A Completely Generalized and Algorithmic Approach to Identifying the Maximum Area Inscribing Ellipse of an Asymmetric Convex Quadrangle

by

Zachary A Copeland

A thesis submitted to
the Faculty of Graduate and Postdoctoral Affairs
in partial fulfilment of
the requirements for the degree of
Master of Applied Science
in
Mechanical Engineering

Ottawa-Carleton Institute for Mechanical and Aerospace Engineering
Department of Mechanical and Aerospace Engineering
Carleton University
Ottawa, Ontario, Canada
January 18th, 2017

Abstract

In this thesis, a novel method is presented whereby the maximum area inscribing ellipse, subject to a set of four linear constraints within a two variable system is automatically generated. Two methods currently exist with varying degrees of utility which provide solutions to generate the maximum area inscribing ellipse within a convex quadrangle: projective transformation of the unit square and unit circle to arbitrary parallelograms and trapezoids and corresponding area maximizing inscribing ellipses; as well as a method whereby a non-metric affine coordinate system is constructed for the identification of the area maximizing inscribing ellipse. Problematically, neither of these methods contain a general metric approach for all inscribing ellipses within any given asymmetric convex quadrangle. An algorithm is developed herein using the projective extension of the Euclidean plane which will always generate the entire one-parameter family of inscribing ellipses, and directly identify the area maximizing one of any given convex quadrangle, within a metric space.

Given four bounding points, no three of which are collinear, four line equations are generated which describe the convex quadrangle. Alongside the definition of a specific polar point, these five constraints identify a pencil of inscribing line conics, which is then transformed into its point conic dual for visualisation and plotting. The pencil of point conics then has its area optimised with respect to the value of its polar point, at which juncture the maximum area inscribing ellipse may be identified from the pencil of inscribing conics.

Acknowledgments

First, I would like to acknowledge the efforts of Professor Hayes in supplying me with the guidance required to produce a thesis that I am genuinely proud of. Secondly, any of the individuals in my office who endured my geometric ramblings, as well as their perpetual comedic relief and moral support. Kendal Wilson, for enduring the tedium of revising the first draft of this thesis, as well as Fred Noe for invaluable contributions to the creativity of my writing process.

Table of Contents

List of Tables

List of Figures

List of Symbols

Nomenclature	Definition
A, B, C, D, E, F	Vertices of a quadrangle
T	Midpoint of Diagonal AC
S	Midpoint of Diagonal BD
$\mathbf{m}$	Position vector of an ellipse centre on line segment TS
O	Intersection of diagonals AC and BD
P	Intersection of diagonals AC and EF
Q	Intersection of diagonals BD and EF
O, P, Q	Vertices of the diagonal triangle
o, p, q	Edges of the diagonal triangle
k	Conic pencil equation
$\mathbf{A}_L$	Line conic shape coefficient matrix
$\mathbf{A}$	Point conic shape coefficient matrix
Δ	Determinant of $\mathbf{A}$

Δ_0	Discriminant of $\mathbf{A}$
$\alpha(u)$	Squared area function

Chapter 1

Introduction and Problem Statement

In this thesis, an algorithm that identifies the ellipse with the largest area from the one parameter family of ellipses that inscribe a general convex quadrangle is presented for the first time. This algorithm and solution methodology have several important applications ranging from determining the maximum area covariance ellipse given a set of variables, or design constraints, to characterising the kinematic isotropy and velocity performance indexes of parallel mechanisms. This kinematic isotropy can be used to determine the area of the workspace in which the parallel mechanism possesses the largest kinematic variability, providing information about where the mechanism is most capable of changes in direction and velocity. Specifically, this algorithm identifies the largest ellipse lying inside of any given convex quadrangle, with tangents on each of the four sides, seen in Figure 1.1.

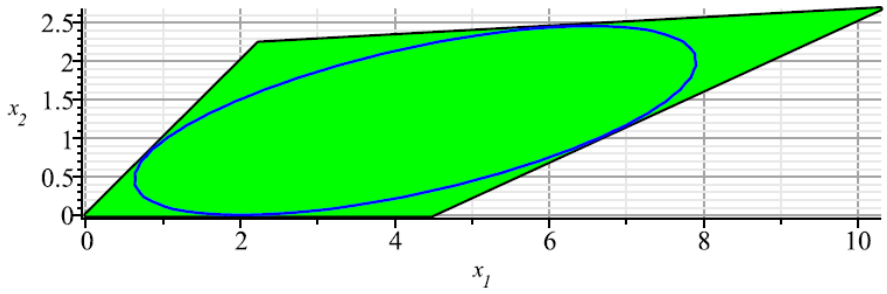


Figure 1.1: The maximum area inscribing ellipse of a convex quadrangle.

1.1 Covariance Ellipses

Within systems of variables, covariance is a measure of how changes within one given 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 its counterpart. Currently, covariance ellipses are generated in many fields of study in order to analyse data sets in an effort to understand the physical processes or relations which are present within a given system. While ellipses are typically thought of as second order curves which bound a closed area, a more relevant representation is that which is obtained from a matrix formulation. This matrix has dimensions $n \times n$, such that n is representative of the number of variables contained within the system in question. Specifically, a covariance matrix, $\mathbf{A}$, for two given variables will have the form,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ \\ a_{12} & a_{22} \end{bmatrix}, \tag{1.1}$$

where $\mathbf{A}$ is a symmetric, square matrix whose values are directly related to the linear correlations measured between the two variables in question. Within statistics, the covariance ellipse of c separate variables, given N data points can be generated as follows:

$$\mathbf{V} = \begin{bmatrix} \frac{\Sigma x_1^2}{N} & \frac{\Sigma x_1 x_2}{N} & \cdots & \frac{\Sigma x_1 x_c}{N} \\ \\ \frac{\Sigma x_2 x_1}{N} & \frac{\Sigma x_2^2}{N} & \cdots & \frac{\Sigma x_2 x_c}{N} \\ \\ \vdots & \vdots & \ddots & \vdots \\ \\ \frac{\Sigma x_c x_1}{N} & \frac{\Sigma x_c x_2}{N} & \cdots & \frac{\Sigma x_c^2}{N} \end{bmatrix}. \tag{1.2}$$

From this representation, it is clear to see that the diagonal of V represents the

variance of each variable within the data set, however, each non-diagonal element v_{ij} represents the covariance of each variable with one of its experimental counterparts. Within a two variable system, $\mathbf{V}$ simplifies to a 2×2 symmetric matrix, which possesses a form identical to that of the quadratic form of an ellipse.

1.2 Velocity Performance Indexes for Parallel Mechanisms

Performance indexes within machine design are often used as a tool whereby information about a specific mechanisms architecture may be obtained, and compared to other dissimilar designs, in order to make an effective decision which will allow for maximum productivity given a variety of constraints. Parallel mechanisms with actuation redundancy are mechanisms whose operational force outputs are not unique; these operational output forces do not correspond to a unique set of joint forces, which may help in reducing the effect of over-mobility singularities [3, 4]. Analysis of the kinematic isotropy, or the capacity of a mechanism to change position, orientation, and velocity given its pose within the workspace, can provide important information about their velocity performance indexes [5].

In this context, the area of the 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. Kinematic isotropy can be used to provide information about the velocity performance of a mechanism during actuation based upon its position within its workspace. In [3] the approach to identifying the maximum area inscribing ellipse is a numerical problem, essentially fitting the inscribing ellipse within the linear constraints defining the velocity profile of the mechanism by starting with the unit circle.

1.3 Literature Review

There are only a handful of papers that report on investigations into determining the maximum area inscribing ellipse contained within arbitrary quadrilaterals within the literature. Determining the maximum area inscribing polygon of a given ellipse has been investigated in [6], but is not necessarily germane to the application of identifying the maximum area inscribing ellipse within an arbitrary quadrangle. Investigations in [7] show that there exists a unique solution to the identity of the maximum area inscribing ellipse which lies tangent to four non-coincident lines, while further investigation in [8] shows that all ellipses lying tangent to these four lines have centres lying on single line contained within the quadrangle. Problematically, these investigations are simply proofs used to showcase that the solutions exist, not explicit algorithms which can be used in order to identify the specific equation of the maximum area inscribing ellipse.

In [9] a projective transformation technique is used in order to map the maximum area inscribing ellipse of a known quadrangle, in this case the unit circle, onto an ellipse inscribing an asymmetric convex quadrangle. Problematically, this solution will only provides the maximum area inscribing ellipse for specific quadrangles, and fails in the event of an asymmetric convex quadrangle. In [2] a solution is identified through the use of non-orthogonal coordinate bases generated based upon the geometry of the quadrangle presented; which, while effective from a purely mathematical standpoint, fails to yield a metric result in general, resulting in a solution that is difficult to implement directly either for the purposes of identifying a covariance ellipse, or for characterising maximum kinematic isotropy.

1.4 Ellipses Tangent to Four Lines

Within its most simple incarnation, the problem of describing an ellipse which adequately represents the covariance between variables in a two variable system simplifies to describing the maximum area ellipse contained within a quadrangle which generates four non-collinear lines. In a system of two linear variables, x and y, these variables can be described as minimum, x_1 and y_1, as well as maximum values, x_2, y_2. The minimum and maximum values of the two variables describe four bounding points. These four bounding points, describe four bounding lines provided that no three of the points are collinear. Given four lines that form a convex quadrilateral, there exists a pencil of conics which lies tangent to all four of these lines [7], and specifically, one single ellipse of maximal area exists within this pencil [8].

Provided that the point couples (x_1, y_1), (x_2, y_1), (x_2, y_2) and (x_1, y_2) generate a convex quadrangle, it is possible to identify a unique ellipse contained entirely within the bounds of this polygon which occupies the maximum possible area within the shape. Given this capacity, it becomes a simple matter to generate a covariance ellipse which covers the largest possible area contained within these linear constraints; or an ellipse describing the kinematic isotropy of a redundantly actuated parallel mechanism.

1.5 Objectives

While it is well accepted that a maximum area inscribing ellipse exists within a convex quadrangle, to the best of the author's knowledge, a stand alone solution as to the exact equation of this ellipse, given the formulation of either a quadrilateral or a quadrangle, is missing from the vast body of literature. Thus, the methodology following herein has been developed in such a way so as to provide a solution for this

problem such that the resulting equation is both metric, and fully general, given any convex quadrangle configuration.

1.5.1 Statement of Originality

Certain aspects of maximum area ellipses inscribing asymmetric convex quadrangles are presented herein for the first time. Of particular interest are the following contributions.

- **Attached, in Appendix A is an algorithm which, given four points, will automatically identify the maximum area inscribing ellipse of the convex quadrangle they define.** This algorithm includes a full solution, carried out in Maple, which identifies the maximum area inscribing ellipse contained within a quadrangle defined by a user as four (x, y) coordinates, input in counter-clockwise order starting from its lowest point.
  - Within this solution, a generalised determinant-based formulation used to determine the area of a given closed central conic is published for the first time in English; its proof can be found in Chapter 2, using a canonically positioned conic section and showcasing that its construction leads to an identical area formulation as the commonly accepted πab.
- **Throughout the summary of previous solutions contained within Chapter 3 several characteristics are identified with respect to the suitability of the solutions, specifically the geometric conditions under which each solution is ineffective.** Specifically, information about the mapping technique presented within [9] and [10] is uncovered and effectively showcases the reasoning for the situational applicability of this methodology given what should be a completely general mapping process.

  - An English translation of portions of [1] is used to develop a Euclidean classification for conic sections given their matrix-vector formulation and based solely upon the values of associated Euclidean invariants derived thereof.

- **The maximum area inscribing ellipse of a trapezoid possesses an area that is, at most, π directly scaled by two numbers which are not necessarily the length of the semi-major and semi-minor axes.** In general, the numerical value of the area for the maximum area inscribing ellipse of a convex quadrangle, when given in rational form, is a relatively cumbersome function comprised of all manner of mathematical operators; in the case of the trapezoid, the maximum area inscribing ellipse possesses an area that can be described by one rational number, the square root of a whole number, and π, multiplying each other.

