CROSS RATIO, HARMONIC SEQUENCE, AND LARGEST AREA ELLIPSES INSCRIBING SPECIFIC QUADRANGLES

M. John D. Hayes[1], Anton Gfrerrer[2]

[1]*Department of Mechanical and Aerospace Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada*
[2]*Institute for Geometry, Graz University of Technology, Graz, Austria*
Email: John.Hayes@carleton.ca; gfrerrer@tugraz.at

ABSTRACT

The projective geometric properties of cross ratio and harmonic sequence are used to explain how an elegant projective geometric computational method works and why. Namely, to answer the question "why does the projective transformation that sends the vertices of a square onto the vertices of an arbitrary convex parallelogram or trapezoid, also map the maximum area ellipse (circle) inscribing the square onto the maximum area ellipse inscribing either the parallelogram or trapezoid, but not for an arbitrary convex quadrangle in general"? In this paper we discuss the nature of the relationships among the points comprising a complete quadrangle. These relationships are examined in the context of the cross ratio of four points on a line in general and the special case of a harmonic sequence in particular. From these relationships, the answers to the posed question, as well as several others, are revealed.

Keywords: cross ratio; harmonic sequence; projective and affine transformations.

RAPPORT ANHARMONIQUE, DIVISION HARMONIQUE ET DES ELLIPSES DE PLUS GRANDE SURFACE INSCRIVIANT DES QUADRANGLES SPÉCIFIQUES

RÉSUMÉ

Les propriétés géométriques projectives du rapport anharmonique et de la division harmonique expliquent le fonctionnement d'une méthode de calcul géométrique projective élégante. Qui peut résoudre la question suivant : "pourquoi la transformation projective qui transpose les sommets d'un carré a ceux d'un parallélogramme convexe arbitraire, ou trapézoïdale, trace également l'ellipse de surface maximale (cercle) insérant le carré sur l'ellipse de surface maximale inscrivant soit le parallélogramme, soit le trapéze, mais pas pour un quadrangle convexe arbitraire en général" ? Dans cet article, nous discutons de la nature des relations entre les points constituant un quadrangle complet. Ces relations sont examinées dans le contexte du rapport anharmonique de quatre points d'une ligne en général et du cas particulier d'une division harmonique en particulier.

Mots-clés : rapport anharmonique ; division harmonique ; les transformations projectives et affines.

1. INTRODUCTION

This paper presents the answers to questions raised by results that appeared to be a geometric anomaly presented in [1], where an algorithm is presented that determines the largest area ellipse inscribing either convex parallelograms or trapezoids. While the algorithm identifies an ellipse that inscribes an arbitrary convex quadrangle with no parallel edges, the identified ellipse does not, in general, possess the largest area. These results underscore that projective and affine transformations which map collinear points onto collinear points do not, in general, preserve area ratios. They are nonetheless preserved by these transformations in two very specific circumstances. The present reader may wish to ask two questions: why is this problem of any interest at all; and what geometric reasons are there to account for these curious results?

The answer to the first question can be found in the literature. For example error and covariance ellipses subject to linear constraints are important in statistical analysis. Consider systems of design or measurement variables in an electrical or mechanical system. An error ellipse is a way of visualising the confidence interval of normally distributed data [2]. While covariance is a measure of how changes within one variable are related to changes in a second; the covariance between two variables, therefore, becomes a measure of to what degree each variable is dependent upon the other. In statistical analysis the covariance ellipse of n separate variables, given distinct data points, can be generated as an $n \times n$ matrix [2]. The diagonal of the matrix represents the variance of each variable within the data set, while each non-diagonal element represents the covariance of each variable with another. The indices of the matrix element indicates which two variables are involved. For a two variable system the matrix is 2×2 and symmetric, possessing a form identical to that of the quadratic form of an ellipse. The largest area ellipse indicates the maximum covariance between the variables.

Performance indices for machine design are used to compare specific elements of capability. Redundantly actuated parallel mechanisms have operational force outputs that are not unique; these forces do not correspond to a unique set of joint forces, which can help reduce the effect of singularities [3–5]. Analysis of kinematic isotropy or the capacity of a mechanism to change position orientation, and velocity given its pose in the workspace yields insight regarding velocity performance [6]. In this context, the area of the ellipse inscribing the arbitrary polygon defined by the reachable workspace of the redundantly actuated parallel mechanism is proportional to the kinematic isotropy of the mechanism. In [3, 5] the approach to identifying the maximum area inscribing ellipse is a numerical problem, essentially fitting the ellipse inscribing the linear constraints defining the velocity profile of the mechanism by starting with the unit circle.

The energy dissipated in an oscillating mechanical system when plotted as a damping force-displacement curve always encloses an area called the hysteresis loop [7]. The area of the loop is the energy loss or work done by the damping force per cycle. For purely mechanical viscous shock absorbers, the loop is an ellipse. If the damping and displacement are linearly constrained then the maximum area inscribing ellipse is of interest. Moreover, the inertia ellipsoid of a rigid body represents the relationship between its moment of inertia about an axis and its instantaneous kinetic energy [7, 8]. The principal axes of the ellipsoid are those of the body's inertia tensor. An ellipsoid is generated by rotating an ellipse about its major axis. The potential for constrained optimisation applications in this area of study is intriguing!

To the best of the author's knowledge, there are only a handful of papers that report investigations into determining maximum area ellipses inscribing arbitrary polygons. The dual problem of determining the polygons of greatest area inscribed in an ellipse is reported in [9]. While interesting, this dual problem is not germane to determining the maximum area ellipse inscribing a polygon. Three papers by the same author [10–12] appear to lead to a solution to the general problem of finding the largest area ellipse inscribing an n-sided convex polygon, however the papers focus on the proof of the existence of a solution rather than an explicit algorithm for computing the ellipse equation or shape coefficients.

