a reference to a proof by Césaro.

2. *Areas*

The conic $\Gamma : ax^2 + 2hxy + by^2 + 2gx + 2fy + c = 0$ can be written in matrix form as $\mathbf{x}\mathbf{M}\mathbf{x}^{\mathrm{T}} = 0$, where $\mathbf{x} = (x, y, 1)$, T stands for transpose, and

$$\mathbf{M} = \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix}.$$

This conic is an ellipse if $\det \mathbf{M} \neq 0$ and $ab - h^2 > 0$, imaginary if $a \det \mathbf{M} > 0$ and real if $a \det \mathbf{M} < 0$.

It is easy to see that applying a translation $(x, y) \to (x + \alpha, y + \beta)$ to Γ does not alter either $\det \mathbf{M}$ or $ab - h^2$, and neither does applying a rotation $(x, y) \to (x \cos\theta - y \sin\theta, x \sin\theta + y \cos\theta)$. So, if Γ is an ellipse, apply a suitable isometry to put it in the standard form $a_1x^2 + b_1y^2 + c_1 = 0$. Now the area of $\dfrac{x^2}{\alpha^2} + \dfrac{y^2}{\beta^2} = 1$ is $\pi\alpha\beta$, so the area Δ of Γ satisfies

$$\left(\frac{\Delta}{\pi}\right)^2 = \frac{c_1^2}{a_1 b_1} = \frac{(a_1 b_1 c_1)^2}{(a_1 b_1)^3} = \frac{(\det \mathbf{M})^2}{(ab - h^2)^3}. \tag{1}$$

Notice that this quantity is unaffected if we multiply the equation of Γ through by a non-zero constant; also note that it would be negative if Γ were a hyperbola, infinite if it were a parabola, and either zero or indeterminate if it were a line pair.

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press

Now a line has *line coordinates* $\mathbf{u} = (u, v, w)$. So, for example, the line $2x + 3y + 5 = 0$ has line coordinates (2, 3, 5). The conic Γ can be given as a conic envelope by its line-equation $\mathbf{uNu}^{\mathrm{T}} = 0$, where $\mathbf{N} = \mathbf{M}^{-1}$. To see this, note that for $\mathbf{x}$ on Γ, the tangent at $\mathbf{x}$ (the polar of $\mathbf{x}$) has line coordinates $\mathbf{u} = \mathbf{xM}$, and then

$$0 = \mathbf{xMx}^{\mathrm{T}} = (\mathbf{xM})\mathbf{M}^{-1}(\mathbf{Mx}^{\mathrm{T}}) = \mathbf{uM}^{-1}\mathbf{u}^{\mathrm{T}} = \mathbf{uNu}^{\mathrm{T}}.$$

We then have

$$\left(\frac{\Delta}{\pi}\right)^2 = \frac{(\det \mathbf{M})^2}{(ab - h^2)^3} = \frac{(\det \mathbf{N})^{-2}}{((\det \mathbf{N})^{-1}\mathbf{N}_{33})^3} = \frac{\det \mathbf{N}}{(\mathbf{N}_{33})^3} \tag{2}$$

where $\mathbf{N}_{33}$ is the (3, 3)-entry of $\mathbf{N}$. For the general theory of conic loci, conic envelopes, and pole and polar, see [5].

3. *Minima and maxima*

Recall that an *affine transformation* is a transformation of the form

$$(x, y) \to (a_{11}x + a_{12}y + b_1, a_{21}x + a_{22}y + b_2),$$

where the a_{ij} and b_i are constants; or, in matrix-vector form, $\mathbf{v} \to \mathbf{Av} + \mathbf{b}$, where

$$\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad \text{and} \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

We require $\det \mathbf{A} \neq 0$, in order to ensure that the transformation is bijective, the inverse map then being given by $\mathbf{v} \to \mathbf{A}^{-1}(\mathbf{v} - \mathbf{b})$, also an affine transformation. We shall need the following well-known properties of affine transformations:

(i) Affine transformations send lines to lines, and conics to conics; indeed, the image of an ellipse/hyperbola/parabola is another ellipse/hyperbola/parabola, respectively. (Here, as always, a circle is just a particular sort of ellipse, of course.) And given two conics, there is an affine transformation sending the one to the other if and only if both are ellipses, or both are hyperbolas, or both are parabolas.

(ii) Although in general ratios of distances are *not* preserved, nonetheless ratios of distances on any particular line (or pair of parallel lines) *are* preserved. (Although in general, angles are *not* preserved.)

(iii) Ratios of areas are preserved; in fact, areas get multiplied by $|\det \mathbf{A}|$.

(iv) Any two triangles are affinely equivalent: given $\triangle ABC$ and $\triangle A'B'C'$, there is an affine transformation sending A to A', B to B' and C to C'.

Given an ellipse circumscribing a quadrilateral, we seek conditions for the ellipse to have least possible area for that quadrilateral. To facilitate the calculation, we apply an affine transformation to turn the ellipse into the unit circle $x^2 + y^2 = 1$, with inscribed quadrilateral $PQRS$. The ratio (area of ellipse) : (area of quadrilateral) is preserved by this transformation. Let the

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press

lines PQ, RS have equations $a_i x + b_i y = c_i$, $i = 1, 2$, so that the line-pair $(PQ)(RS)$ has equation $(a_1x + b_1y - c_1)(a_2x + b_2y - c_2) = 0$. Then the conic $\Gamma_\lambda : (x^2 + y^2 - 1) + \lambda(a_1x + b_1y - c_1)(a_2x + b_2y - c_2) = 0$ circumscribes $PQRS$, and any circumscribed conic can be put in this form.