1.5.2 Thesis Outline

In order to adequately describe the generalised solution for determining the maximum area inscribing ellipse within a convex quadrangle, its significance, and furthermore its novel nature, a large body of mathematical background is presented in Chapter 2; this background information includes, but is not necessarily limited to, an overview of linear geometries, transformation groups, line and point coordinates, second order curves, closed conic sections (both in their line and point forms), and concludes with a description and proof of a generalized area function based upon the matrix-vector formulation of any conic section.

Once the requisite mathematical background information has been presented, Chapter 3 presents two previously identified, but non-general, solutions for generating the maximum area inscribing ellipse within a convex quadrangle. Chapter 4

outlines, in full, the novel general solution to identify the maximum area ellipse inscribing a convex quadrangle. Several working examples are included so as to show that the method produced herein agrees with geometrically accepted solutions, while a general convex quadrangle has its maximum area inscribing ellipse identified to conclude the discussion of the methodology.

Finally, Chapter 5 provides a summary, conclusions, and a comparison of each solution methodology, as well as interesting properties of inscribing ellipses within specific convex quadrangles, specifically the trapezoid, which were observed throughout the course of the testing procedure.

Chapter 2

Mathematical Background

Complete comprehension of the methodologies presented herein necessarily requires a large body of knowledge pertaining to quadrilaterals, quadrangles, and second order curves, and in particular, geometry. Although the background presented within this chapter is sufficient for the purposes of communicating the significance of the problem at hand, it is to be noted that this is, by no means, a comprehensive overview of these topics.

2.1 Geometric Definitions

Though algebraic and geometric representations of functions are invaluable within engineering and design, most engineers are only familiar with Euclidean spatial constructions, as they tend to be exceedingly advantageous for describing functions and real world problems. Euclidean spaces are an extremely useful subset of spaces for their metric qualities, however they represent only a small subset of spaces within an infinite number of spaces.

Within the scope of this problem, discussion will be limited to three spaces; projective, affine, and Euclidean. Although theorems will be presented upon which the construction of any given space may be undertaken.

2.1.1 Euclidean Geometry

Overwhelmingly, within engineering, Euclidean spaces are used to represent and define functions and geometric constructions. Euclid first pioneered a formal concept for geometry nearly 2300 years ago. Originally, five postulates, or self evident truths, which lay the foundation for geometry are presented within Book I of The Elements [11].

1. A straight line may be drawn between two distinct points.
2. A finite straight line may be produced to any length in a straight line.
3. A circle may be described with any centre and any distance from that centre.
4. All right angles are equal.
5. For each line L and point P not on L, there exists one and only one line passing through P, parallel to L.

Through inspection, it is obvious that for what is now known as Euclidean geometry, the above postulates hold true. Euclid did propose many additional postulates and axioms which provide a truly complete view of Euclidean Geometry as a whole, which are all built upon these five postulates.

Non-Euclidean Geometries

However, the most commonly debated topic within Euclidean geometry was the fifth postulate, also known as the *parallel postulate*; throughout the course of this debate, discussion lead to the discovery of additional non-Euclidean geometries [12, 13], such as hyperbolic and elliptical geometries, where this postulate is pushed to opposing extremes. Specifically, the non-Euclidean geometries were developed in such a way

so as to attempt find a contradiction within Euclidean geometry itself, but this development resulted in spaces which, while neglecting the concept of parallelism, did not contradict any of the additionally posed postulates.

Let the points of the hyperbolic plane be the points on the interior of a circle in an ordinary Euclidean plane, while hyperbolic lines are the chords of the circle, with their end points on the circumference of the circle excluded. Within hyperbolic geometry, given a line L and a point P not on L, there are infinitely many distinct lines passing through P, parallel to L. Figure 2.1 is a two dimensional depiction of what hyperbolic spaces may look like, with infinitely many lines passing through point P parallel to the L [14].

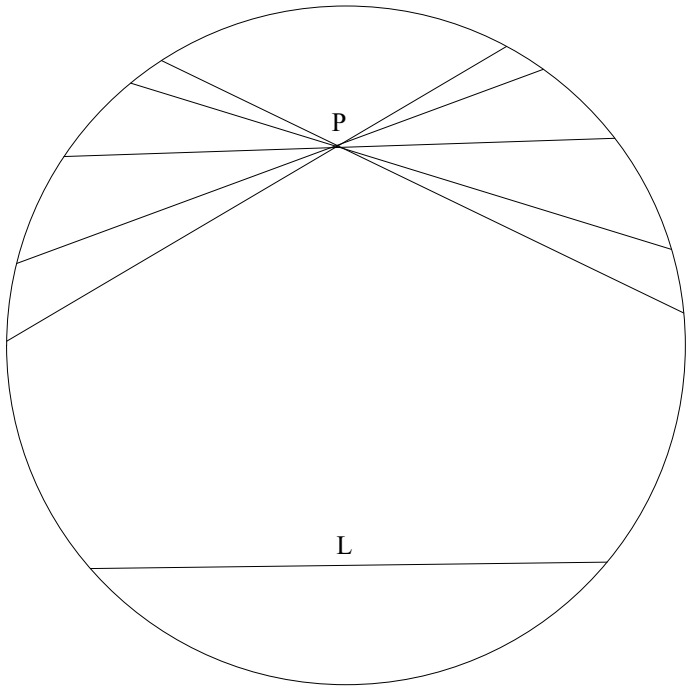


Figure 2.1: Illustration of the hyperbolic plane, with line L and a subset of the infinitely many lines passing through point P which lie parallel to L.

On the opposing extreme, elliptical geometry contains no lines passing through P which are parallel to L [14]. Figure 2.2 is a three dimensional depiction of what elliptical spaces may look like. It is easiest to think of elliptical space, in this sense, as

a central projection of an ordinary Euclidean plane onto the surface of a hemisphere; a line on the surface of the sphere is a great circle of the sphere, and any great circle intersects every other great circle. In Figure 2.2, the line on the surface of the sphere represents the line at infinity of the projective extension of the Euclidean plane, from which lines a and b are projected.

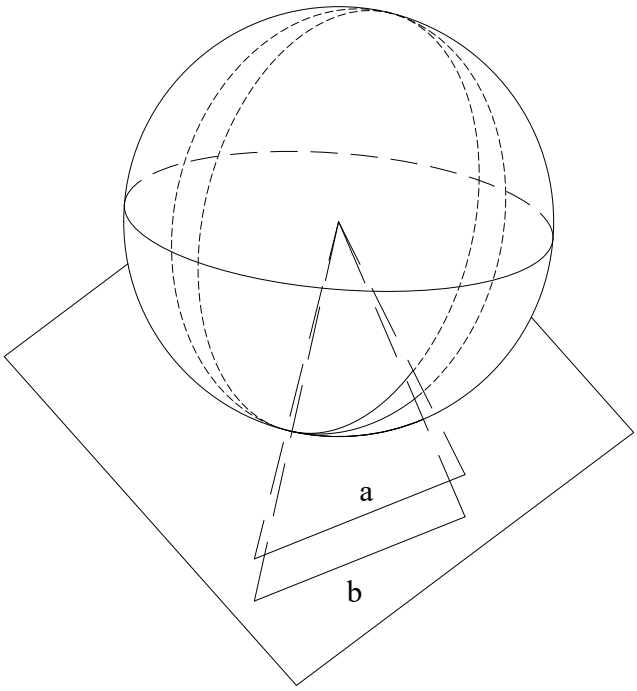


Figure 2.2: Illustration of the elliptical plane using a central projection of (parallel) lines a and b onto the surface of a sphere, creating pair of intersecting great circles.

Elliptical and hyperbolic geometries both give rise to the concept of curvature in space. The Euclidean plane can be thought of as having no curvature, while the hyperbolic plane is considered to have negative curvature (as though you are viewing the projection of a horse saddle shaped space onto a flat surface), indicating that lines grow more distant from each other as they tend towards the line at infinity, where the elliptical plane is considered to have positive curvature, indicating that the lines grow more close as they tend towards the line at infinity [14]. While straight

lines are the foundations upon which these spaces are derived, Figure 2.3 shows what triangles constructed of straight lines in Euclidean, hyperbolic, and elliptical spaces look like [14].

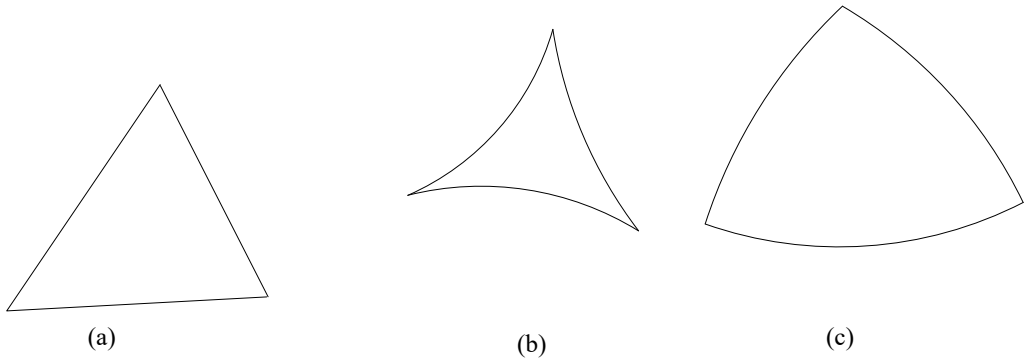


Figure 2.3: Euclidean planar triangle (a), hyperbolic planar triangle (b), and elliptical planar triangle (c).

In Euclidean geometry, the internal angles of a triangle will always sum to 180 degrees, while in hyperbolic geometry the sum of the internal angles will always be less than 180 degrees, and this same sum will be in excess of 180 degrees within elliptical geometry [14]. Although the subject of non-Euclidean geometries is a rich one, the remainder of this discussion will be constrained to the hierarchy of linear spaces. Spaces where the *parallel postulate* is either modified slightly, or not at all, namely projective and affine spaces.

2.1.2 Synthetic Projective Geometry

Although seemingly ubiquitous, Euclidean geometry is inherently limited. Specifically, Euclidean geometry fails to account for the concept of a perspective within the viewing of geometry itself. Development of projective geometry began several hundred years ago, by scientists such as Johann Kepler, Blaise Pascal, Albrecht Durer, and Gerard Desargues. Their attempts to reconcile geometry with the world around them lead to the production of axioms which were proven by Felix Klein [15] and unified under the theory of transformation groups in 1872 [16].

Projective geometry lead to the development of a hierarchy within linear geometries. In order to develop the governing theorems of projective geometry, five non-metric theorems (involving no measurement of either angles or distance) can be taken from Euclidean geometry, specifically the three dimensional Euclidean space, E_3, and modified so as to exclude the concept of parallelism. The five Euclidean theorems to be modified are what follows.

- **E1**: Two distinct points determine one and only one line.
- **E2**: Three distinct, non-collinear points determine one and only one plane.
- **E3**: Two distinct coplanar lines, existing in the same plane, either intersect in exactly one point, or are parallel.
- **E4**: A line not contained in a given plane either intersects the plane in exactly one point, or is parallel to it.
- **E5**: Two distinct planes either intersect in one line, or are parallel to each other.

When modified to exclude the concept of parallelism, the final three theorems define the corresponding projective theorems.

- **P1**: Two distinct points determine one and only one line.

- **P2**: Three distinct, non-colinear points determine one and only one plane.
- **P3**: Two distinct lines, existing in the same plane, intersect in one point.
- **P4**: A line not contained in a given plane intersects the plane in one point.
- **P5**: Two distinct planes intersect in one line.

Comparing the two sets of theorems above yields an obvious simplification within the projective theorems; they are devoid of either/or constructions. Most importantly, however, is the emergence of the concept of *duality*. For example, **P3** is obtained from **P1** through exchanging the words "point" and "line", which are the dual elements contained within the projective plane, P_2. The dual elements of projective space, P_3 are "point" and "plane"; by exchanging these two words, **P5** is obtained from **P1**.