In this paper we revisit the projective extension of the Euclidean plane approach to identifying the coordinate transformation that maps the vertices of a unit square to those of a convex quadrangle, then using the transformation to map the parametric equation of the unit circle inscribing the square to one of an ellipse inscribing the quadrangle, originally presented in [13]. The curious result that the approach leads to the maximum area ellipses that inscribe convex parallelograms and trapezoids observed in [1] is explained in terms of projective geometry, namely the cross ratio and harmonic sequence of four points on a line, along with the dual entities of polar lines and pole points. Note that in this work the ellipse must touch all edges of the quadrangle, there can be other interior ellipses with larger area that don't touch all edges [14].

2. MATHEMATICAL BACKGROUND

The geometry of planar four sided polygons and curves, along with the projective geometric invariants of cross ratio and harmonic sequence, have all been the subject of intense study since antiquity [15]. In this section we briefly discuss the relevant properties in subsections describing: the *complete quadrangle*; the *cross ratio* and a special germane case called *harmonic sequence*; *projective* and *affine transformations*; and finally relevant properties of *parallelograms* and *trapezoids*.

2.1. Complete Quadrangle

A *complete convex quadrangle* is a configuration consisting of four coplanar points, no three of which are collinear, defining the four quadrangle vertices, together with its six sides determined by pairs of these vertices. Sides not on the same vertex are called *opposite*. The edges are the four line segments joining sequential vertices. A convex quadrangle is one whose edges do not intersect, except at its vertices. In other words, the edges of a convex quadrangle do not self-intersect. The triangle determined by the points of intersection of opposite sides is called the *diagonal triangle*, whose vertices are the *diagonal points* of the quadrangle. For the complete quadrangle illustrated in Fig. 1, the vertices are points A, B, C, and D, while the diagonal points are O, E, and F. Point Q is the point of intersection of the diagonal line f containing vertices B and D and the diagonal line g containing the two diagonal points E and F.

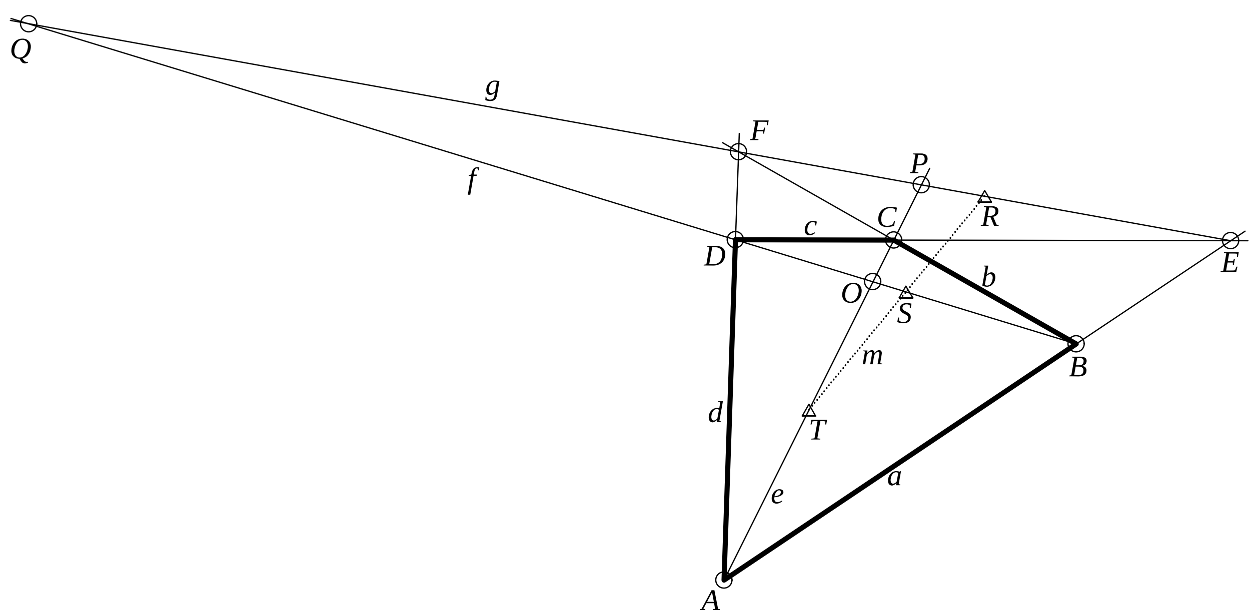

Fig. 1. Complete convex quadrangle.

A projective correlation in a plane in which the points and lines correspond doubly is called a *polarity*

[16]. An important relevant polarity concerning conics are the dual projective elements of polar lines and corresponding pole points. The polar line to a point on a nondegenerate point conic is the tangent line to the conic at that point, while the pole point to a polar line of a line conic is the tangent point on that line . In the context of conics that inscribe a quadrangle, the edges are polar lines with respect to the conic, lines a, b, c, and d in Fig. 1 and dually the corresponding tangent points are the pole points.

It is well known that all second order planar curves, i.e. conics, which contain the four vertices A, B, C, and D of a given convex quadrangle establish a one parameter family of curves, a so called *pencil of point conics*: $\mathscr{P}_{ABCD}$ [17]. Its dual counterpart is a *pencil of line conics* $\mathscr{P}_{abcd}$, which is the set of all line conics that contain the four edges of a given quadrilateral. The diagonal trilateral efg, see Fig. 1, is a common polar trilateral of all curves in $\mathscr{P}_{abcd}$, which means that e, f, and g are the polar lines of $Q := f \cap g$, $P := e \cap g$, and $O := e \cap f$, respectively, with respect to any curve $k \in \mathscr{P}_{abcd}$. The corresponding *pole triangle* has vertices O, P, and Q. The midpoints of the three diagonals of the complete quadrangle are all always collinear. These are the midpoints of the distances between points A and C on line e, points B and D on line f and points E and F on line g. We label these midpoints T, S, and R, and the line on which they lie, m.