The matrix of Γ_λ is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} + \lambda \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix} = \begin{pmatrix} 1 + \lambda a & \lambda h & \lambda g \\ \lambda h & 1 + \lambda b & \lambda f \\ \lambda g & \lambda f & \lambda c - 1 \end{pmatrix}$$

where $a = a_1a_2$, $h = \frac{1}{2}(a_1b_2 + a_2b_1)$, $g = -\frac{1}{2}(a_1c_2 + a_2c_1)$, etc. From (1), working to first order in λ, we have, for Γ_λ, that $(\Delta / \pi)^2$ is

$$\frac{((1 + \lambda a)(1 + \lambda b)(\lambda c - 1))^2}{((1 + \lambda a)(1 + \lambda b))^3} = \frac{(\lambda c - 1)^2}{(1 + \lambda a)(1 + \lambda b)} = 1 - \lambda(a + b + 2c).$$

We want conditions on $PQRS$ to ensure that our circle is the circumscribed ellipse of least area. We need the following lemma, which will be proved in the next section:

Lemma 1: For ellipses Γ_λ circumscribing $PQRS$, there is just one value of λ for which the λ-derivative of $(\Delta / \pi)^2$ vanishes.

That being so, the condition we seek is that this derivative vanishes when $\lambda = 0$. So the condition is $a + b + 2c = 0$, or $a_1a_2 + b_1b_2 + 2c_1c_2 = 0$.

Now the line coordinates of PQ are $(a_1, b_1, -c_1)$; so the pole K of this with respect to the circle (that is, the point K where the tangents to the circle at P and Q meet) has homogeneous coordinates (a_1, b_1, c_1), or cartesian coordinates $\left(\frac{a_1}{c_1}, \frac{b_1}{c_1}\right)$. Apply a central dilation to this, centre O and scale-factor $-\frac{1}{2}$, to obtain the point $\left(-\frac{a_1}{2c_1}, -\frac{b_1}{2c_1}\right)$. This lies on $RS : a_2x + b_2y = c_2$ if $a_2\left(-\frac{a_1}{2c_1}\right) + b_2\left(-\frac{b_1}{2c_1}\right) = c_2$, that is, if $a_1a_2 + b_1b_2 + 2c_1c_2 = 0$, which is the same condition as before. So the condition for minimality is this: if $K' = KO \cdot RS$ (the meet), then $KO : OK' = 2 : 1$.

Since affine transformations preserve ratios of lengths on a line, and ratios of areas, we have the following result (see Figure 2):

Proposition 1: Let Γ be an ellipse, centre O, circumscribing the quadrilateral $PQRS$, and let the pole of PQ with respect to Γ be K, that is, the tangents to Γ at P and Q meet at K. Let $K' = KO \cdot RS$. Then Γ is the ellipse of least area circumscribing $PQRS$ if, and only if, $KO : OK' = 2 : 1$.

As a consequence of this proposition, there are six equivalent conditions involving the 2 : 1 ratio. For example, if L is the pole of QR and $L' = LO \cdot SP$, we must also have $LO : OL' = 2 : 1$. To see this directly, go back to the unit circle and use the rational parametrisation $\left(\frac{1 - t^2}{1 + t^2}, \frac{2t}{1 + t^2}\right)$. If P, Q, R, S have parameters p, q, r, s respectively, it is an exercise for the reader to

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press

show that the condition $KO : OK' = 2 : 1$ is equivalent to the symmetrical condition $3pqrs + pq + pr + ps + qr + qs + rs + 3 = 0$.

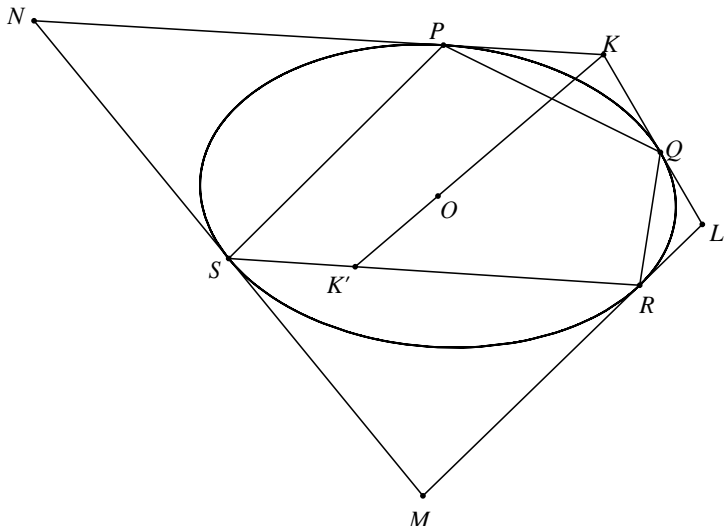


FIGURE 2

Note that there may be an ellipse containing $PQRS$, rather than circumscribing it, of yet smaller area, for example the Steiner circumellipse of ΔPQR if S happens to be inside it.

We now turn to the problem of maximising the area of an ellipse inscribed in a convex quadrilateral $KLMN$. The calculation that follows is remarkably similar to the preceding one.

So, apply an affine transformation so that the ellipse becomes the circle $x^2 + y^2 = 1$ or, as an envelope, $u^2 + v^2 = w^2$. Let K, M have homogeneous coordinates (a_i, b_i, c_i), $i = 1, 2$, so that the point-pair $(K)(M)$ has equation $(a_1u + b_1v + c_1w)(a_2u + b_2v + c_2w) = 0$. Then the conic envelope $\Gamma_\lambda : (u^2 + v^2 - w^2) + \lambda(a_1u + b_1v + c_1w)(a_2u + b_2v + c_2w) = 0$ touches the sides of $KLMN$, and every inscibed conic of the quadrilateral can be put in this form.

The matrix of Γ_λ is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} + \lambda \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix} = \begin{pmatrix} 1 + \lambda a & \lambda h & \lambda g \\ \lambda h & 1 + \lambda b & \lambda f \\ \lambda g & \lambda f & \lambda c - 1 \end{pmatrix}$$

where $a = a_1a_2$, $h = \frac{1}{2}(a_1b_2 + a_2b_1)$, $g = \frac{1}{2}(a_1c_2 + a_2c_1)$, etc.

From (2), working to first order in λ, we have, for Γ_λ, that (Δ / π^2) is

$$\frac{(1 + \lambda a)(1 + \lambda b)(\lambda c - 1)}{(\lambda c - 1)^3} = \frac{(1 + \lambda a)(1 + \lambda b)}{(\lambda c - 1)^2} = 1 + \lambda(a + b + 2c).$$

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press

We want conditions on *KLMN* to ensure that our circle is the inscribed ellipse of greatest area. We need the following lemma, which will be proved in the next section:

Lemma 2: For ellipses Γ_λ inscribed in *KLMN*, there is just one value of λ for which the λ-derivative of $(\Delta/\pi)^2$ vanishes.