The important implication of these statements is that for every valid theorem which exists in projective geometry, and equally valid dual theorem is guaranteed to exist [17].

2.1.3 Transformation Groups

While highly general, a discussion of the properties of different spaces within geometry is not inherently useful from the standpoint the computational application at hand. Therefore, the concept of a generalised transformation group, the unifying theory presented by Felix Klein in his Erlangen Programme [16], is the method whereby geometries will be classified herein. Transformation groups are also referred to as *structure matrices* [18], due to their description of the structure of the space which they represent.

Each linear geometry has a corresponding linear transformation group associated with it; *projective geometry*, in P_3, may be represented by a 4×4 matrix operator called a

collineation,

$$\mathbf{P} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix}, \tag{2.1}$$

where collinear points are transformed to collinear points. All elements of matrix $\mathbf{P}$ are arbitrary real numbers. The only restriction placed on $\mathbf{P}$ is that its determinant can not vanish, and due to the arbitrary nature of each individual element, relative to the value of any one of the 16 entries, there are only 15 independent elements contained within this matrix. Thus, the *projective group* represents the most general group of linear geometries in three-dimensional space, referred to as G_{15} [12].

From the concept of a transformation group, affine geometry may be developed; A_3, or three dimensional affine space is represented by a 4×4 collineation,

$$\mathbf{A} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{2.2}$$

It is clear from Equation (2.2) that the affine transformation group has four constrained elements; in this example the final row of the matrix. Thus, the affine transformation group is a member of G_{12}. Due to the unconstrained nature of the values contained within G_{15}, it is apparent that $G_{12} \subset G_{15}$, thus affine geometry is

a constrained subset of projective geometry, and any affine group is necessarily also a projective group, however it is also inherently less general than G_{15}. From this definition, the *principal Euclidean group* can be developed as yet another subset of the *affine group*, and logically, the *projective group*:

$$E_3 = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{2.3}$$

While Equation (2.3) and Equation (2.2) are identical in their representation, the *principal Euclidean group* has its values constrained so as to produce a 3×3 proper orthogonal submatrix, with a determinant of exactly +1 [19]. This constraint yields the transformation group G_7, representing the most general Euclidean collineations, and it is obvious that $G_7 \subset G_{12} \subset G_{15}$. Typically, the principal Euclidean group is further constrained in order to preserve distances, and is referred to as the Euclidean displacement group; this transformation group is what is commonly used within geometric transformations within the plane. Thus, within the linear spaces, a hierarchical nature is uncovered; projective spaces can yield affine spaces, which can, in turn, yield Euclidean spaces. Thus, every Euclidean space is necessarily a specific subset of affine space, and every affine space is a specific subset of projective space [12].

2.1.4 Invariants

Within the context of spaces, and their defining transformation groups, the concept of an *invariant* arises naturally [20]. While transformation groups are useful for classifying the geometry at hand, these transformation groups preserve certain properties

of the points contained within their respective spaces [17, 21]; upon the application of a transformation group on a set of points, certain properties of these points will remain *invariant*. The number of invariants is inversely proportional to the number of independent elements of the transformation group; affine spaces possess more invariants than projective spaces, and fewer invariants than Euclidean spaces.

The most general three dimensional linear space, P_3, contains only one invariant in the transformation group G_{15}, namely the cross ratio of four collinear points. Given four collinear points, A, B, C, D, at least three of which are distinct, having homogeneous coordinates on the line; $(a_0 : a_1)$, $(b_0 : b_1)$, $(c_0 : c_1)$, $(d_0 : d_1)$, then the real number defined by

$$CR(\mathrm{AB};\mathrm{CD}) = \frac{\begin{vmatrix} a_0 & a_1 \\ c_0 & c_1 \end{vmatrix} \begin{vmatrix} b_0 & b_1 \\ d_0 & d_1 \end{vmatrix}}{\begin{vmatrix} b_0 & b_1 \\ c_0 & c_1 \end{vmatrix} \begin{vmatrix} c_0 & c_1 \\ d_0 & d_1 \end{vmatrix}} \tag{2.4}$$

is referred to as the cross ratio of the four points, A, B, C, D in that order; this number is always preserved by any transformation group that is a subset of G_{15}. Each subset of G_{15} possesses additional invariants. Affine transformations, for example, preserve the plane at infinity, and therefore parallelism, as well as preserving the ratios of distances along parallel lines, and thus ratios of areas, and finally, the similarity between shapes. Additionally, Euclidean displacement transformations possess the same invariants as those of affine geometry, but with the addition of invariants such as distance, angle, and congruence; two objects are congruent if they can be transformed into each other through a combination of rotations and translations.

2.2 Homogeneous Coordinate Triples

Given any coordinate system, (x_1,x_2) containing the origin, O, and a point S, it is possible to construct a ray, OS, to describe the coordinates of point S, (x, y). Now, given any point Q which is not incident with the origin, but lies on the ray OS, which is represented by the point $(\mu x, \mu y)$, such that μ is a real number, it can be seen that ray OQ is an extension of ray OS, by a factor of μ, seen in Figure 2.4.

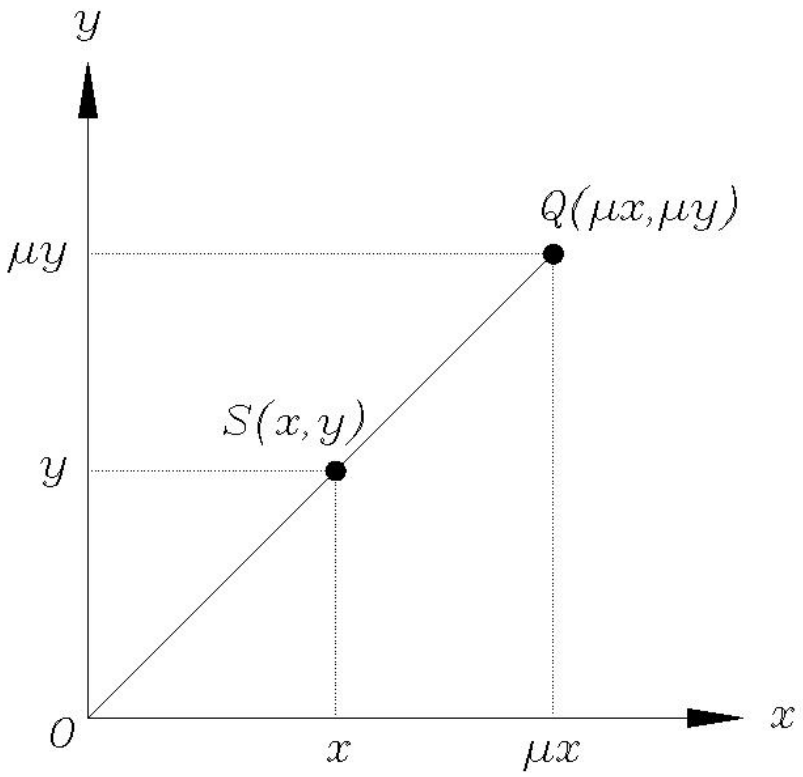


Figure 2.4: Ray OS, scaled by μ, yields ray OQ.

While assuming that the scaling factor μ is any real number, the extension necessarily will yield the seemingly meaningless coordinate pair (∞,∞) as $\mu \to \pm\ \infty$. However, to make sense of this, the Euclidean plane can be extended to include the projective line at infinity. Thus, in order to include this point at infinity, the Euclidean space can be extended and represented by the ordered triple (x_0:x_1:x_2). Provided that x_0 is not equal to zero:

$$\begin{aligned} x &= \frac{x_1}{x_0}; \\ y &= \frac{x_2}{x_0}. \end{aligned} \tag{2.5}$$

Given Equation (2.5), any scalar multiple of the ordered triple, (μx_0,μx_1,μx_2), will recover the same coordinate pair as (x_0,:x_1:x_2). That is to say,

$$\begin{aligned} \frac{\mu x_1}{\mu x_0} &= \frac{x_1}{x_0}, \\ \frac{\mu x_2}{\mu x_0} &= \frac{x_2}{x_0}. \end{aligned} \tag{2.6}$$

Therefore, for any μ, or scalar multiple of the ordered triple, the same (x,y) coordinate pair will be obtained once this ratio is computed. The coordinates (x_0:x_1:x_2) are referred to as *homogeneous coordinates*. Considering the previous example of ($\mu x, \mu y$):

$$\begin{aligned} (\mu x, \mu y) &= (x_0, \mu x_1, \mu x_2); \\ (x_0, \mu x_1, \mu x_2) &= (\frac{x_0}{\mu}, x_1, x_2). \end{aligned} \tag{2.7}$$

Within the context of homogeneous coordinates, as $\mu \to \pm \infty$, the trivial point of (∞,∞) is represented with the homogeneous point triplet of (0:x_1:x_2). Referred to as an *ideal point*, this point represents the point at which the ray OS intersects the line at infinity, $x_0 = 0$; regardless of whether or not $\mu \to \pm \infty$, the exact same homogeneous point triple is recovered. Thus, any line can be thought of as representing a closed curve with either end intersecting at the same point on the line

at infinity. Any two parallel lines within the plane share a common ideal point, and thus intersect on this point on the line at infinity.

While this construction facilitates the use of many mathematical operations, it does not have an effect on the Cartesian point coordinate value. Provided that the triple's ratio is evaluated with $x_0 = 1$, the original Cartesian coordinate (x, y) will be recovered. Though this construction has taken place entirely within E_2, or 2-D Euclidean space, it is entirely possible to extend the generalization to E_3, and use the homogeneous quadruple of $(x_0{:}x_1{:}x_2{:}x_3)$.

2.3 Four Sided Polygons

Although it is commonly understood that a four sided polygon is a figure comprised of four sides alongside four vertices which serve to necessarily create an enclosed area, the specificity of this problem necessitates a more precise depiction of the properties of four sided figures, alongside how they are constructed.

2.3.1 Quadrangles and Quadrilaterals

In general conversation it is typical to refer to any four sided figure as a quadrilateral, however, most cases of the creation of a four sided figure are actually undertaken in such a way so as to properly define a quadrangle [17]. The differences between a quadrangle and its line counterpart, a quadrilateral, are as subtle as they are fundamental to the nature of these shapes [17]. Regardless, one is the dual of the other.

Specifically, four sided figures may be broken into two categories based on their construction; they may either be a complete quadrangle, or a complete quadrilateral. Complete quadrangles are created through four mutually coplanar *points*, typically referred to as vertices, no three of which are collinear. These four points are coupled

together to generate four lines, where the triangle determined by the intersections of the opposite sides of the complete quadrangle is referred to as its diagonal triangle. This diagonal triangle has three vertices, referred to as the diagonal points of the quadrangle [17]. Figure 2.5 shows a quadrangle comprised of points A, B, C, and D, its vertices, while dotted lines connect the vertices of the diagonal triangle, points E, F, and G.

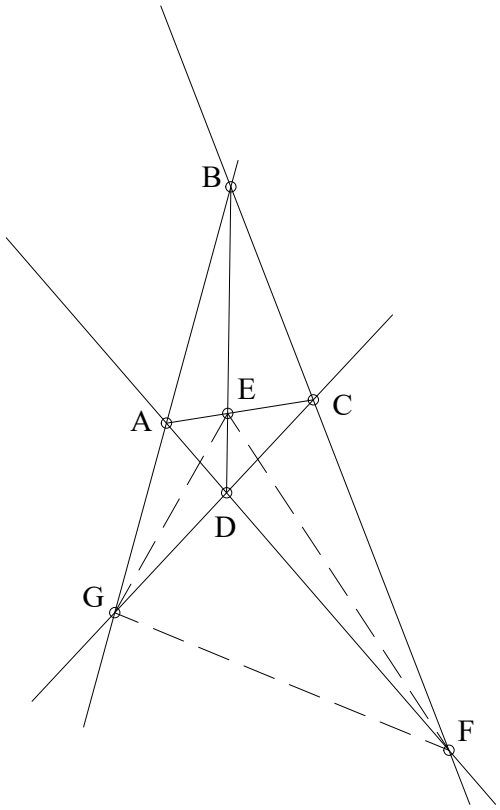


Figure 2.5: Convex quadrangle comprised of points A, B, C, D, and its diagonal triangle, composed of points E, F, G.