Since we are imposing the condition that the quadrangle $ABCD$ is convex, point O always lies on the interior, whereas P, Q, E, and F always lie on the exterior. The points P, Q, E, and F are always collinear, laying on line g. Again, because of the convexity condition, point O is always a finite (proper) point, whereas points P, Q, E, and F can all be either finite or infinite (improper). The relative locations of these four points on line g yields a classification scheme for quadrangles. A convex quadrangle can either possess no parallel edges, a pair of opposite parallel edges or two pairs of opposite parallel edges. The presence of a pair, or pairs, of opposite parallel edges will place some or all of these four points on the line at infinity. Because these four distinct points lie on the common line g, there are only four cases that can occur [18].

1. Parallelogram case: all four points P, Q, E, and F are points at infinity and hence line g is the line at infinity $\mathscr{L}_\infty$.
2. Trapezoid case: points P and Q are finite, while one of points E or F is finite and the other is at infinity.
3. Kite case: points E and F are finite, while one of points P or Q is finite and the other is at infinity.
4. General case: all four points P, Q, E, and F are finite points.

Since we are interested in answering the questions raised in [1], our focus in this paper will be on the first two cases: parallelograms and trapezoids.

2.2. Cross Ratio

The concept of *cross ratio* is derived from the relative positions of four collinear points and is widely believed to have been known since antiquity [19]. The cross ratio of four points on a line in a specified sequence is the only invariant of projective geometry and is the fundamental invariant in every linear geometry [20], including Euclidean. It can be analytically defined as [19]: given four distinct collinear points in the sequence A, B, C, and D on a projective line having homogeneous coordinates $(a_0 : a_1)$, $(b_0 : b_1)$, $(c_0 : c_1)$ and $(d_0 : d_1)$, respectively, then the real number

$$CR(A,B;C,D) = \frac{\begin{vmatrix} a_0 & a_1 \\ c_0 & c_1 \end{vmatrix} \begin{vmatrix} b_0 & b_1 \\ d_0 & d_1 \end{vmatrix}}{\begin{vmatrix} b_0 & b_1 \\ c_0 & c_1 \end{vmatrix} \begin{vmatrix} a_0 & a_1 \\ d_0 & d_1 \end{vmatrix}} \tag{1}$$

is the *cross ratio* of the four points in the order A, B, C, D. Evaluating the first determinant in Eq. (1) yields $a_0c_1 - c_0a_1$ which can be interpreted in a metric way as the directed distance from A to C. With the use of this metric concept the cross ratio of the four collinear points can also be expressed as the ratios of directed distances along the line [21]:

$$CR(A,B;C,D) \;=\; \left(\frac{AC}{BC}\right)\left(\frac{BD}{AD}\right). \tag{2}$$

The values of $CR = 0$, $CR = \infty$, and $CR = 1$ cannot occur for four distinct and finitely separated points. Consider the four distinct and collinear points in the order A, B, C, D, and the Euclidean directed distance interpretation of cross ratio. The value of the cross ratio is given by the ratio of the directed distances as the product of ratios in Eq. (2). Considering that equation, the value of $CR = 0$ requires that points A and C coincide and/or points B and D coincide. If $CR = \infty$ then points A and D coincide and/or points B and C coincide. Finally, if $CR = 1$ then points A and B coincide and/or points C and D coincide. Hence, these three values are not in the set of possible values for the cross ratio of four distinct collinear points.

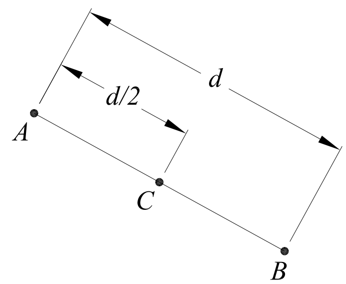

Fig. 2. Planar 4R linkage.

If one of the points along the line is at infinity, then the ratio containing the homogenising coordinate that is 0 is simply not included in the computation [19, 20, 22]. Moreover, when C is midway between A and B while D is at infinity then $CR = -1$ and the four points are said to be a *harmonic quadruple* or in a *harmonic sequence* [23]. Consider this situation illustrated in Fig. 2. Let the coordinate origin be at point A and AB the direction of increasing coordinates. Because point D is at infinity, we ignore the ratio of the directed distances BD and AD, leaving only the ratio

$$CR(A,B;C,D) \;=\; \frac{AC}{BC} \;=\; \frac{d/2-0}{d/2-d} \;=\; \frac{d/2}{-d/2} \;=\; -1. \tag{3}$$

Regardless, any four finite points on a line whose cross ratio is $CR = -1$ are in a harmonic sequence. Complete quadrangle points E, F, P, and Q as in Fig. 1, regardless of finiteness, always lie in a harmonic sequence in the order E, F, P, Q [23]. Since this is the value of the cross ratio of these particular four collinear points, this value is preserved under any linear transformation whatsoever. This is a harmonic property of all complete convex quadrangles and is stated more formally here in Proposition 1 [22].

Proposition 1 *Two opposite vertices of a complete quadrangle are separated harmonically by the points in which their diagonal is met by the other two diagonals.*

Proof of Proposition 1: the proof of Proposition 1 requires Lemma 1.

Lemma 1 *If α, β, γ, and δ are distinct collinear points or concurrent coplanar lines, then*

$$CR(\alpha,\beta;\gamma,\delta) \;=\; CR(\gamma,\delta;\alpha,\beta) \;=\; \frac{1}{CR(\alpha,\beta;\delta,\gamma)}.$$

Proof of Lemma 1: to prove Lemma 1 we use two theorems from projective geometry [19, 22, 24].