That being so, the condition we seek is that this derivative vanishes when $\lambda = 0$. So the condition is $a + b + 2c = 0$, or $a_1a_2 + b_1b_2 + 2c_1c_2 = 0$. This looks the same as before! Now if the points of contact of the circle with *KL*, *LM*, *MN*, *NK* are *Q*, *R*, *S*, *P*, then *PQ*, *RS* are the polars of *K*, *M* and thus have equations $a_ix + b_iy = c_i, i = 1, 2$, which are the same as before.

Since affine transformations preserve ratios of lengths on a line, and ratios of areas, we have the following result (see Figure 2 again):

Proposition 2: Let Γ be an ellipse, centre *O*, inscribed in the quadrilateral *KLMN*, and let the points of contact of Γ with *LM*, *MN* be *R*, *S* respectively. Let $K' = KO{\cdot}RS$. Then Γ is the ellipse of greatest area inscribed in *KLMN* if, and only if, $KO : OK' = 2 : 1$.

Note that there may be an ellipse *contained in KLMN*, rather than inscribed in it, of yet larger area, for example the Steiner inellipse of ΔMNX, if $X = LM{\cdot}NK$ and the segment *KL* is inside the triangle but does not meet the ellipse.

Combining Propositions 1 and 2, we have our main result:

Theorem: Let Γ be an ellipse circumscribing the convex quadrilateral *PQRS* and inscribed in the convex quadrilateral *KLMN*, where *Q*, *R*, *S*, *P* are on *KL*, *LM*, *MN*, *NK* respectively. Then the following conditions are equivalent:

1. Γ is the ellipse of least area circumscribing *PQRS*.
2. Γ is the ellipse of greatest area inscribed in *KLMN*.
3. If the centre of Γ is *O*, and $K' = KO{\cdot}RS$, then $KO : OK' = 2 : 1$.

4. *Proofs of the lemmas*

This section is rather long and technical and is essentially independent of the others, so it can be omitted by readers who are prepared to take Lemmas 1 and 2 on trust.

First then, Lemma 1. The set of conics $\mathcal{P} = \{\Gamma_\lambda : \lambda \in \mathbb{R} \cup \{\infty\}\}$ is called a *pencil* of conics; it consists of all the conics circumscribing *PQRS*. (Readers nervous about putting $\lambda = \infty$ should instead divide through by λ and then let $\lambda \to \infty$, which has the same effect.) For the general theory of pencils (and ranges, which we use below) see [5, pp.156-165].

Members of $\mathcal{P}$ are linear combinations of any two of its members; we have used the unit circle and the line-pair $(PQ)(RS)$ as a basis, but if we take a different pair of members, this is equivalent to using a new parameter $\mu = \dfrac{a_0\lambda + b_0}{c_0\lambda + d_0}$, where $a_0d_0 - b_0c_0 \neq 0$. It is easily seen that $\dfrac{d\mu}{d\lambda}$ never vanishes, so that we can use μ-derivatives instead of λ-derivatives to find extremals.

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press

If we rotate the figure about O so that PQ, RS are equally inclined to either axis, then the equation of $(PQ)(RS)$ will have no xy-term, and the equation of Γ_λ will take the form

$$\left(x^2 + y^2 - 1\right) + \lambda\left(ax^2 + by^2 + \text{linear terms}\right) = 0.$$

Putting $\lambda = -\frac{1}{b}, -\frac{1}{a}$ yields two parabolas (or, possibly, parallel line-pairs) in $\mathcal{P}$, which are of the form $x^2 + 2gx + 2fy = c$ and $y^2 - 2gx - 2fy = d$ for some (new) constants f, g, c, d, with $c + d = 1$, since adding their equations must give the unit circle. So every *ellipse* Γ_λ has the form

$$x^2 + 2gx + 2fy - c + \mu\left(y^2 - 2gx - 2fy - d\right) = 0,$$

for some $\mu > 0$, or

$$x^2 + \mu y^2 + 2(1-\mu)gx + 2(1-\mu)fy - (c + \mu d) = 0.$$

Completing the squares,

$$(x + (1-\mu)g)^2 + \mu\left(y + (\mu^{-1} - 1)f\right)^2 = c + \mu d + (1-\mu)^2\left(g^2 + \mu^{-1}f^2\right).$$

The area of this ellipse is $\pi F(\mu)$, where

$$F(\mu) = \mu^{-\frac{1}{2}}\left(c + \mu d + (1-\mu)^2\left(g^2 + \mu^{-1}f^2\right)\right) \tag{3}$$

$$= \mu^{\frac{3}{2}}g^2 + \mu^{\frac{1}{2}}\left(d + f^2 - 2g^2\right) + \mu^{-\frac{1}{2}}\left(c + g^2 - 2f^2\right) + \mu^{-\frac{3}{2}}f^2. \tag{4}$$

Note that $F(\mu)$ and $F'(\mu)$ are continuous for $\mu > 0$. Also, $F(\mu) \to \infty$ as $\mu \to \infty$ provided $g \neq 0$. But if $g = 0$ then the parabola $y^2 - 2gx - 2fy - d = 0$ degenerates into the parallel line-pair $(y-f)^2 = d + f^2$, distinct parallel lines, since P, Q, R, S are distinct. So then $d + f^2 > 0$, and from this it follows that $F(\mu) \to \infty$ as $\mu \to \infty$ in this case also. Likewise, $F(\mu) \to \infty$ as $\mu \to 0^+$, and a similar argument with the other parabola shows that this is the case even if $f = 0$. We deduce that there must be at least one value of $\mu > 0$ for which $F'(\mu) = 0$. By taking a suitable affine transformation we may assume this happens when $\mu = 1$, that is, when our ellipse is the unit circle.