Complete quadrilaterals, however, are defined by the dual configuration consisting of four mutually coplanar *lines*, no three of which are concurrent. These four lines intersect in six points, and form one four sided closed figure; the quadrilateral. Joining the opposite vertices of the quadrilateral, it is possible to construct a diagonal trilateral, such that the sides of the diagonal triangle are the diagonal lines of the

quadrilateral [17]. Figure 2.6 shows a complete quadrilateral generated through lines a, b, c, and d, whose intersection points define a four sided closed figure, while lines e, f, and g define the diagonal trilateral of the complete quadrilateral.

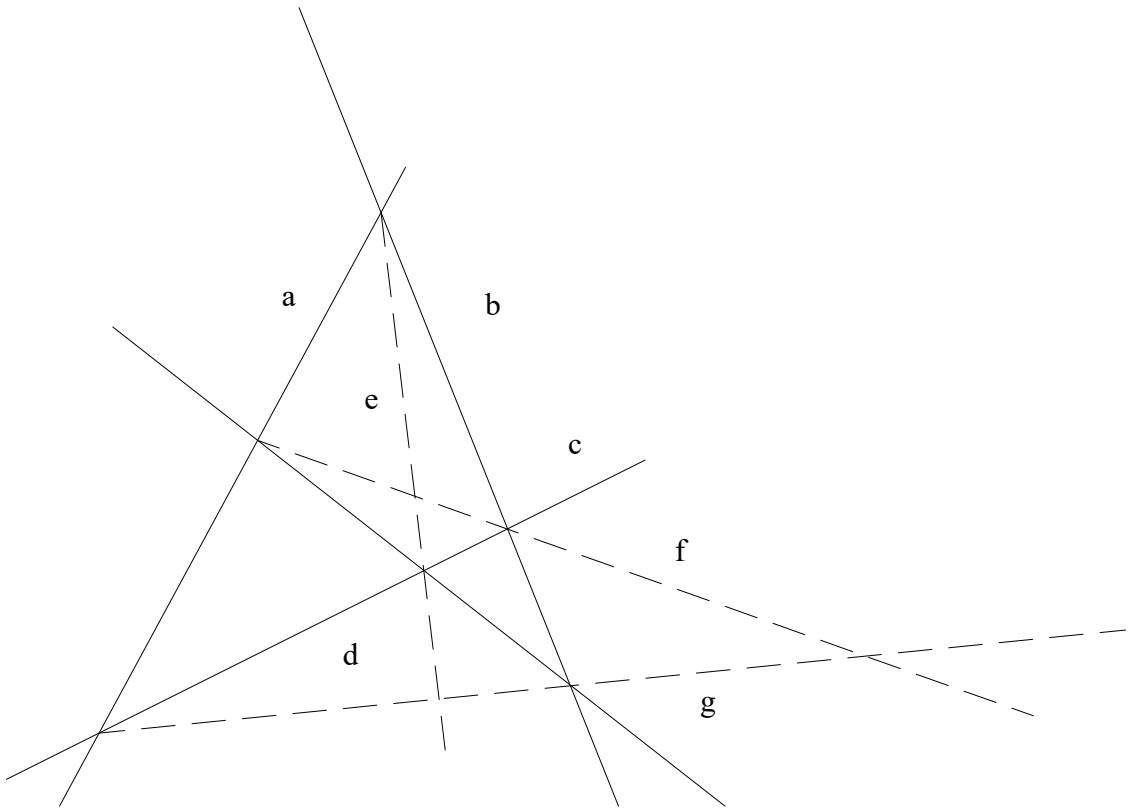


Figure 2.6: Convex quadrilateral comprised of lines a, b, c, d, and its diagonal trilateral, composed of lines e, f, and g.

In order to fully characterize the proposed solutions and their methods, it is important to describe, in detail, the differences and defining characteristics of complete quadrangles in a general sense. Specifically, the methods proposed herein will focus on shapes referred to as *convex*. Convex quadrangles differ from general quadrangles in that all of their internal angles are less than 180 degrees, and no two opposite sides intersect with each other. Concave quadrangles exist in the event that any one interior angle of the quadrangle exceeds 180 degrees, while no two opposite sides intersect, and crossed quadrilaterals exist in the event that two opposite sides intersect

in a point that is not a vertex of the quadrangle. While quadrangles and quadrilaterals are fundamentally different, they both lead to a construction of a closed, four sided figure, that can be classified as one of these three types; Figure 2.7 depicts the differences between these three types of quadrangles.

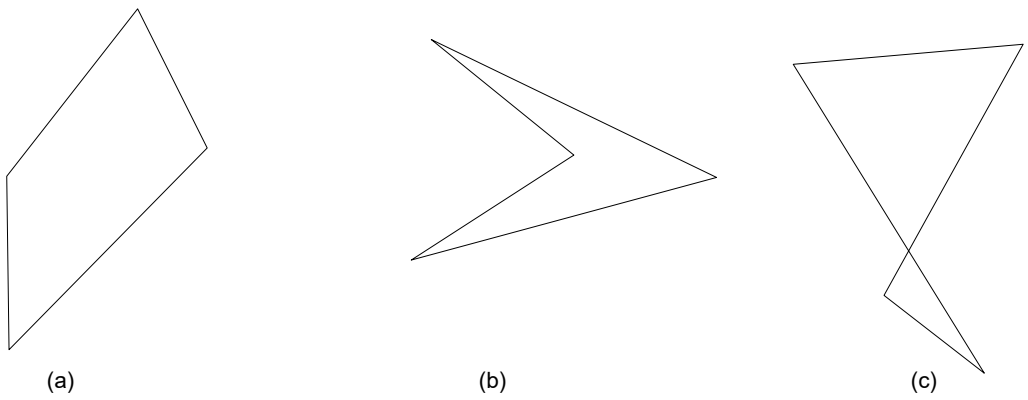


Figure 2.7: Convex (a), concave (b), and crossed quadrangles (c).

While all three types of quadrangles are geometrically valid shapes, the analysis contained henceforth will be concerned solely with the examination of the properties of convex quadrangles. While quadrangles and quadrilaterals both facilitate the construction of a closed, four sided figure, quadrangles inherently lend themselves to characterisation more easily than their quadrilateral counterpart. As such, they will be used to provide more specific characterisations of four sided convex shapes. Characterisation of quadrangles can be undertaken through the examination of the closed four sided figure created through joining lines defined by four mutually coplanar points, no three of which are collinear, alongside the two external intersection

points of these lines. Special care will be taken in order to expound upon the information contained in the construction of the quadrangle so as to provide a location for the locus of the centres of the pencil (one parameter family) of inscribing ellipses bound by the quadrangle.

In the Euclidean plane, given that all of the internal angles of a quadrangle must sum to 360 degrees, the primary variables that will define the shape, and ultimately the properties of the quadrangle, are the relative magnitudes of the side lengths to each other, as well as the angles separating them. Specifically, Figure 2.8 shows a square; squares arise when all internal angles and side lengths of a quadrangle are equal. Cutting the square into sections by using the diagonals also shows a fifth point, directly in the centre of the quadrangle, where the midpoints of the diagonals intersect. This point is indicated as the doubly mapped points S and R, in Figure 2.8, and represents the unique centre of all inscribing ellipses contained within the square.

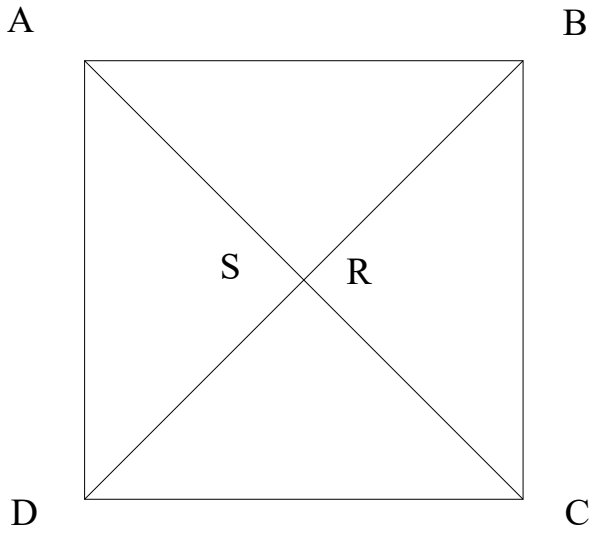


Figure 2.8: Square ABCD with its diagonals AD and BC, alongside the centre of all inscribing ellipses, the midpoints of the diagonals, S and R, which lie incident with each other in this case.

Holding all angles equal, but allowing two of the side lengths to vary from each

other yields the rectangle. Similar to the square, cutting the rectangle into sections via its diagonals shows a single point which represents the unique centre of all inscribing ellipses within this quadrangle. Figure 2.9 shows this critical point as the identical points S and R. Much like the square, the centre of all inscribing ellipses contained within the rectangle lies in the geometric centre of the quadrangle, equidistant from each opposite set of parallel edges.

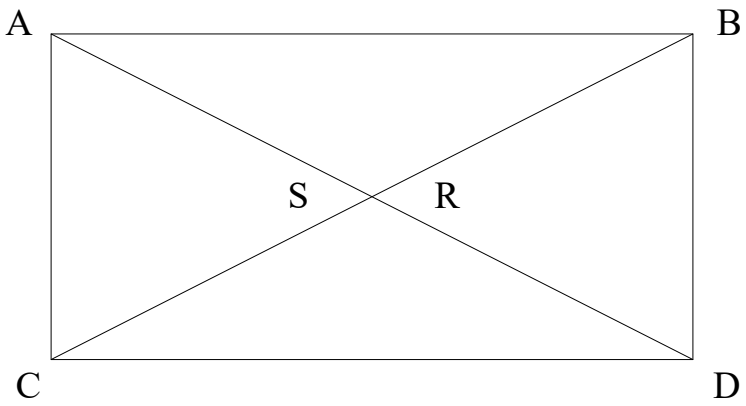


Figure 2.9: Rectangle ABCD with its diagonals AD and BC, alongside the centre of inscribing ellipses, the midpoints of the diagonals, S and R, which lie incident with each other in this case.

Conversely, allowing the interior angles to vary while maintaining the equal side lengths of the square facilitates the creation of a parallelogram; each pair of sides remains parallel to each other, though these quadrangles do not contain a single right angle. Figure 2.10 also shows that the locus of centres for the inscribing ellipses lies on a single point, once again in the geometric centre of the parallelogram. While a parallelogram with two pairs of unequal edge lengths is presented in Figure 2.10, it should be noted that the same properties exist within any parallelogram with each pair of opposite sides being equal in length. In the special case of the parallelogram with equal sides but interior angles which vary from 90 degrees, the rhombus is generated;

squares, rectangles, and rhombuses are all specific cases of parallelograms.

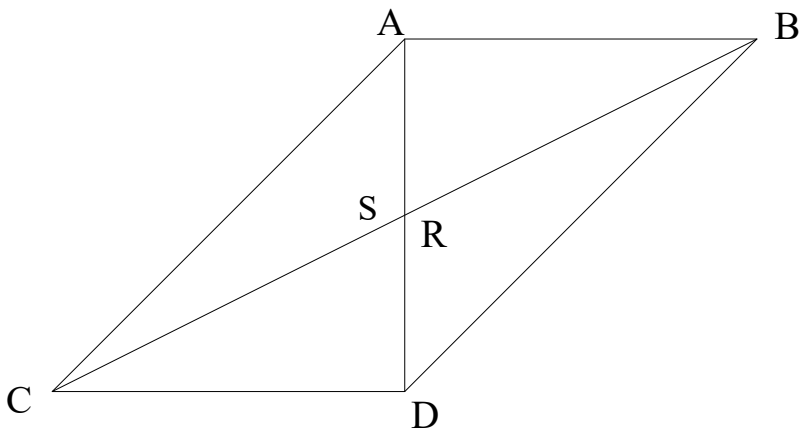


Figure 2.10: Parallelogram ABCD with its diagonals AD and BC, alongside the centre of all inscribing ellipses, the midpoint of the diagonals, S and R, which in this case are incident with each other.