Theorem 1 *The value of the cross ratio of four distinct collinear points or concurrent coplanar lines, remains unchanged when any two of the four elements are interchanged simultaneously with the other two.*

Theorem 1 states that the value of a cross ratio is unchanged by reversing the order of elements in both the first and second pairs or of the inner and outer elements or of the first and second pairs. This means

$$CR(\alpha,\beta;\gamma,\delta) \;=\; CR(\beta,\alpha;\delta,\gamma) \;=\; CR(\delta,\gamma;\beta,\alpha) \;=\; CR(\gamma,\delta;\alpha,\beta). \tag{4}$$

Theorem 2 *The value of the cross ratio of four distinct collinear points or concurrent coplanar lines, is changed to its reciprocal when the order of the elements in either the first or second pair is reversed.*

Theorem 2 means that

$$CR(\beta,\alpha;\gamma,\delta) \;=\; \frac{1}{CR(\alpha,\beta;\gamma,\delta)} \quad \text{or,} \quad CR(\alpha,\beta;\delta,\gamma) \;=\; \frac{1}{CR(\alpha,\beta;\gamma,\delta)}. \tag{5}$$

Clearly, Lemma 1 follows directly from Theorems 1 and 2.

Now, to prove Proposition 1 we consider the complete quadrangle illustrated in Fig. 1 with edges AB, AD, BC, and CD. We consider the diagonal points E and F to be *opposite vertices* as well as the points P and Q in which their diagonal EF is met by diagonals AC and BD and show that $CR(E,F;P,Q) \;=\; -1$. Points E and F can be considered opposite vertices since they are the points of intersection of opposite edges AB, CD and AD, BC. Choose vertex A as the centre of projection and centrally project line g onto line f. Points E, F, P, and Q then project centrally through vertex A onto points B, D, O, and Q, respectively. This projection is represented in projective geometric symbols as

$$EFPQ \;\overset{A}{\underset{\wedge}{=}}\; BDOQ. \tag{6}$$

Central projection is a collineation that preserves the cross ratio [22], hence

$$CR(E,F;P,Q) \;=\; CR(B,D;O,Q). \tag{7}$$

Now, with vertex C as the projection centre, project line f onto line g so that

$$BDOQ \;\overset{C}{\underset{\wedge}{=}}\; FEPQ. \tag{8}$$

The projection through C yields

$$CR(B,D;O,Q) \;=\; CR(F,E;P,Q). \tag{9}$$

From Equations (7) and (9) we have

$$CR(E,F;P,Q) \;=\; CR(F,E;P,Q). \tag{10}$$

Now, relying on Lemma 1 we can write

$$CR(F,E;P,Q) \;=\; \frac{1}{CR(E,F;P,Q)}. \tag{11}$$

Finally, substituting Eq. (11) into Eq. (10) we obtain

$$(CR(E,F;P,Q))^2 \;=\; 1. \tag{12}$$

Since the cross ratio of four distinct points on a line in a specific order can be thought of as representing the ratio of division of directed distances then the cross ratio $CR(A,B;C,D) < 0$ if and only if the point pairs A, B separate C, D. Otherwise the cross ratio will be a positive number. Points E and F always separate points P and Q in the way we have defined a general complete quadrangle. Moreover, the cross ratio of four distinct points must exclude the values 0, 1, and ∞. In the case of Eq. (12), we therefore must infer that $CR(E,F;P,Q) = -1$. This completes the proof of Proposition 1.

The value of the cross ratio depends upon the order of the four points. There are 24 distinct permutations of four distinct points and there is the same number of cross ratios. However, there are only six distinct cross ratio values for four distinct collinear points, occurring in sets of four, as required by Theorem 1.

2.3. Projective and Affine Transformations

The projective transformation group in the projective plane P_2 may be thought of as 3×3 matrix operators that are collineations. It is important to note that an $n+1$ dimensional homogeneous coordinate space is required to analytically describe the elements of an n dimensional projective space. These matrices are non-singular by definition. They are sometimes referred to as *structure matrices* [25] since changing the structure of the matrix changes the character of the geometry it represents. A coordinate transformation of P_2 may be represented as

$$\mathbf{T} = \begin{bmatrix} t_{00} & t_{01} & t_{02} \\ t_{10} & t_{11} & t_{12} \\ t_{20} & t_{21} & t_{22} \end{bmatrix}, \tag{13}$$

where the nine matrix elements t_{ij} are arbitrary real numbers meeting the only condition $\det(\mathbf{T}) \neq 0$. As homogeneous coordinates are used the matrix $\mathbf{T}$ corresponding to a particular projective transformation is only determined up to a scalar factor $\rho \neq 0$. Hence, the projective group of all collineations in P_2 has eight independent parameters and the cross ratio of four collinear points is the only invariant.

The affine transformation group in the affine plane A_2 is a subset of the projective group. Because of the use of homogeneous coordinates, only six of the elements are arbitrary, with two exceptions: $t_{01} = t_{02} = 0$; and the determinant $t_{11}t_{22} - t_{21}t_{12} \neq 0$. The affine group may be considered as being richer than the projective group because the geometry defined on the affine plane is due to the existence of more invariants belonging to the transformation group. While the distance between two points is not invariant under an affine transformation, ratios of distances are. In particular, affine transformations send equal distances into equal distances and preserve ratio of division, thus sending midpoints into midpoints [22]. Affine transformations preserve the *between*-relation and hence the property of being a segment, directed distance, angle, n-sided polygon or pair of equal vectors. These transformations multiply the areas of all triangles by a constant value, which is equal to the determinant of the associated transformation. Under any affine transformation the image of a conic section is a conic section of the same type, but not necessarily of the same eccentricity and is degenerate only if the original conic is. Finally, affine transformations preserve the line at infinity, and of course, the cross ratio [22].