From (3), it is clear that $F'(1) = \frac{1}{2}(d - c)$, so since $c + d = 1$ we must have $c = d = \frac{1}{2}$. Then, from (4),

$$F'(\mu) = \frac{3\mu^{\frac{1}{2}}}{2}g^2 + \frac{\mu^{-\frac{1}{2}}}{2}\left(\frac{1}{2} + f^2 - 2g^2\right) - \frac{\mu^{-\frac{3}{2}}}{2}\left(\frac{1}{2} + g^2 - 2f^2\right) - \frac{3\mu^{-\frac{5}{2}}}{2}f^2$$

$$= \tfrac{1}{4}\mu^{-\frac{5}{2}}(\mu - 1)\left(6\mu^2 g^2 + \mu\left(2f^2 + 2g^2 + 1\right) + 6f^2\right).$$

This is clearly negative for $0 < \mu < 1$ and positive for $1 < \mu$. So for $\mu > 0$, $F'(\mu) = 0$ if, and only if, $\mu = 1$, and $F(\mu)$ takes its least value when $\mu = 1$. This completes the proof of Lemma 1.

The reader might like to show that the centres of the conics of $\mathcal{P}$ lie on a certain rectangular hyperbola, the *centre-locus* of $\mathcal{P}$, with the ellipses of $\mathcal{P}$ having their centres on one branch, and the hyperbolas and line-pairs of $\mathcal{P}$ having their centres on the other.

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press

Now for the proof of Lemma 2. The concept dual to that of a pencil of conics, the set of conics through four points, is a *range* of conics, the set of conics touching four lines. It is tempting, therefore, to see Lemma 2 as being possibly the dual of Lemma 1, in some sense, since it concerns the range $\mathcal{R}$ of conics touching the lines KL, LM, MN and NK. However, things are not quite that simple, as the dual of an ellipse is not necessarily another ellipse, and there is no easy way to dualise the concept of area.

We shall see shortly that, in general, the centre-locus of a range is a straight line. Putting $V = KL{\cdot}MN$ and $W = LM{\cdot}NK$, we have that the point-pairs $(K)(M)$, $(L)(N)$ and $(V)(W)$ belong to $\mathcal{R}$, so that the centre-locus of $\mathcal{R}$ must be the line through the centres of these point-pairs, that is, through the midpoints O_1, O_2 and O_3 of KM, LN and VW respectively. However, in the special case where $KLMN$ is a parallelogram, we have $O_1 = O_2$ and O_3 is at infinity, so let us deal with this first. To simplify the algebra, we apply an affine transformation to make $KLMN$ a square; in fact, let us take $K = (1, 0)$, $L = (0, 1)$, $M = (-1, 0)$ and $N = (0, -1)$. Then the line-equations of the point-pairs $(K)(M)$ and $(L)(N)$ are $(u + w)(-u + w) = 0$ and $(v + w)(-v + w) = 0$ respectively, so the general member Γ of $\mathcal{R}$ has the form $w^2 - u^2 + \lambda(w^2 - v^2) = 0$, for $\lambda \in \mathbb{R} \cup \{\infty\}$. Changing the basis of $\mathcal{R}$, put $\lambda = \dfrac{1 - \mu}{1 + \mu}$, and Γ becomes $(1 + \mu)u^2 + (1 - \mu)v^2 - 2w^2 = 0$, whose area is given by $(\Delta/\pi)^2 = \frac{1}{4}(1 - \mu^2)$. So Γ is an ellipse if $|\mu| < 1$ and a hyperbola if $|\mu| > 1$; and $(\Delta/\pi)^2$ has a unique maximum, at $\mu = 0$, when Γ is the inscribed circle of $KLMN$. This completes the proof of Lemma 2 in this case.

So now assume $KLMN$ is *not* a parallelogram. Then $O_1 \neq O_2$, and obviously we cannot have all four vertices of $KLMN$ on O_1O_2, so assume that L is not on O_1O_2. Apply an affine transformation so that $O_1 = (0, 0)$, $O_2 = (1, 0)$ and $L = (1, 1)$. So $N = (1, -1)$, $K = (\alpha, \beta)$ and $M = (-\alpha, -\beta)$, for some α, β with $\alpha \neq 0$; in fact (swapping K and M if necessary), let us assume $\alpha > 0$. Then $KM{\cdot}LN = \left(1, \frac{\beta}{\alpha}\right)$, and since this point is between K and M we must in fact have $\alpha > 1$; and since it is between L and N we must also have $-1 < \frac{\beta}{\alpha} < 1$. Thus $\alpha^2 - 1 > 0$, and also $\alpha^2 - \beta^2 > 0$.

The line-equations of the point-pairs $(L)(N)$ and $(K)(M)$ are $(u + v + w)(u - v + w) = 0$ and $(\alpha u + \beta v + w)(\alpha u + \beta v - w) = 0$, respectively. Thus a typical member Γ of $\mathcal{R}$ is

$$(u + v + w)(u - v + w) + \lambda(\alpha u + \beta v + w)(\alpha u + \beta v - w) = 0,$$

for $\lambda \in \mathbb{R} \cup \{\infty\}$; or, replacing λ by $1 - \lambda^{-1}$ and multiplying through by λ,

$$\lambda(u+v+w)(u-v+w) + (\lambda-1)(\alpha u+\beta v+w)(\alpha u+\beta v-w) = 0,$$

for $\lambda \in \mathbb{R} \cup \{\infty\}$ (still). Multiplying out, Γ has equation

$$\big((\lambda-1)\alpha^2+\lambda\big)u^2 + \big((\lambda-1)\beta^2-\lambda\big)v^2 + w^2 + 2\lambda wu + 2(\lambda-1)\alpha\beta uv = 0.$$

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press

The matrix of this conic envelope is

$$\mathbf{N} = \begin{pmatrix} (\lambda - 1)\alpha^2 + \lambda & (\lambda - 1)\alpha\beta & \lambda \\ (\lambda - 1)\alpha\beta & (\lambda - 1)\beta^2 - \lambda & 0 \\ \lambda & 0 & 1 \end{pmatrix}.$$

Multiplying on the left by $\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$, we see that the centre of Γ (the pole of the line at infinity) is $(\lambda,\ 0)$, which is between O_1 and O_2 if, and only if, $0 < \lambda < 1$. And the centre-locus of $\mathcal{R}$ is indeed a line; in fact, the x-axis.