While the Euclidean definitions of these three shapes have distinct differences in order to separate their representations, the square, rectangle, and rhombus are simply special cases of the parallelogram. Thus, if viewed from the perspective of constructing a complete quadrangle with four coplanar points, the distinction between their shapes is ultimately lost. Specifically, given four coplanar points, as defined previously, each of the parallelogram, rectangle, square, and rhombus shapes possess four definite vertices, which establish the closed four sided figure represented in Euclidean space. However, the remaining two intersection points lie as distinct, unique points on the line at infinity, representing the intersection of the pencil of every line parallel to the opposite sides of the parallelogram. Figure 2.11 shows three parallelogram cases, and that their quadrangle definitions yield ideal intersection points for the vertices of their diagonal triangles.

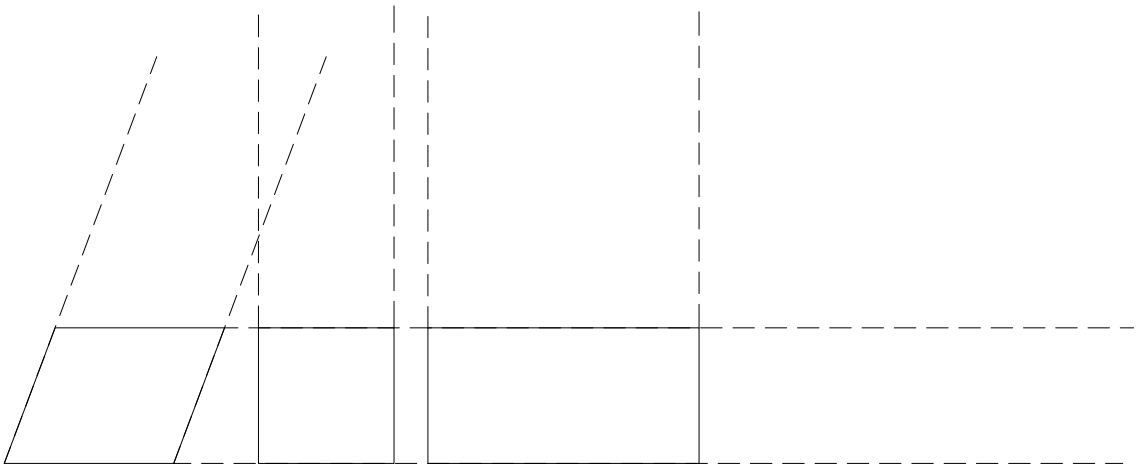


Figure 2.11: Three separate, unique cases of parallelogram construction which showcase affine similarity.

While each of the three cases presented in Figure 2.11 represent distinct quadrangles, they share the idea of symmetry along at least one axis [22]. In the case of the rectangular parallelogram, the symmetry is in the form of a reflection about either of its diagonals. Problematically, this symmetry condition necessarily restricts the locus of the centre of ellipses to be contained in exactly one point; the geometric centre of the quadrangle. Though this property is decidedly interesting, and can simplify the analysis of some specific cases of quadrangles, it is necessarily not the general case for all convex quadrangles.

In the event that two pairs of sides have equal lengths, but lie adjacent to one

another as opposed to opposite one another, the quadrangle is referred to as a symmetric kite. Figure 2.12 shows kite ABCD, alongside the intersection of diagonals BC and AD, with the locus of the centres of all inscribing ellipses, line segment SR, lying incident with diagonal BD.

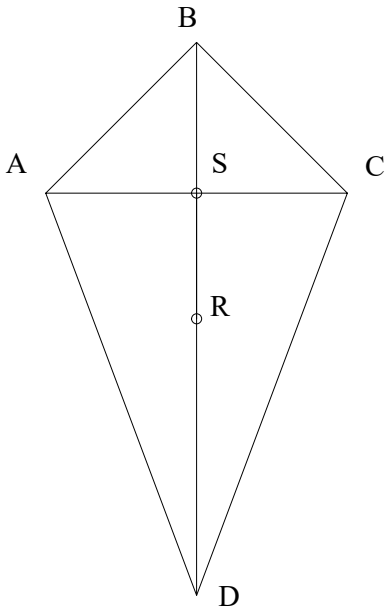


Figure 2.12: The kite quadrangle, ABCD, alongside the line segment SR connecting the midpoint of diagonals BD and AC, which lies incident with diagonal BD.

Given a quadrangle with no two sides equal in length, as well as having no angles equal, two cases can arise. Firstly, an asymmetric trapezoid may be constructed in the event that any two opposite sides of the quadrangle are constrained to be parallel. Upon connecting the midpoints of the diagonals of this trapezoid, seen in Figure 2.13, the line segment SR arises. Line segment SR represents the locus of the centres [8] for all inscribing ellipses contained within the trapezoid, and is constructed through connecting the midpoints of lines AD and BC, the diagonals of the quadrangle. Note that here, the line segment SR lies parallel to lines AB and CD, the parallel pair of sides in the quadrangle. Moreover, the mid points of parallel edges AB, CD, and

line segment SR are all collinear, with a finite point of intersection between the non-parallel edges.

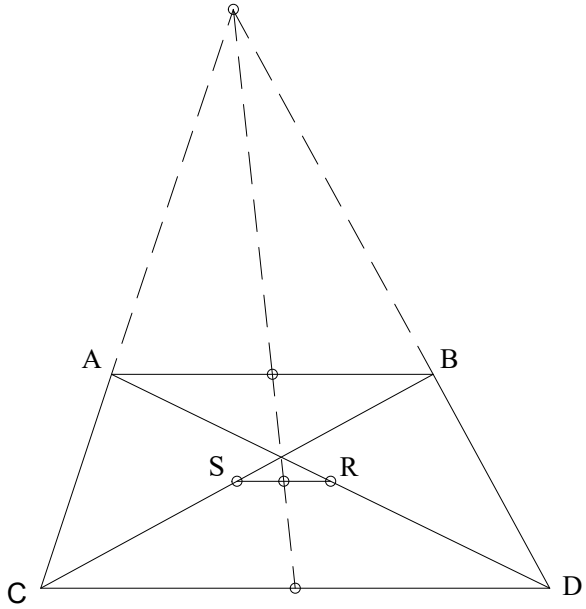


Figure 2.13: Trapezoid ABCD with its diagonals BC and AD, alongside the locus of centres of inscribing ellipse, line segment SR, connecting the midpoints of BC and AD, respectively.

More generally still, is the case presented within Figure 2.14; no sides lie parallel to each other, and no lengths or angles are equal. Line segment SR still represents the locus of centres for all inscribing ellipses within the quadrangle, however it is no longer parallel to any of the sides, nor is it oriented in any immediately evident fashion relative to the geometry of the rest of the quadrangle.

While the various parallelogram cases represent archetypes through which the derivation of the maximum area inscribing ellipse can be somewhat simplified, their simplicity forces the methods used to determine their maximum area inscribing ellipses to be decidedly unsuitable as a general solution. Given this fact, the asymmetric quadrangle will be used for the remainder of this procedure, in order to maintain generality of the solution.

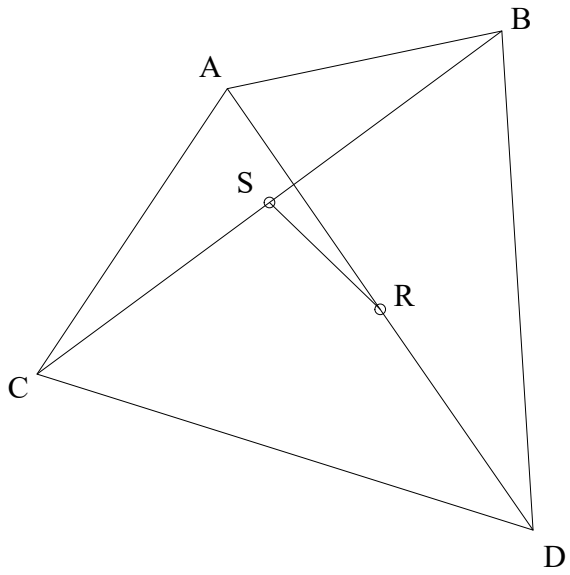


Figure 2.14: An example of a general asymmetric quadrangle ABCD, with line segment SR containing the locus of centres of all inscribing ellipses, connecting the midpoints of diagonals BC and AD, respectively.

2.4 Second Order Curves

Second order curves, including ellipses, can be described through equations with no more than a second order term contained within them. These second order variable curves are, therefore, all a part of a family of functions known as conic sections [23]. Conic sections include ellipses, parabolas, hyperbolas, as well as degenerate conics. Several classification techniques exist within both the algebraic formulations and the quadratic expressions of these curves [1]. Although inherently different in their depiction, these classification techniques are interchangeable in their effectiveness.

Due to the usage of geometric transformations, the quadratic formulation for second order curves will be discussed herein; this is due to their inherent capacity to remain shape invariant under the Euclidean displacement group of geometric transformations, and therefore their utility in describing conic sections throughout a wide range of spaces. Through this generalised representation, all conic sections: hyperbolas; parabolas; ellipses; as well as degenerate conic sections are simply different affine

projections of the same second order projective curve [23].

2.4.1 Conic Section Geometric Representation

Conic sections are, by definition, the result the intersection of a complete cone constructed in three dimensions, with a plane. While the exact cone is variable, in general, a cone, in its canonical form, can be described through the following equation,

$$z^2 = x^2 + y^2. \tag{2.8}$$

This equation specifically represents a cone whose apex is present at the origin of the coordinate system, and is aligned in a right circular fashion, such that the axis of the cone contains the origin and lies incident with the z axis [23]. This position is known as the canonical position; for ellipses and conic sections, their canonical position is the position in which the semi-major axis is aligned with the x coordinate axis, the semi-minor axis with the y coordinate axis, and its centre at the origin of the coordinate system [23]. Figure 2.15 shows what one such complete cone might look like in its canonical form.

Now, from this three dimensional complete cone, two dimensional planar representations of second order curves can be derived. Specifically, if a plane is used to cut this complete cone, said cone will intersect with the plane in a second order curve that is either an ellipse, a parabola, a hyperbola, or in degenerate cases, pairs of straight lines. Figure 2.16 shows a planar intersection that results in the development of an ellipse on the intersecting plane. This plane intersects the complete cone above (or below) its apex, and at an angle other than 90 degrees to the central axis of the cone.

Within the context of this example, circles result as a special case of the ellipse; namely, that the plane being used to cut the complete cone intersects the cone at a position which is removed from its apex, but also intersects the main axis of the cone

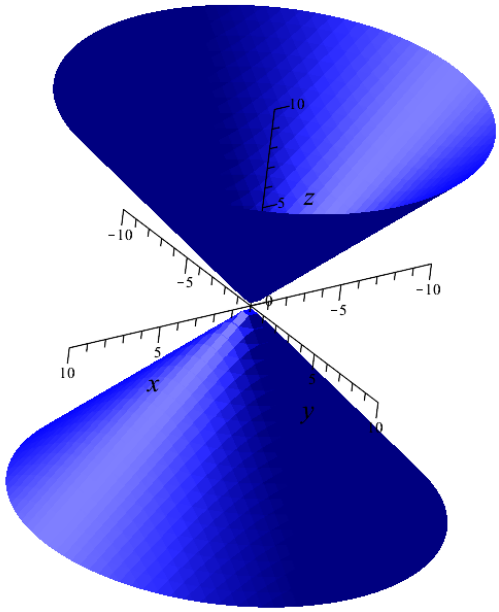


Figure 2.15: A graphical representation of a complete cone in three dimensions, in its canonical position.

at a right angle. Figure 2.17 shows this intersection on a complete cone.