2.4. Relevant Properties of Parallelograms and Trapezoids

It is well known that the closed second order curve possessing the largest area that inscribes a square is a circle. The pole points between the circle and the square it inscribes are the midpoints of the four edges, while the centres of both circle and square are coincident. Five distinct pole points in general, no three collinear, are needed to uniquely determine a point conic. Dually, five polar lines, no three coincident, are required to uniquely define a line conic. This means that a convex quadrangle, which provides only four of the five required constraints, must possess a one parameter family of inscribing ellipses. One and only one possesses the largest area [11].

Given an arbitrary convex quadrangle, if the interior diagonals intersect at their midpoints then the quadrangle is a parallelogram and all inscribing ellipses are centred at the diagonal's point of intersection [24], as in Fig. 3. Alternately, with reference to Fig. 1, if the interior diagonals do not intersect at their midpoints the one parameter family of inscribing ellipses are all centred at points along the line segment joining the diagonal midpoints T and S on line m [24]. The degenerate bounding ellipses in the one parameter pencil inscribing any convex quadrangle, are the two diagonals whose centres are at the diagonal midpoints T and S, respectively as in Fig. 3.

Consider the parallelogram illustrated Fig. 3. The affine transformation that maps a square to a parallelogram preserves the property that the largest inscribing ellipse, among the pencil of inscribing ellipses

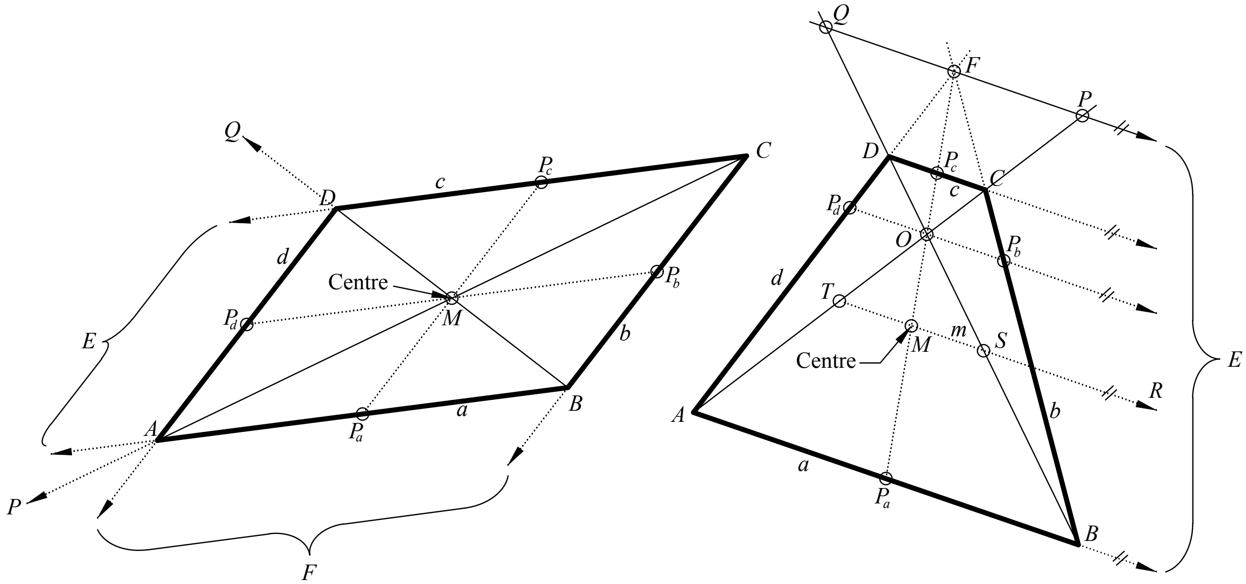

Fig. 3. Pole points and centre of largest area ellipse inscribing a parallelogram and a trapezoid.

centred at the intersection of the diagonals AC and BD, has pole points P_a, P_b, P_c, and P_d located at the mid points of the edges [22]. The diagonal line g is the line at infinity, $\mathscr{L}_\infty$, containing points P, Q, E, and F.

Two opposite edges in a trapezoid are always parallel, see Fig. 3. The lines containing the two non parallel edges always intersect in a finite point, which we call F. As one can easily check, the diagonal's intersection point O is distinct from both their midpoints T and S. The centres of the ellipses inscribing the trapezoid lie on the open line segment with endpoints T and S. The inscribing ellipse with the greatest area has its centre M at the midpoint of T and S [24]. The coordinates of this point are the geometric centre, or barycentre, of the trapezoid determined by the relation

$$\begin{bmatrix} M_0 \\ M_1 \\ M_2 \end{bmatrix} = \frac{1}{4}\left(\begin{bmatrix} A_0 \\ A_1 \\ A_2 \end{bmatrix} + \begin{bmatrix} B_0 \\ B_1 \\ B_2 \end{bmatrix} + \begin{bmatrix} C_0 \\ C_1 \\ C_2 \end{bmatrix} + \begin{bmatrix} D_0 \\ D_1 \\ D_2 \end{bmatrix}\right).$$

The following statements apply to the trapezoid in Fig. 3, but apply equally to any trapezoid with appropriate relabelling. The pole points of the area maximising ellipse on the parallel edges AB and CD are the midpoints P_a and P_c of these edges. The pole points P_b and P_d on the non parallel edges AD and BC are on the line through O parallel to AB and CD. Moreover, the five points P_a, M, O, P_c, and F are always collinear [24].