Now we know that the quantity $(\Delta/\pi)^2$ is positive for an ellipse, negative for a hyperbola, and zero (or indeterminate) for a degenerate conic. Further, $(\Delta/\pi)^2 = \det \mathbf{N} / \mathbf{N}_{33}^3$; but here we have cunningly arranged things so that $\mathbf{N}_{33} = 1$, and so $(\Delta/\pi)^2 = \det \mathbf{N}$. Then

$$\begin{aligned} \det \mathbf{N} &= \big((\lambda-1)\alpha^2+\lambda\big)\big((\lambda-1)\beta^2-\lambda\big) - (\lambda-1)^2\alpha^2\beta^2 - \lambda^2\big((\lambda-1)\beta^2-\lambda\big) \\ &= \lambda(1-\lambda)\big(\alpha^2-\beta^2\big) - \lambda^2 + \lambda^2(1-\lambda)\beta^2 + \lambda^3 \\ &= \lambda(1-\lambda)\big(\alpha^2-\beta^2+\lambda\big(\beta^2-1\big)\big) \\ &= \lambda(1-\lambda)\big((1-\lambda)\big(\alpha^2-\beta^2\big)+\lambda\big(\alpha^2-1\big)\big). \end{aligned}$$

If $KL \parallel MN$ then $\beta = 1$, and if $LM \parallel NK$ then $\beta = -1$, and in either case $\det \mathbf{N} = \lambda(1 - \lambda)\big(\alpha^2 - 1\big)$. This is zero for $\lambda = 0$ or 1, positive for $0 < \lambda < 1$, and negative otherwise, and it has a unique maximum at $\lambda = \frac{1}{2}$. This completes the proof of Lemma 2 when $KLMN$ is a trapezium.

If $KLMN$ is not a trapezium, then $\beta^2 - 1 \neq 0$, so $\det \mathbf{N}$ is a cubic in λ, with distinct zeros at 0, 1, and $\gamma = \dfrac{\alpha^2 - \beta^2}{1 - \beta^2}$. So here $O_3 = (\gamma,\ 0)$. (If $1 - \beta^2 > 0$ then $\gamma > 1$ and O_2 is between O_1 and O_3; and if $1 - \beta^2 < 0$ then $\gamma < 0$ and O_1 is between O_2 and O_3.) Now a cubic with distinct real zeros has one local maximum and one local minimum. In this case, we know $\alpha^2 - \beta^2 > 0$ and $\alpha^2 - 1 > 0$, so $\det \mathbf{N}$ is positive for $0 < \lambda < 1$, so must have its unique maximum in that interval, corresponding to the required inscribed ellipse of greatest area, which completes the proof of Lemma 2. (Note that $\det \mathbf{N}$ is also positive — and unbounded — when λ is on the opposite side of O_3 from O_1 and O_2, but that these values of λ correspond to *escribed* ellipses of $KLMN$. And, at its local minimum, $\det \mathbf{N}$ is negative, so this corresponds to a hyperbola.)

5. *Construction of the circumscribed ellipse*

We shall now describe a way, given a convex quadrilateral, of constructing the ellipse that featured in Proposition 1. The construction needs a dynamic geometry program such as Geometer's Sketchpad.

We need to be able to draw a conic through five given points. This facility is often provided by dynamic geometry packages, but if not it is easily done, using Pascal's theorem, which states that six points A, B, C, D,

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press

E, F lie on a conic (in some order) if, and only if, the meets $AB{\cdot}DE$, $BC{\cdot}EF$ and $CD{\cdot}FA$ lie on a line, the *Pascal line* of the hexagon $ABCDEF$. See [5, p.143]. So, given A, B, C, D, E, draw an arbitrary line ℓ through A and try to find the point F where this meets the required conic again: the Pascal line m is the join of $AB{\cdot}DE$ and $CD{\cdot}\ell$; if $BC{\cdot}m = X$, then $EX{\cdot}\ell = F$. The required conic is the locus of F as ℓ varies.

We also need to be able to draw a conic through four points, touching a given line at one of them. This is a limiting case of the above: let $AB = n$, and then, fixing n, let $A \to B$. So, given B, C, D, E, and a line n through B, draw an arbitrary line ℓ through B and try to find the point F where this meets the required conic again: the Pascal line m is the join of $n{\cdot}DE$ and $CD{\cdot}\ell$; if $BC{\cdot}m = X$, then $EX{\cdot}\ell = F$. The required conic is the locus of F as ℓ varies.

The construction that follows depends on some projective geometry, namely harmonic properties of conics, (complete) quadrilaterals and (complete) quadrangles. A complete quadrangle is a set of four points P, Q, R, S, no three collinear, joined in pairs by six lines, which come in pairs of opposites which meet in three diagonal points, $X = PQ{\cdot}RS$, $Y = PR{\cdot}QS$, and $Z = PS{\cdot}QR$, the vertices of its diagonal triangle. Dually, a complete quadrilateral is a set of four lines p, q, r, s, no three concurrent, meeting in pairs in six points, which come in pairs of opposites whose joins are three diagonal lines, $x = (p{\cdot}q)(r{\cdot}s)$, $y = (p{\cdot}r)(q{\cdot}s)$ and $z = (p{\cdot}s)(q{\cdot}r)$, the sides of its diagonal trilateral.

Suppose P, Q, R, S lie on p, q, r, s respectively. Then there is a conic touching p at P, q at Q, r at R and s at S if, and only if, $x = YZ$, $y = ZX$, and $z = XY$, in other words, if, and only if, the diagonal triangle of $PQRS$ and the diagonal trilateral of $pqrs$ are one and the same. This triangle (trilateral) is then self-polar with respect to the conic, that is, x, y, z are the polars of X, Y, Z respectively. (We shall not prove this here, but refer the reader to [5] for the background theory, especially Chapter V, §4.) In the notation of the theorem in section 3, we have that K, M lie on YZ; L, N lie on XY; and $KL{\cdot}MN$, $LM{\cdot}NK$ lie on ZX. Notice also that, if O is the centre of the conic, then KO passes through the mid-point T of PQ. See Figure 3.

So, given a convex quadrangle $PQRS$, we wish to construct Γ, the circumscribed ellipse of least area. There are three cases.