Now, the remaining second order curves are developed through an increasingly extreme rotation of the cutting plane away from perpendicular to the axis of the complete cone. First, a parabola is generated by using an angle of cutting plane that forces the plane to lie parallel to the side of the cone. Due to the parallelism between the cone and its intersecting plane, the resulting curve remains open, and only generates a curve based on one side of the cone. Figure 2.18 shows how this plane intersects the complete cone, in three dimensions.

Upon rotation beyond parallel with the side of the complete cone, a hyperbola will be generated on the cutting plane. Figure 2.19 shows how this curve falls onto the cutting plane from the initial complete cone.

In addition to the aforementioned curves generated through planar intersections

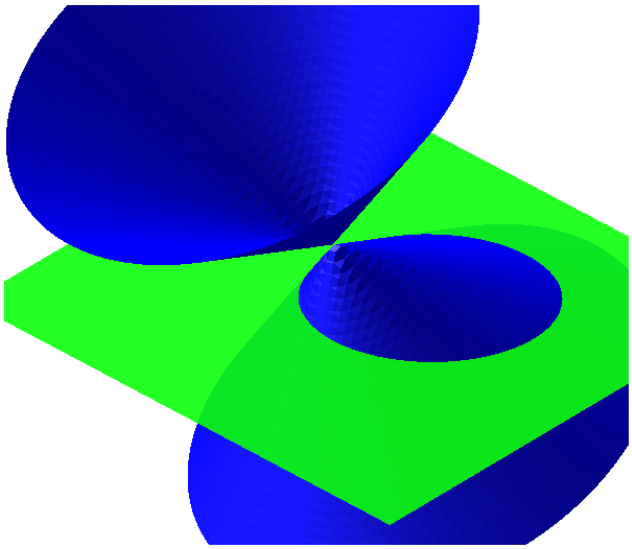


Figure 2.16: A graphical depiction of the intersection of a plane, light, with a complete cone, dark, generating a elliptical conic section.

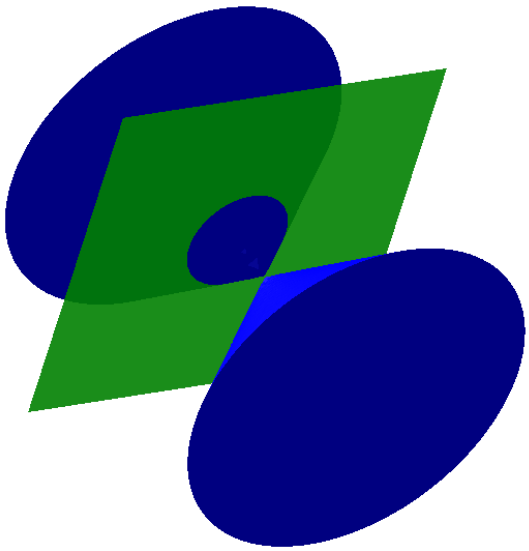


Figure 2.17: A graphical depiction of the intersection of a plane (light) with a complete cone (dark) generating a circular conic section.

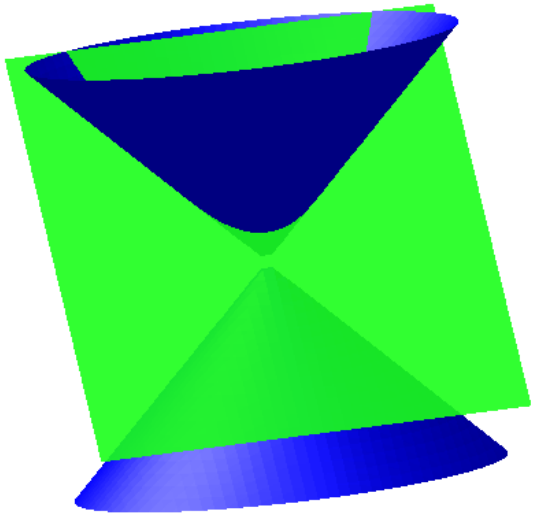


Figure 2.18: A graphical depiction of the intersection of a plane (light) with a complete cone (dark) generating a parabolic conic section.

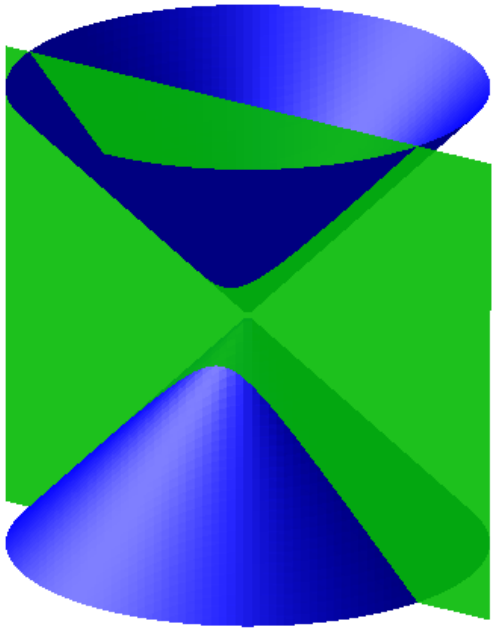


Figure 2.19: A graphical depiction of the intersection of a plane (light) with a complete cone (dark) generating a parabolic conic section.

with the complete two additional degenerate conic sections may be generated in special cases of planar orientations. First, given a planar orientation which is parallel to the edge of the cone while also lying incident with it, a doubly mapped real line is generated. This orientation is depicted in Figure 2.20.

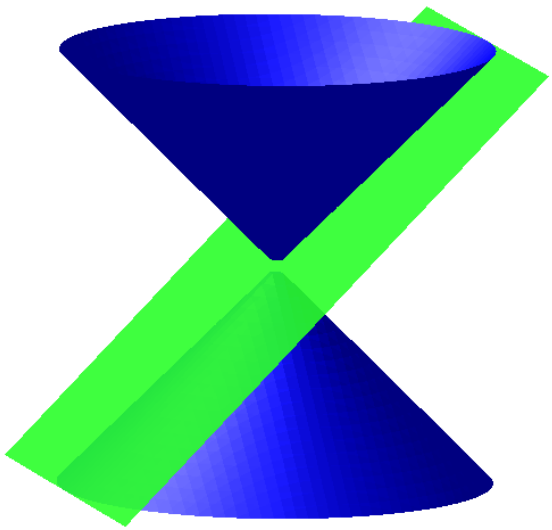


Figure 2.20: A graphical representation of a plane (light) with a complete cone (dark) resulting in the generation of a doubly mapped real line.

Secondly, given a plane parallel and incident to the axis of the cone, a pair of real lines will be mapped onto the intersecting plane, shown in Figure 2.21.

Assuming that the complete cone is being viewed in such a way so as to ensure that the cutting planes are all running normal to the viewing plane, Figure 2.22 arises. From this perspective, it is trivial to see how each of the three unique conic section types can be defined from a single complete cone.

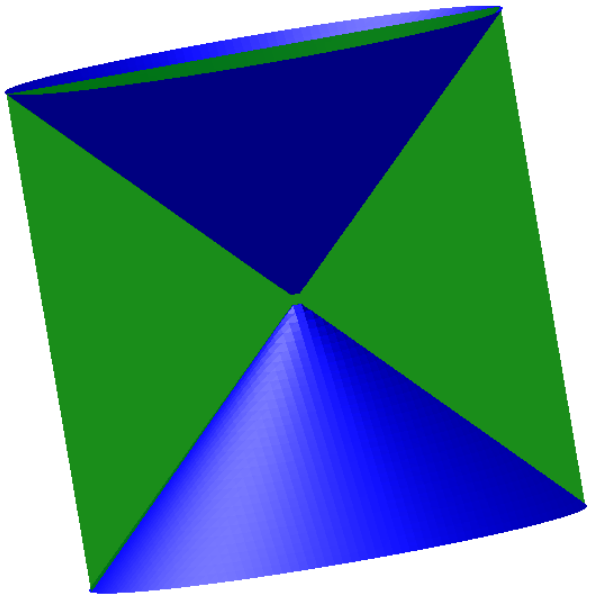


Figure 2.21: A graphical representation of a plane (light) with a complete cone (dark) resulting in the generation of a pair of real lines.

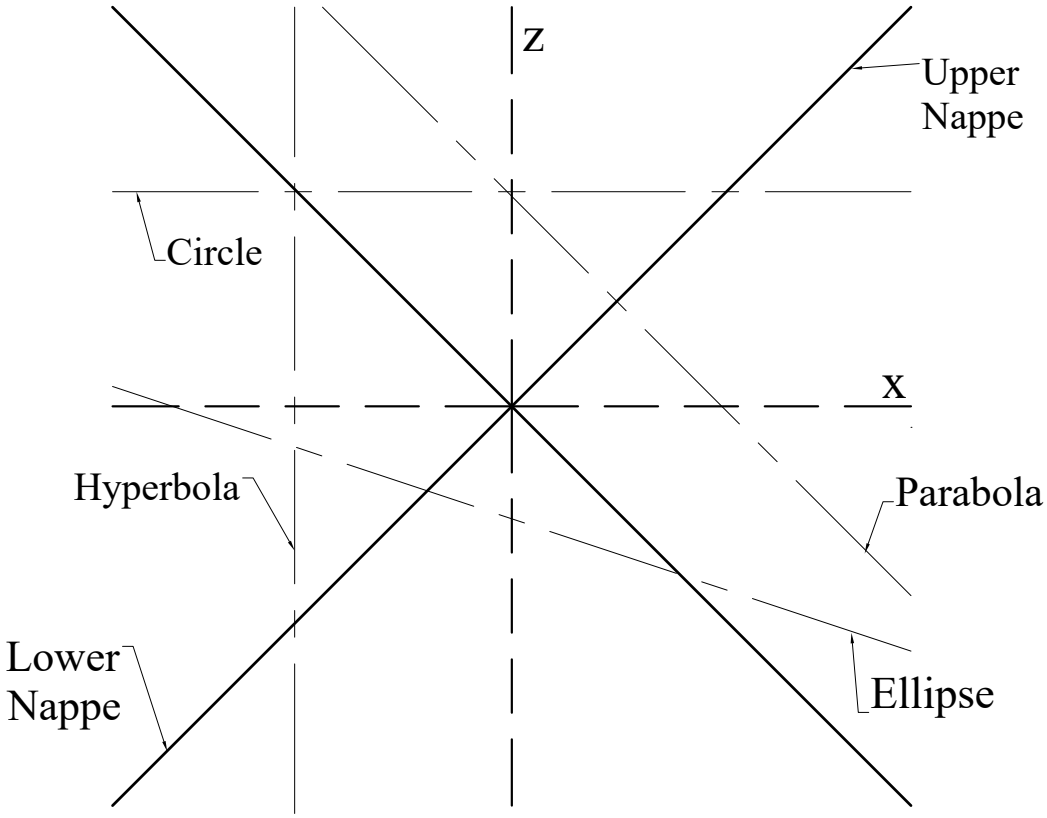


Figure 2.22: Right angled view showcasing how each planar conic section is created through differing orientations of the plane intersecting the cone in three dimensions.

Figure 2.22 shows a complete cone marked by solid lines, as a projection onto the $x - z$ plane, shown by centre dashed lines. Each possible planar position used to generate every type of non-degenerate conic section is shown by dotted lines. The solid lines themselves represent the degenerate conic sections as pairs of distinct lines. The ellipse plane intersects the axis of the cone at an angle other than 90 degrees relative to the z axis, where the circle's plane intersects the axis in a right angle. The parabolic conic section is generated when the intersecting plane lies parallel to the side of the cone, whereas the hyperbolic conic section is generated with the intersecting plane lies parallel to, but displaced from, the axis of the cone itself.

2.4.2 Conic Section Algebraic Characterisation

While the geometric definitions of conic sections are useful for demonstrating that a given parametric pencil of conic sections may, at any point, be an ellipse, a parabola, or a hyperbola, an exact mathematical definition must be given in order to facilitate their computation. Typically, second order functions are represented through defined equations such as those of the parabola, circle, and hyperbola; unfortunately, these representations are less efficient at a computational classification approach. Thus, in general, conic sections will be represented in their matrix-vector form,

$$k := \mathbf{x}^T \mathbf{A} \mathbf{x} = 0, \tag{2.9}$$

where $\mathbf{x}$ is a given homogeneous point triple, and $\mathbf{A}$ is the conic shape coefficient matrix that is symmetric and positive definite given non-degenerate conics.