3. ANSWERS TO THE QUESTIONS

Two distinct sets of four distinct points in the projective plane P_2 uniquely determine a projective transformation if no three of the points are on the same line [22]. Let an arbitrary pair of distinct finite points (x, X) have the coordinates $x(x_0 : x_1 : x_2)$ and $X(X_0 : X_1 : X_2)$. The projective transformation can be represented by the vector-algebraic relationship

$$\lambda \begin{bmatrix} X_0 \\ X_1 \\ X_2 \end{bmatrix} = \mu \begin{bmatrix} t_{00} & t_{01} & t_{02} \\ t_{10} & t_{11} & t_{12} \\ t_{20} & t_{21} & t_{22} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix}, \quad (14)$$

where λ and μ are arbitrary scalars. Without loss in generality, we can set $\rho = \lambda/\mu$ and express Eq. (14) more compactly as

$$\rho \mathbf{X} = \mathbf{T}\mathbf{x}, \quad \text{or} \quad \mathbf{T}\mathbf{x} - \rho \mathbf{X} = \mathbf{0}. \tag{15}$$

If we wish to determine the transformation $\mathbf{T}$ given the coordinates of four points and their images, we must determine 12 unknowns: the eight independent elements of the transformation matrix and the four independent scaling factors, ρ_i, $i \in \{1,2,3,4\}$.

We consider the four points a, b, c, and d, to be the vertices of the square containing the unit circle centred on the origin of the Cartesian coordinate system in which an arbitrary quadrangle is defined. The image of the vertices of the square are those of the quadrangle represented by the four points A, B, C, and D. Now a set of equations must be written so that the elements of $\mathbf{T}$ can be computed in terms of the point and image coordinates:

$$\begin{aligned}
t_{00}a_0 + t_{01}a_1 + t_{02}a_2 - \rho_1 A_0 &= 0, \\
t_{10}a_0 + t_{11}a_1 + t_{12}a_2 - \rho_1 A_1 &= 0, \\
t_{20}a_0 + t_{21}a_1 + t_{22}a_2 - \rho_1 A_2 &= 0, \\
t_{00}b_0 + t_{01}b_1 + t_{02}b_2 - \rho_2 B_0 &= 0, \\
&\vdots \\
t_{20}d_0 + t_{21}d_1 + t_{22}d_2 - \rho_4 D_2 &= 0.
\end{aligned} \tag{16}$$

Equations (16) represent 12 linear equations in 13 unknowns, 12 of which are independent, hence we can arbitrarily scale the elements of $\mathbf{T}$ by $1/t_{00}$, thereby setting $t_{00} = 1$. Two examples are considered now, where the ellipse bounding quadrangles are a parallelogram and a trapezoid, respectively. Once the transformation has been identified the parametric equation of the unit circle is transformed to the parametric equation of the maximum area ellipse inscribing either the parallelogram or trapezoid used to identify the transformation. The resulting parametric ellipse equation is easily re-expressed as a homogeneous second order implicit equation. But it is important to emphasise that $\mathbf{T}$ is not the matrix of point conic shape parameters, rather it is the transformation that maps the parallelogram inscribing ellipse parametric equation to that of the square inscribing circle. The parametric equation of the desired ellipse is obtained by multiplying the parametric equation of the unit circle centred at the origin by $\mathbf{T}^{-1}$.

3.1. Parallelogram

Consider the square and parallelogram in Fig. 4. The homogeneous coordinates of the square vertices are $a(1:-1:-1)$, $b(1:1:-1)$, $c(1:1:1)$, and $d(1:-1:1)$ and the image points are those of the boundary quadrangle, which are $A(1:2:0)$, $B(1:10:-1)$, $C(1:14:3)$, and $D(1:6:4)$. The projective collineation defined by the vertices of the two quadrangles is computed to be

$$\mathbf{T}^{-1} = \frac{1}{2}\begin{bmatrix} 2 & 0 & 0 \\ 16 & 8 & 4 \\ 3 & -1 & 4 \end{bmatrix}. \tag{17}$$

The transformation in Eq. (17) is clearly affine, so the properties of betweenness and of being the midpoint between two others are preserved by the transformation [22]. Indeed, the corresponding maximum area ellipse inscribing the parallelogram has the homogeneous implicit equation

$$-1075.7100x_0^2 - 24.3179x_1^2 - 114.4372x_2^2 + 354.7554x_0x_1 + 160.2121x_0x_2 + 22.8874x_1x_2 = 0. \tag{18}$$

The four pole points are computed to be located at the mid points of the edges, with centre at the intersection of the two diagonals, by virtue of these two facts the identified inscribing ellipse is the one possessing the maximum area. This transformation method will always result in the largest area ellipse inscribing a parallelogram because it maps pole points located at the midpoint of each edge and will always result in an affine transformation matrix.

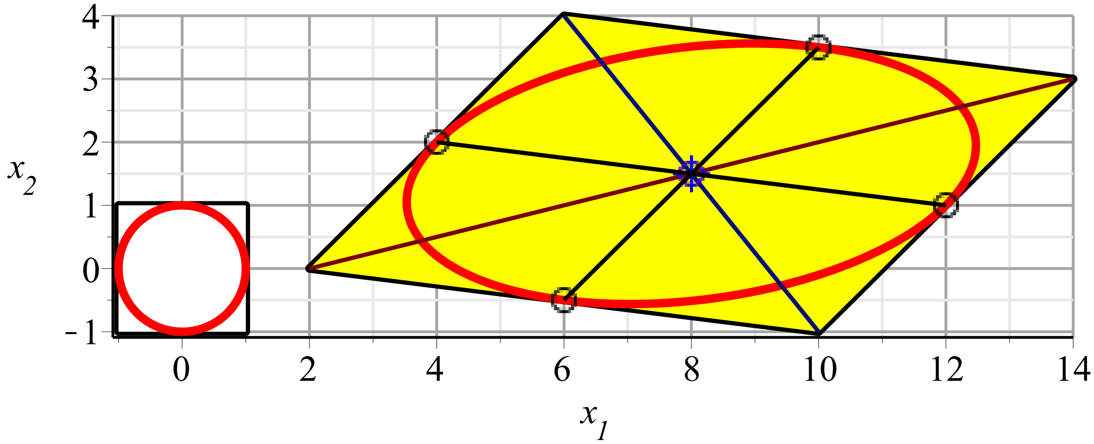

Fig. 4. Maximum area ellipse inscribing a parallelogram.