(i) Suppose $PQRS$ is a parallelogram. Then every proper conic circumscribing $PQRS$ has its centre at $O = PR{\cdot}QS$. Let the midpoint of RS be K', and mark K on $K'O$ so that $KO : OK' = 2 : 1$. Then the required ellipse is the conic through P, Q, R and S, touching PK at P. (If $PQRS$ is a square, we obtain its circumcircle.)

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press

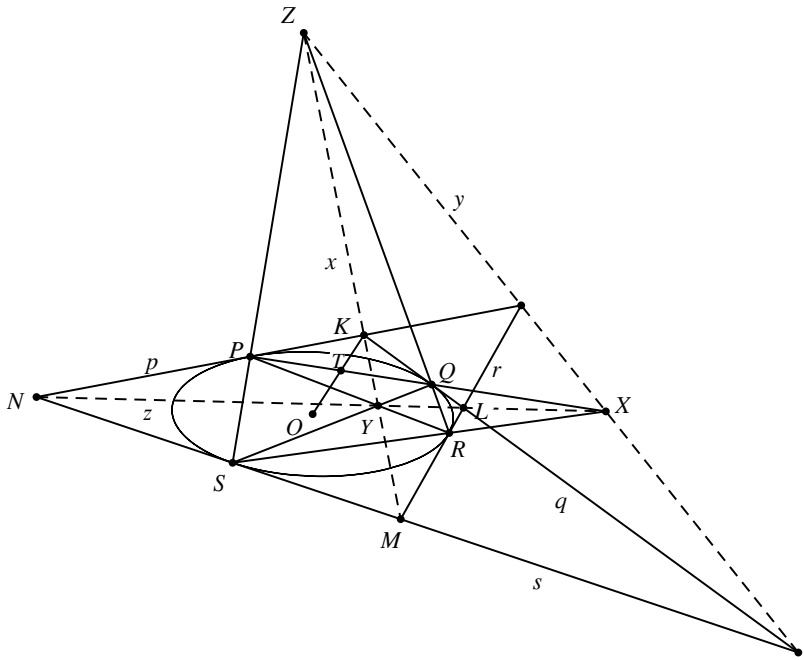


FIGURE 3

(ii) Suppose $PQRS$ is a trapezium, but not a parallelogram: without loss of generality, suppose that $QR \parallel SP$, and $QR < SP$. See Figure 4. Let $X = PQ{\cdot}RS$, let $Y = PR{\cdot}QS$, and let x be the line through Y parallel to QR. (Here the diagonal point Z is at infinity on QR.) Let T, U, V be the midpoints of PQ, QR, SP respectively, and let W be the midpoint of UV. (So U, V, W are on the diagonal line $z = XY$.) Then every proper conic circumscribing $PQRS$ has its centre on z, and is an ellipse if, and only if, its centre is between V and W. Put $\Lambda_0 = VW$, the segment; the construction will be continued in the paragraph after next.

(iii) If $PQRS$ has no parallel sides, construct the centre-locus Λ of the pencil P of conics through P, Q, R and S. This passes through the mid-points of PQ, PR, PS, QR, QS, RS and the meets $PQ{\cdot}RS$, $PR{\cdot}QS$ and $PS{\cdot}QR$, so pick any five of these nine points and construct the conic through them, which will be a hyperbola. See Figure 5. (If instead we had taken P, Q, R, S as the vertices of a triangle together with its orthocentre, then Λ would be the nine-point circle of the triangle.) This time, let Λ_1 be the branch of Λ that contains the diagonal points X, Y, Z of $PQRS$, and let Λ_0 be the *other* branch, so that any member of $\mathcal{P}$ with its centre on Λ_0 is an ellipse. Let T be the midpoint of PQ.