Line and Point Coordinates

Given the duality of lines and points within the plane, any point or line can be derived as either a variable line passing through a fixed point, or as a variable point on a fixed

line. Specifically, the equation used to represent any line or point coordinate,

$$x_0X_0 + x_1X_1 + x_2X_2 = 0, \tag{2.10}$$

can be described either in terms of the variable lines passing through the fixed point (x_0, x_1, x_2), or as the variable points lying on the fixed line $[X_0, X_1, X_2]$. In order to provide clarity of expression, line coordinates are delimited by brackets, i.e, $[X_0, X_1, X_2]$, while point coordinates are delimited by parentheses, i.e, (x_0, x_1, x_2). Given a point (a, b, c), any line which satisfies,

$$aX_0 + bX_1 + cX_2 = 0, \tag{2.11}$$

is a line that passes through the fixed point (a, b, c). Conversely, given a line $[A, B, C]$, any point which satisfies,

$$x_0A + x_1B + x_2C = 0, \tag{2.12}$$

is a point that lies on the line (A, B, C). Given the planar duality of lines and points, the conversion of line coordinates to point coordinates and point coordinates to their respective line coordinates is easily undertaken. Given two coplanar and distinct points, (x_0, x_1, x_2) and (y_0, y_1, y_2), the line coordinate of the line passing through both of these points can be expressed through Grassmanian expansion of,

$$\begin{vmatrix} X_0 & X_1 & X_2 \\ x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \end{vmatrix} = 0, \tag{2.13}$$

where the cofactor corresponding to $[X_0, X_1, X_2]$ is the value of X_i within the line

equation. In other words, any additional point, (b_0, b_1, b_2) on the line connecting points (x_0, x_1, x_2) and (y_0, y_1, y_2), must satisfy the equation,

$$b_0(x_1y_2 - x_2y_1) - b_1(x_0y_2 - x_2y_0) + b_2(x_0y_1 - x_1y_0) = 0. \tag{2.14}$$

This exact procedure may be dualised in order to describe the unique point at which two lines intersect, and from this equation, provide the point equation of the family of lines passing through the fixed point defined by the initial two lines.

Line and Point Conics

Although computationally efficient, the matrix-vector formulation of conic sections, seen in Equation (2.9), is capable of presenting the conic sections in both their line and point conic forms. While these two forms represent the exact same curve, they each have distinct advantages and disadvantages in terms of their computational capacities, as well as their visual representations. In keeping with the case at hand, the discussion of the differentiation of pole points and polar lines, the elements upon which point and line conics themselves are constructed, respectively, will be limited to a discussion of the pole points and polar lines lying on the curve k, in Equation (2.9).

Each of the line and point conics possess a conic shape coefficient matrix, $\mathbf{A}$, which, when using homogeneous coordinate triples in a plane, are comprised of at most six independent variables, A_{ij}. Equation (2.15) shows a line conic shape coefficient matrix, $\mathbf{A}_L$, which is always a square symmetric matrix,

$$\mathbf{A}_L = \begin{bmatrix} A_{00} & A_{01} & A_{02} \\ A_{01} & A_{11} & A_{12} \\ A_{02} & A_{12} & A_{22} \end{bmatrix}, \tag{2.15}$$

where each element, A_{ij}, is a line conic shape coefficient. Although it is typical within engineering and design to describe functions through their point equations, Equation (2.15) will represent the curve k in terms of the family of lines lying tangent to the curve at any given point. Substituting values into this expression will supply you with the equation of the line coordinates of the line that lies tangent to the curve at the point in question, illustrated in Figure 2.23, with a small subset of the resulting polar lines plotted along the complete curve, k.

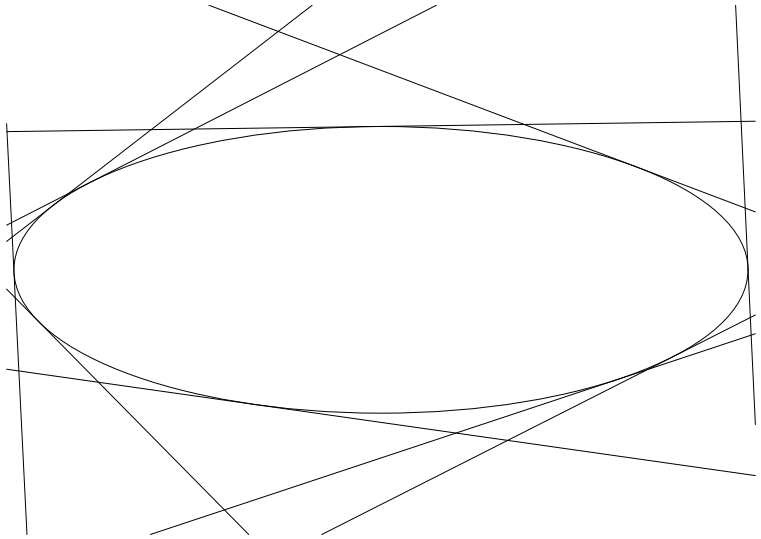


Figure 2.23: A graphical depiction of what a line conic shape equation produces, with a small subset of the polar lines belonging the second order curve, k.

In general, the complete curve k would only be displayed by way of its polar lines, thus eliminating the closed curve depicted in 2.23, leaving only the polar lines in place. Ultimately, in order to provide a high enough resolution for displaying this curve, the line conic must be displayed along a small enough interval to provide the resolution, resulting in a cumbersome and unclear display consisting of an infinite family of unbounded lines enclosing the desired function. Upon pre and post multiplication with the three element triple of the line coordinates, X, the curve k is defined as,

$$k := A_{00}X_0^2 + 2A_{01}X_0X_1 + 2A_{02}X_0X_2 + A_{11}X_1^2 + 2A_{12}X_1X_2 + A_{22}X_2^2 = 0, \quad (2.16)$$

where k is a second order function in terms of the line coordinates, X_i, and linearly dependent upon the line conic shape coefficients, A_{ij}. However, for display purposes, it is important to obtain the point conic representation of the second order curve, k. The point conic shape matrix, $\mathbf{A}$, is the dual of the line conic shape coefficient matrix, $\mathbf{A}_L$,

$$\mathbf{A} = \begin{bmatrix} a_{00} & a_{01} & a_{02} \\ a_{01} & a_{11} & a_{12} \\ a_{02} & a_{12} & a_{22} \end{bmatrix}, \quad (2.17)$$

whose point conic shape coefficients are represented by each of the six point conic shape coefficients, a_{ij}. Figure 2.24 shows what curve k produces when pre and post multiplied by the homogeneous point triple, $\mathbf{x}$.

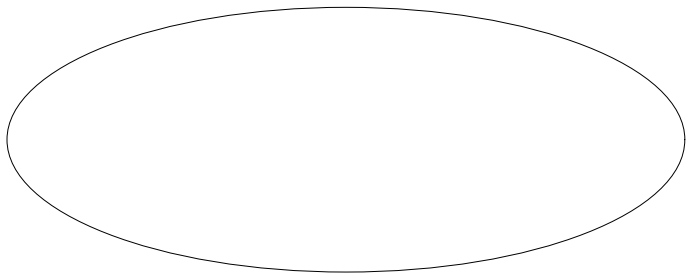


Figure 2.24: A graphical depiction of what a point conic shape coefficient matrix produces, the second order curve, k.

The point form of the conic as defined by the shape coefficients in $\mathbf{A}$ is k:

$$k := a_{00}x_0^2 + 2a_{01}x_0x_1 + 2a_{02}x_0x_2 + a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = 0, \tag{2.18}$$

which is also linearly dependent upon the point conic shape coefficients, a_{ij}, and homogeneously quadratic in the point coordinates, x_i, in an identical manner to the line conic shape coefficients and line coordinates in Equation (2.16).

Although line conic matrix-vector equations can be more easily defined and parameterised given a set of linear constraints, displaying a line conic equation is less effective for describing the curve it represents than displaying its point conic counterpart. Conveniently, duality within the plane supplies a method whereby the point conic equation may be computed directly from its corresponding line conic shape coefficient matrix.

Specifically, the point conic shape coefficient matrix is the inverse of the line conic shape coefficient matrix. Typically, given a matrix of line conic shape coefficients, $\mathbf{A}_L$, the point conic shape coefficient matrix, $\mathbf{A}$ is expressed as,

$$\mathbf{A} = \mathbf{A}_L^{-1} = \frac{\mathrm{adj}\mathbf{A}_\mathrm{L}}{\det \mathbf{A}_L}, \tag{2.19}$$

where the adjoint of $\mathbf{A}_L$, is the matrix of cofactors, or the (ij) position subdeterminants of $\mathbf{A}_L$ with the i and jth column of $\mathbf{A}_L$ excluded. While Equation (2.19) is exact, it can be further reduced in complexity through the elimination of the determinant of $\mathbf{A}_L$. Due to the use of the homogeneous coordinates in the expression of both line and point conics, any scalar multiple of two points can be represented by the same line or point coordinate ratio, $\mathbf{x}$. This implies that the constant value of the determinant of $\mathbf{A}_L$ is a scalar that can be factored out since $\det \mathbf{A}_L \neq 0$ for non-degenerate conic sections, and thus the matrix of point conic shape coefficients

can be expressed as,

$$\mathbf{A} \propto \text{adj}\mathbf{A}_\text{L}, \tag{2.20}$$

proportional to the line conic shape coefficient matrix, $\mathbf{A}_L$. Importantly, nothing is lost in this conversion, so the matrix of point conic shape coefficients may be obtained from the transpose of the matrix of cofactors of $\mathbf{A}_L$. However, since the matrix $\mathbf{A}_L$ is symmetric, its matrix of cofactors is also symmetric, and is, hence, equal to its transpose.

Point Conic Algebraic Characterisation

While line and point conic shape coefficients may be computed through the exact same approach, line conics are not useful for visualisation, and therefore the remainder of this characterisation will deal with point conics; the rules presented herein may be extended to their dual line conics through the principles of duality presented in Section 2.1.2. The symmetric point conic shape coefficient matrix, $\mathbf{A}$ (Equation (2.17)), possesses a 2×2 sub determinant, the quadratic form of the conic section, resulting from elimination of the first row and column of $\mathbf{A}$, named $\mathbf{A}_0$,

$$\mathbf{A}_0 = \begin{bmatrix} a_{11} & a_{12} \\ \\ a_{12} & a_{22} \end{bmatrix}, \tag{2.21}$$

which is used for characterisation and evaluation of the conic section represented by $\mathbf{A}$. This equation generates a set of Euclidean invariants [1, 24] which provide the classification. Using these Euclidean invariants allows for the development of a set of rules which facilitates the classification of any given conic section, regardless of whether or not there is a displacement from the canonical position.

Specifically, the Euclidean invariants which are useful for this problem are [1],

$$\Delta := \det \mathbf{A}; \tag{2.22}$$

$$\Delta_0 := \det \mathbf{A_0} = a_{11}a_{22} - a_{12}^2; \tag{2.23}$$

$$H := a_{11} + a_{22}, \tag{2.24}$$

where Δ is the determinant of the conic shape coefficient matrix, while Δ_0 and H are the discriminant of trace of the quadratic form of the conic, respectively. These quantities will remain invariant under any Euclidean transformation applied to the conic in question, or if the original Cartesian coordinate system is transformed. Special care must be taken in the event that $\Delta = \Delta_0 = 0$, arising in the event of the creation of a special subset of degenerate conic sections. This special case gives rise to yet another Euclidean invariant which facilitates the classification of these degenerate conic sections [1],

$$K := a_{00}H - (a_{01}^2 + a_{02}^2). \tag{2.25}$$

While the invariants themselves are not necessarily useful for determining the physical attributes of these conic sections, the signs of the invariants are important. Specifically, Δ and Δ_0 determine the physical shape characteristics of the conic sections, with H and, in some specific instances, K, providing additional information about the nature of the sections. These values and their relative magnitudes are listed in Table 2.1.