3.2. Trapezoid

Consider the square and trapezoid illustrated in Fig. 5. The square has the same coordinates as before, while the image of these points are the vertices of the trapezoid, $A(1:2:-1)$, $B(1:6:1)$, $C(1:5:4)$, and $D(1:3:3)$. The projective collineation defined by the vertices of the two quadrangles is

$$\mathbf{T}^{-1} = \frac{1}{3}\begin{bmatrix} 3 & 0 & 1 \\ 12 & 4 & 4 \\ 7 & 2 & 7 \end{bmatrix}. \tag{19}$$

The corresponding maximum area ellipse inscribing the trapezoid has the homogeneous implicit equation

$$-2.7943x_0^2 - 0.2315x_1^2 - 0.1210x_2^2 + 1.6246x_0x_1 - 0.0650x_0x_2 + 0.1210x_1x_2. = 0. \tag{20}$$

The Cartesian coordinates of the ellipse centre are computed to be $M\left(\frac{1}{4}(16,7)\right)$, which are the same as those of the midpoint of the open line segment joining the midpoints of the diagonals, the barycentre of the trapezoid. Moreover, the Cartesian coordinates of the pole points of the ellipse are computed to be $P_a(4,0)$, $P_b\left(\frac{1}{3}(16,9)\right)$, $P_c\left(\frac{1}{2}(8,7)\right)$, and $P_d\left(\frac{1}{3}(8,5)\right)$, the first and third corresponding to the midpoints of the two parallel trapezoid edges, while the second and fourth pole points are on a line through point O that is parallel to edges AB and CD, see Fig. 5. By virtue of these facts, the computed ellipse is the one in the pencil of inscribing ellipses possessing the greatest area. But, why did this work?

Examining the transformation matrix $\mathbf{T}^{-1}$ for the trapezoid it is to be seen that it is a projective transformation, so its only invariant is the cross ratio. We already know that if A, B, P_a, and E are points on a line such that P_a is the midpoint of the segment AB and E is the point at infinity of this line, then $CR(A,B;P_a,E) = -1$. Moreover, the cross ratio of four points on a line is invariant under a central projection through some point O to corresponding points on another line. If we project points of the line AB in Fig. 5 onto those of the line AD using the projection centre O, the points $A,\ P_a,\ B,\ E \in AB$ are mapped to the points $A,\ F,\ D,\ P_d \in AD$.

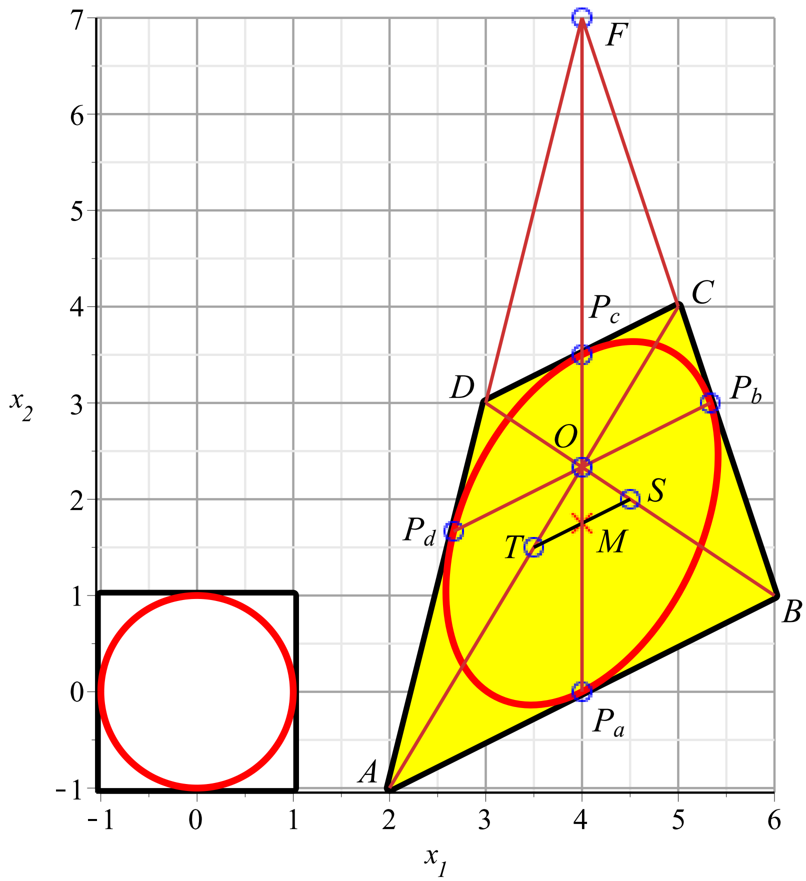

Fig. 5. Maximum area ellipse inscribing a trapezoid.

Since $CR(A,B;P_a,E) = -1$, we then also have $CR(A,D;P_d,F) = -1$. This proves that the projective transformation that maps the vertices of a square onto those of an arbitrary convex trapezoid will also always map the circle inscribing the square to the largest area ellipse inscribing the trapezoid.