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press

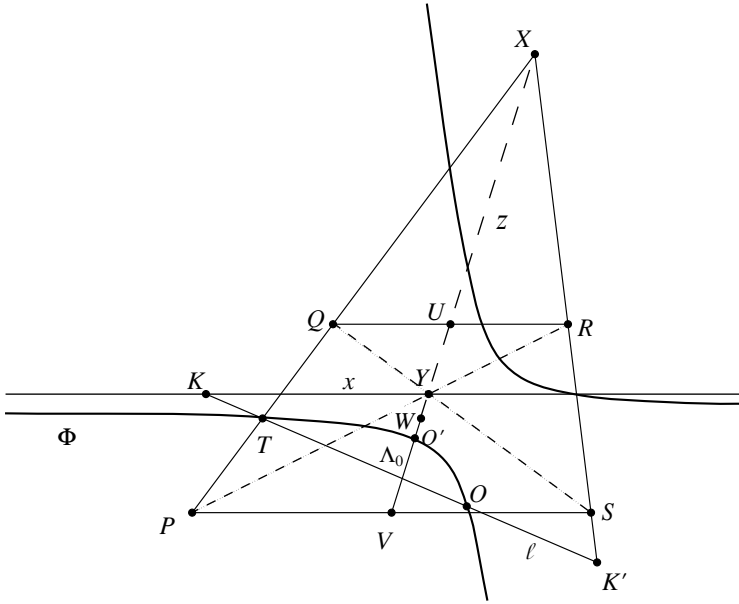


FIGURE 4

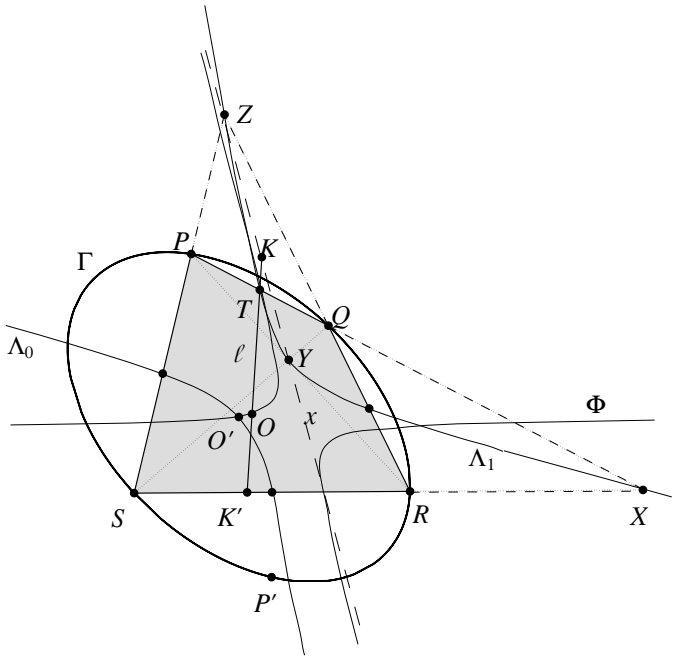


FIGURE 5

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press

To complete either (ii) or (iii), put an arbitrary line ℓ through T, and let it meet $x = YZ$ at K and RS at K'. Mark O on ℓ so that $KO : OK' = 2 : 1$. It would be very nice if O turned out to be on Λ_0, because this would then be the centre of the required ellipse Γ; but this is unlikely to happen. So instead, plot the locus Φ of O as ℓ varies; the correct position of the centre of Γ is the point O' where Φ meets Λ_0. Then mark P' on PO' so that O' is the midpoint of PP', and finally plot Γ as the conic through P, Q, R, S and P'.

The locus Φ is actually another conic — for a proof, use Steiner's theorem (see [5, p. 136]) — though we do not need to know that to perform the above constructions. Since the intersection of two conics is in general given by a quartic equation, the construction of Γ (or, rather, of the fifth point P' needed to determine Γ) by ruler and compasses is not possible, in general, in case (iii); it is possible in case (ii). Actually, Geometer's Sketchpad will not mark the intersection of two loci, though it will mark the intersection of a line or circle with a locus. So an alternative to the above in case (iii) is to take an arbitrary point O on Γ_0, let OT meet YZ at K, and mark K' on OT such that $KO : OK' = 2 : 1$. Then plot the locus of K' as O varies, and the correct position of K' is where this locus (which is not a conic in general) meets RS; the correct position of O now being where $K'T$ meets Λ_0.

As a limiting case, given ΔABC and a line ℓ through C, we can construct the ellipse of least area circumscribing ΔABC and with tangent ℓ at C. The method is arrived at by taking D on ℓ and applying the foregoing construction to $ABCD$, and then letting $D \to C$.

This time, the pencil $\mathcal{P}$ consists of conics circumscribing ΔABC and touching ℓ at C. Its centre-locus Λ is the conic through C, the midpoints of AB, BC and CA, and the meet $\ell \cdot AB$. Let Λ_0 be the branch of $\Lambda = \Lambda_0 \cup \Lambda_1$ that does not contain C or $\ell \cdot AB$. Find the centroid Γ of ΔABC and put in the line m through G, parallel to AB. Let Λ_0 meet m at O. Then if $CO \cdot AB = C'$, we have $CO : OC' = 2 : 1$, so O is the centre of the required ellipse Γ, and all we have to do is mark A', B' on AO, BO respectively so that O is the midpoint of AA' and of BB' , and Γ is then the conic through the five points A, B, C, A' and B'. See Figure 6.

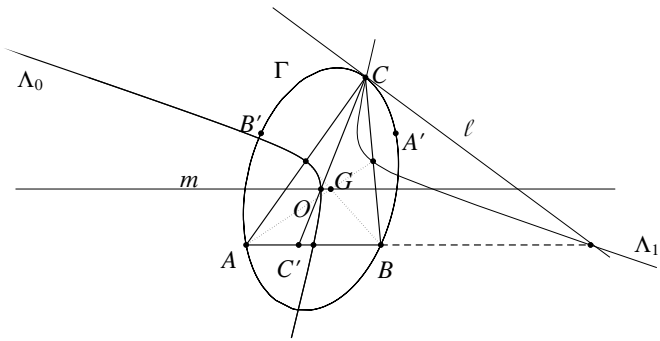


FIGURE 6

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press

The above only works if ℓ is not parallel to AB, but if we rotate ℓ so that it becomes parallel to AB, then Λ degenerates into a line-pair, the line joining the midpoints of BC and CA (every point of which is a centre for the parallel line-pair $(\ell)(AB)$, which belongs to $\mathcal{P}$), and the median through C, which contains the centres of all the other conics of $\mathcal{P}$. This meets m at G, so now $O = G$, and Γ is the Steiner circumellipse of ΔABC, as we should expect.

6. *Construction of the inscribed ellipse*

Now for the ellipse of Proposition 2. We need to be able to construct a conic touching five lines, which is done with the help of Brianchon's theorem, the dual of Pascal's. This says that six lines a, b, c, d, e, f touch a conic (in some order) if, and only if, $(a{\cdot}b)(d{\cdot}e)$, $(b{\cdot}c)(e{\cdot}f)$ and $(c{\cdot}d)(f{\cdot}a)$ (the joins of the meets) concur at a point, the *Brianchon point* of *abcdef*. See [5, p. 143] again. We need to be able to locate the points of contact of the conic touching five lines with each of them. We do this by taking the limiting case of Brianchon's theorem, as above, when $f \to a$. So the Brianchon point J (where the three joins meet) is the meet of $(a{\cdot}b)(d{\cdot}e)$ and $(b{\cdot}c)(e{\cdot}a)$, and now the join $(c{\cdot}d)J$ meets a at the required point of contact. So now, given five lines, which determine a unique conic touching them, we can construct the five points of contact, which enables us to draw the conic through these five points, which is the conic touching the five lines.

If $KLMN$ is a parallelogram, the inscribed ellipse of greatest area is the conic touching the four sides of the parallelogram at their respective midpoints. Since we know how to construct a conic through four points, touching a given line at one of them, this case is easy. If the parallelogram is a square, we get its inscribed circle.

Given a convex quadrilateral $KLMN$, *not* a parallelogram, the first thing to do is construct the centre-locus of the range $\mathcal{R}$ of conics touching the lines KL, LM, MN, NK, or the relevant part of the centre-locus for our purposes, which is the line segment joining O_1, the midpoint of KM, to O_2, the midpoint of LN.

Put $p = NK$, $q = KL$, $r = LM$ and $s = MN$. Let x, y, z be the diagonal lines of the quadrilateral $pqrs$; specifically, let $x = (p{\cdot}q)(r{\cdot}s) = KM$. $y = (p{\cdot}q)(r{\cdot}s)$ and $z = (p{\cdot}s)(q.r) = NL$, and then let $X = y{\cdot}z$, $Y = z{\cdot}x$ and $Z = x{\cdot}y$. Choose P on p, and let $Q = PX{\cdot}q$, $R = PY{\cdot}r$ and $S = PZ{\cdot}s$. Let T be the mid-point of PQ, put $K' = KT{\cdot}RS$, and mark O on KK' so that $KO : OK' = 2 : 1$. It would be very nice if O were on the segment O_1O_2, because then it would be the centre of Γ, the inscribed ellipse of greatest area we are looking for; but this is unlikely to happen. So instead plot the locus Φ of O as P moves on p, and the point O' where Φ crosses O_1O_2 is the correct position for the centre of Γ. See Figure 7. To finish, let p' be the 180° rotation of p about O', and construct Γ as the conic touching p, q, r, s and p'.

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press

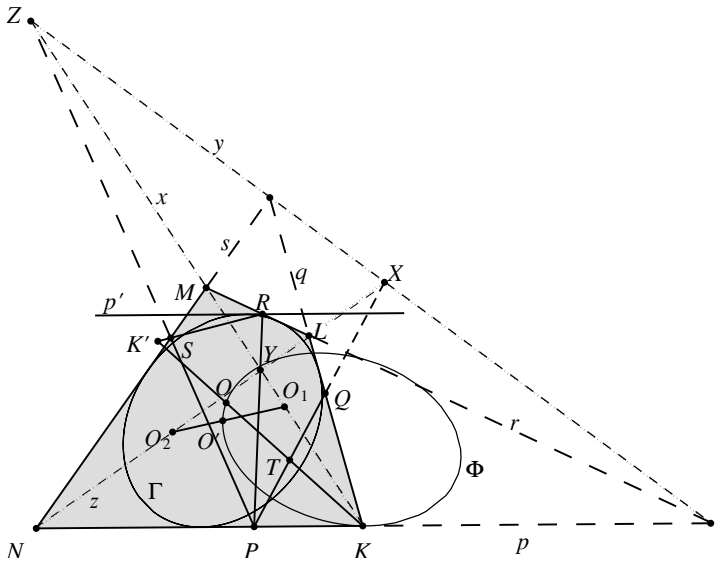


FIGURE 7

The locus Φ is yet another conic; this involves Steiner's theorem again. Given a line and five points, it is possible with ruler and compasses to locate the two points of intersection, where the line meets the conic through the five points (provided the points of intersection are real). That being so, it is possible in principle to construct O', as above, and an arbitrary number of points on Γ, with ruler and compasses.

As another limiting case, given $\triangle ABC$ and a point D on BC, we can construct the ellipse of greatest area inscribed in $\triangle ABC$ and touching BC at D. The method is arrived at by taking D_0 *not* on BC and applying the foregoing construction to ABD_0C (note the order), and then letting $D_0 \to D$.

This time, the range $\mathfrak{R}$ consists of conics touching the lines AB and CA, and touching BC at D. Its centre-locus is the line O_1O_2, where O_1 is the midpoint of AD and O_2 is the midpoint of BC. Let m be the line through the centroid G of $\triangle ABC$, parallel to BC. Then $O = m \cdot O_1O_2$ is the centre of the required ellipse Γ. Rotate any two sides of $\triangle ABC$ by 180° about O, and Γ can be constructed as the conic touching these two rotated lines and the three sides of the triangle. See Figure 8.

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press

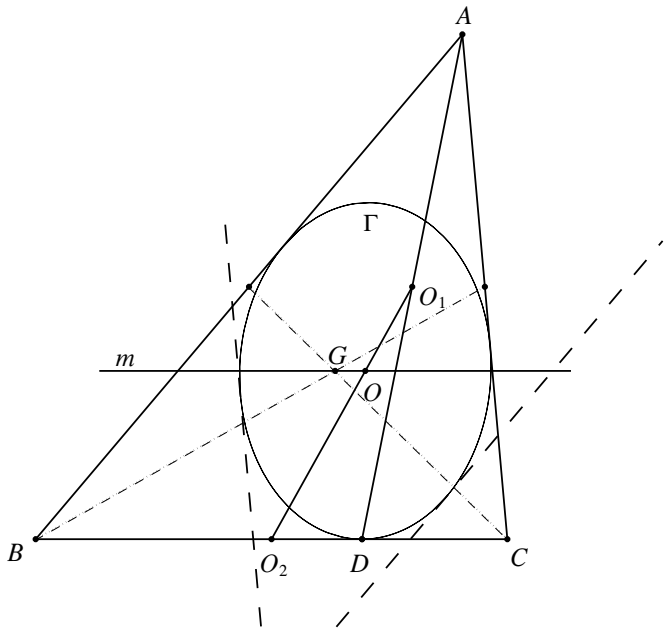


FIGURE 8

Note finally that if we take D as the midpoint of BC, then the lines AD and O_1O_2 coincide, $O = G$, and Γ is then the Steiner inellipse of $\triangle ABC$, again as we should expect.

References

1. Eric W. Weisstein, 'Steiner circumellipse' *Wolfram MathWorld*, accessed October 2016 at http://mathworld.wolfram.com/SteinerCircumellipse.html
2. Eric W. Weisstein, 'Steiner inellipse' *Wolfram MathWorld*, accessed October 2016 at http://mathworld.wolfram.com/SteinerInellipse.html
3. Alan Horwitz, Ellipses of minimal area and of minimal eccentricity circumscribed about a convex quadrilateral, *Aust. J. Math. Anal. App.*, **7** (2010) Article 8, also available at: http://arXiv.org/abs/0707.2092v1
4. G. H. Hardy, *Pure mathematics* (10th edn.), Cambridge University Press (1955), p. 497, Ex. 9.
5. J. G. Semple and G. T. Kneebone, *Algebraic projective geometry*, Oxford University Press (1952).

10.1017/mag.2017.2

JOHN R. SILVESTER
Department of Mathematics, King's College, Strand, London WC2R 2LS
e-mail: jrs@kcl.ac.uk

https://doi.org/10.1017/mag.2017.2 Published online by Cambridge University Press