4. CONCLUSIONS

In this paper the concepts of cross ratio and harmonic sequence were used to prove that the transformation which maps the vertices of a square to those of either a convex parallelogram or trapezoid also maps the circle inscribing the square to the maximum area ellipse inscribing the quadrangle, touching all four edges. While these results may seem obvious to a geometer, they are far from obvious in the mechanism and machine theory community.

Hence, the main contribution of this paper has been uncovering the affine and projective geometric insight required to provide the answers to the questions posed by earlier work on quadrangle inscribing ellipse area maximisation. The important questions raised in 2003 and 2016 in [1, 13] have finally been answered and laid to rest. The projective transformation method can only be applied to parallelograms and trapezoids since the pole points in a convex quadrangle with no parallel edges are not at the midpoints of the edges in general. These results provide a new mathematical tool for velocity performance analysis of redundantly actuated parallel mechanisms and covariance analysis, among other applications involving linear constraints such as shock absorber design and inertia ellipsoid analysis, which can now always be used with confidence.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Professor Eric Lanteigne of the University of Ottawa for generously providing the French translation of the title page.

REFERENCES

1. Hayes, M.J.D. "Maximum Area Ellipses Inscribing Specific Quadrilaterals." *Proceedings of the 2016 CSME International Congress, Kelowna, BC, Canada*, 2016.
2. Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. *Numerical Recipes in C, 2nd Edition*. Cambridge University Press, Cambridge, England, 1992.
3. Marquet, F., Krut, S. and Pierrot, F. "ARCHI: A Redundant Mechanism for Machining with Unlimited Rotation Capacities." *Proceedings of ICAR*, 2001.
4. Krut, S., Company, O. and Pierrot, F. "Velocity Performance Indexes for Parallel Mechanisms with Actuaion Redundancy." *Proceedings of the Workshop on Fundamental Issues and Future Research Directions for Parallel Mechanisms and Manipulators,* Québec, QC, Canada, pp 46-56, Oct. 2002.
5. Yoshikawa, T. "Manipulability of Robotic Mechanisms." *International Journal of Robotics Research*, Vol. 4, No. 2, 1985.
6. Krut, S., Company, O. and Pierrot, F. "Velocity Performances Indexes for Parallel Mechanisms with Actuation Redundancy." *Robotica*, Vol. 22, No. 2, pp. 129–139, 2004.
7. Arnold, V.I. *Mathematical Methods of Classical Mechanics, 2nd ed., translated from original 1974 Russian edition by K. Vogtmann and A. Weinstein*. Springer-Verlag, New York, N.Y., U.S.A., 1989.
8. Ginsberg, J.H. *Advanced Engineering Dynamics*. Harper & Row, Publishers, New York, N.Y., U.S.A., 1988.
9. Parker, V.W. and Pryor, J.E. "Polygons of Greatest Area Inscribed in an Ellipse." *The American Mathematical Monthly*, Vol. 51, No. 4, pp. 205–209, April, 1944.
10. Horwitz, A. "Finding Ellipses and Hyperbolas Tangent to two, three, or four Given Lines." *Southwest Journal of Pure and Applied Mathematics*, Vol. 1, No. 1, 2002.
11. Horwitz, A. "Ellipses of Maximal Area and of Minimal eccentricity Inscribed in a Convex Quadrilateral." *Australian Journal of Mathematical Analysis and Applications*, Vol. 2, No. 1, January, 2005.
12. Horwitz, A. "Ellipses Inscribed in Parallelograms." *Australian Journal of Mathematical Analysis and Applications*, Vol. 9, No. 1, January, 2012.
13. Hayes, M.J.D., Zsombor-Murray, P.J. and Gfrerrer, A. "Largest Ellipse Inscribing an Arbitrary Polygon." *Proceedings of the 19th Canadian Congress of Applied Mechanics (CANCAM 2003), University of Calgary, Calgary, AB, Canada*, 2003.
14. Behrend, F. "Über die kleinste umbeschriebene und die größte einbeschriebene Ellipse eines konvexen Bereichs." *Math. Ann.,* vol 115, pp 379-411, 1938.
15. Ceccerelli, M., ed.. *Distinguished Figures in Mechanism and Machine Science, Their Contributions and Legacies Part 1*. Springer, New York, U.S.A., 2007.
16. Veblen, O. and Young, J.W. *Projective geometry*. Volume I. Blaisdell, 1938.
17. Gfrerrer, A. "The Area Maximizing Inellipse of a Convex Quadrangle." private communication, December 19, 2002.
18. Glasser, G., Stachel, H. and Odehnal, B. *The Universe of Conics*. Springer-Verlag, New York, Berlin, DE., 2016.
19. Fishback, W.T. *Projective and Euclidean Geometry,* 2nd edition. John Wiley & Sons, Inc., New York, N.Y., U.S.A., 1969.
20. Coxeter, H.S.M. *Non-Euclidean Geometry,* 5th edition. University of Toronto Press, Toronto, On., Canada, 1965.
21. Ayres, F. *Projective Geometry*. Schaum's Outline Series in Mathematics, McGraw-Hill Book Company, New York, N.Y., U.S.A., 1967.
22. Gans, D. *Transformations and Geometries*. Appleton-Century-Crofts, New York, N.Y., U.S.A., 1969.
23. Salmon, G. *Lessons Introductory to the Modern Higher Algebra,* 4th edition. Hodges, Foster, and Figgis, Dublin, Rep. of Ireland, 1885.
24. Salmon, G. *A Treatise on Conic Sections, 6th edition*. Longmans, Green, and Co., London, England, 1879.
25. Chari, V. and Pressley, A. *A Guide to Quantum Groups*. Cambridge University Press, London, England, 1994.