While the Euclidean invariants provide valuable information which facilitates a complete classification of all conic sections, the sections which will be of primary interest for the remainder of this classification are the sections where $\Delta \neq 0$.

Table 2.1: Conic section classes and their corresponding Euclidean invariants [1].

	$\Delta \neq 0$**:**	$\Delta = 0$**:**
	Regular conic sections	***Singular conic sections***
$\Delta_0 \neq 0$**: Second order curve with a real center**	$\Delta_0 > 0$ and $H \times \Delta > 0$, section is a nullpartite ellipse	$\Delta_0 > 0$: Section is a pair of conjugate complex lines with a proper intersection point
	$\Delta_0 > 0$ and $H \times \Delta < 0$, section is a unipartite ellipse	
	$\Delta_0 < 0$: section is a hyperbola	$\Delta_0 < 0$: Section is a pair of real lines with a proper intersection point
$\Delta_0 = 0$**: Second order curve with one axis of symmetry**	Section is a parabola	$K > 0$: Pair of conjugate complex parallel lines
		$K = 0$: Doubly mapped real line
		$K < 0$: Real parallel lines

These conic sections exist, with exception of the case of $\Delta_0 > 0$ with $H \times \Delta > 0$, as second order curves within the real plane. If $\Delta_0 > 0$ and $H \times \Delta > 0$, the conic section generated by this point conic shape coefficient matrix is a nullpartite ellipse [1] [24]; this ellipse lies purely in the imaginary plane, possessing no real points.

Given that $\Delta_0 \neq 0$ defines conic sections with a real center (i.e. central conics), $\mathbf{m}$, this case bears additional investigation. First, the homogeneous point coordinates of the center of this conic section may be computed as the solution to the two partial derivatives of k in x_1 and x_2,

$$\frac{\partial k}{\partial x_1} = 2a_{01}x_0 + 2a_{11}x_1 + 2a_{12}x_2 = 0, \tag{2.26}$$

$$\frac{\partial k}{\partial x_2} = 2a_{02}x_0 + 2a_{12}x_1 + 2a_{22}x_2 = 0, \tag{2.27}$$

yielding,

$$\mathbf{m} = \begin{bmatrix} m_0 \\ m_1 \\ m_2 \end{bmatrix} = \begin{bmatrix} a_{11}a_{22} - a_{12}^2 \\ a_{12}a_{02} - a_{22}a_{01} \\ a_{01}a_{12} - a_{02}a_{11} \end{bmatrix}. \tag{2.28}$$

From Equation (2.21), there exist two distinct, non-zero, real eigenvalues, λ_1 and λ_2, with corresponding mutually perpendicular eigenvectors, representing the major and minor axis directions of the conic $\mathbf{A}$. Due to the symmetric nature of $\mathbf{A}$, there will always be two distinct eigenvalues and therefore eigenvectors. While solutions for the eigenvalues of this matrix are possible through solution of the characteristic equation [25], they may also be expressed in terms of the Euclidean invariants used to characterize the conic section $\mathbf{A}$ itself,

$$\lambda^2 - H\lambda + \Delta_0 = 0. \tag{2.29}$$

Equation (2.29) shows this construction for the characteristic equation in terms of the Euclidean invariants mentioned above. Therefore, by means of Vieta's formulae [26], the solutions to this characteristic equation may be described by

$$\lambda_1 + \lambda_2 = H, \tag{2.30}$$

$$\lambda_1\lambda_2 = \Delta_0. \tag{2.31}$$

From this description of the eignevalues for the aforementioned conic section, it is possible to provide additional specificity regarding the classifications found within Table 2.1. While each case can be expounded upon, only real conic sections and a

special case of the singular conics, particularly relevant to the problem at hand, will be described herein.

- Case **1**, $\Delta \neq 0$, **A** describes a regular second order curve.
  - Case **1.1**, $\Delta_0 < 0$, from Equation 2.31, λ_1 and λ_2 have different signs, indicating that **A** describes a hyperbola.
  - Case **1.2**, $\Delta_0 > 0$, from Equation 2.31, λ_1 and λ_2 have identical signs, and **A** describes a unipartite ellipse, lying in the real plane.
- Case **2** $\Delta_0 = 0$, second order curves with exactly one line of symmetry, occurs when one of λ_1 or λ_2 is exactly equal to zero.
  - Case **2.1**, $\Delta \neq 0$, occurs when Δ and H have different signs, and **A** is therefore a description of a parabola.
  - Case **2.2**, $\Delta = 0$, and $K = 0$, then the section defined by **A** describes a pair of doubly mapped real lines.

Within Cases 1.1 and 1.2, the eigenvalues, λ_1 and λ_2, of the characteristic equation, can be used to generate the semi major and semi minor axis lengths of **A**, α_1 and α_2, [1]:

$$\alpha_i = \sqrt{\left|\frac{\Delta}{\Delta_0 \lambda_i}\right|}. \tag{2.32}$$

These axes lie parallel to the corresponding eigenvectors of the matrix $\mathbf{A}_0$. In the event that the two eigenvalues are identical in magnitude and sign, a circle is defined by **A**. This special case elliptical conic section occurs not only when both eigenvalues agree in magnitude and direction, but also in the case that a_{01}, a_{02} and a_{12} are all identically zero.

Within Cases 2.1 and 2.2, however, only one non-zero eigenvalue can be obtained. Therefore, one of λ_i is precisely equal to zero, and thus these two special cases possess only one non-zero eigenvector, which in turn defines their axis of symmetry. This axis of symmetry can be defined by its homogeneous equation, m,

$$m := (a_{12}a_{01} + a_{22}a_{02})x_0 + a_{12}(a_{11} + a_{22})x_1 + (a_{12}^2 + a_{22}^2)x_2 = 0. \tag{2.33}$$

In Case 2.1, this homogeneous equation represents the axis of symmetry of the parabola defined by $\mathbf{A}$, and within Case 2.2, this equation directly represents the equation of the doubly mapped lines. While Equation (2.33) defines the axis of symmetry for both line conics within Cases 2.1 and 2.2, the latter case always generates an unbounded line, and the only line of symmetry for an infinite line is that line itself.

2.5 Area of a Conic

In order to effectively define an algorithm that is capable of providing the maximum area inscribing ellipse within a convex quadrangle, it is necessary to develop a method whereby the area of any given conic section may be computed. While the area of an ellipse, particularly relevant to the problem at hand, may be calculated by,

$$A = \pi ab, \tag{2.34}$$

this equation requires the lengths of both the semi-major and semi-minor axes, which for an ellipse not in canonical position, are not immediately evident. Instead, given the conic shape coefficient matrix which describes an ellipse (or a parameterised pencil of conics) the following equation is proposed [2] [1],

$$Area(\mathbf{A}) = \left| \frac{\pi \det \mathbf{A}}{(\sqrt{\Delta_0})^3} \right|, \tag{2.35}$$

where,

$$\mathbf{A} = \begin{bmatrix} a_{00} & a_{01} & a_{02} \\ a_{01} & a_{11} & a_{12} \\ a_{02} & a_{12} & a_{22} \end{bmatrix}$$

and, Δ_0 is the 1-1 cofactor of the quadratic form of $\mathbf{A}$,

$$\Delta_0 = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}.$$

The following two subsections of this chapter describe the derivation of a special case pencil of conics, a central conic in its canonical position, as well as a proof for Equation (2.35), which always computes the area of the conic section if it is a closed, central conic.

Ellipse In Canonical Position

While the position and orientation of an ellipse is completely arbitrary, it is sometimes valuable to express the ellipse itself in its canonical position. For ellipses, the canonical position is that in which the centre of the ellipse is incident with the origin, while each semi-major and semi-minor axes of the ellipse are incident with the coordinate axes. While it is possible to transform any ellipse into canonical position, performing such a transformation requires the directions of the semi-major and semi-minor axes of the ellipse. These axis directions are not, however, immediately evident given the point (or line) conic shape coefficient matrices.

However, although the orientation of the ellipse is not immediately evident upon inspection of the matrix-vector form, recall that matrix $\mathbf{A}$ is a square symmetric matrix, such that $\mathbf{A} = \mathbf{A}^T$, indicating that it is always diagonalisable [25]. This diagonalisation of $\mathbf{A}$ positions the conic section, k, in its canonical position, and is accomplished by computing the eigenvalues of $\mathbf{A}$, and aligning them on the diagonal of the matrix.

A square matrix $\mathbf{A}$ is said to be diagonalisable if there is a matrix $\mathbf{P}$ such that, $\mathbf{P}^{-1}\mathbf{AP}$ is diagonal. Matrix $\mathbf{P}$ is said to diagonalise $\mathbf{A}$. It can be shown [25] that the columns of matrix $\mathbf{P}$ are the eigenvectors of $\mathbf{A}$, and the resulting diagonal matrix $\mathbf{D}$ contains the corresponding eigenvalues on its diagonal,

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}. \tag{2.36}$$

2.5.1 Derivation of an Origin Centred Pencil of Ellipses

Given the computational complexity of Equation (2.35), it is necessary to use a special case pencil of ellipses in order to prove this area function applies to every non degenerate ellipse. Specifically, a pencil of conics whose centre lies coincident with the origin of the coordinate system, with semi-major and semi-minor axes incident with the axes of the coordinate system.

Given a coordinate system defined by the axes, x and y, with homogeneous coordinate components x_0,x_1, and x_2, the general quadratic expression for a conic section can be written as Equation (2.9), reproduced here for convenience.

$$k := a_{00}x_0^2 + 2a_{01}x_0x_1 + 2a_{02}x_0x_2 + a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = 0. \tag{2.37}$$

Another representation of an ellipse is often described as,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0. \tag{2.38}$$

These two equations are bound to represent the same curve when the Cartesian coordinates are recovered for a value of $x_0 = 1$, while both have their centres incident with the origin of the coordinate system. Given that Equation (2.37) must simplify to agree with Equation (2.38), it can be inferred that the point conic shape coefficients a_{01}, a_{02} and a_{12} must all be equal to zero, while taking a projectivity of $x_0 = 1$ to recover the Cartesian coordinates of the function shows that the conic shape coefficient a_{00} is equal to -1.

Arranging these values in a matrix provides the corresponding diagonalised matrix vector form of an origin centred conic, $\mathbf{A}$,

$$\mathbf{A} = \begin{bmatrix} a_{00} & 0 & 0 \\ 0 & a_{11} & 0 \\ 0 & 0 & a_{22} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{a^2} & 0 \\ 0 & 0 & \frac{1}{b^2} \end{bmatrix}, \tag{2.39}$$

in terms of the Cartesian semi-major and semi-minor axis lengths, a and b. This diagonalisation can also be used to represent the matrix-vector form of the conic section in Equation (2.37) in terms of the Euclidean axis lengths, as in Equation (2.38). Upon substitution of this matrix $\mathbf{A}$ into Equation (2.35) such that:

$$\Delta = \frac{-1}{a^2 b^2}; \tag{2.40}$$

$$\Delta_0 = \frac{1}{a^2 b^2}. \tag{2.41}$$

The area of the conic section defined by the point conic shape coefficient matrix, **A** can be rewritten as,

$$Area(\mathbf{A}) = \pi \left| \frac{\frac{-1}{a^2b^2}}{\frac{1}{a^2b^2}^{\frac{3}{2}}} \right|. \tag{2.42}$$

Applying the laws of exponents, it can be show that Equation (2.42) reduces to,

$$Area(\mathbf{A}) = \pi \frac{ab^{-2}}{ab^{-3}} = \pi ab. \tag{2.43}$$

Considering that any conic shape coefficient matrix **A** corresponding to an ellipse can be transformed into canonical position through diagonalisation, Equation (2.35) will always provide the area of the conic section k, given its conic shape coefficient matrix, **A**.

Chapter 3

Existing Solutions For Generating the Maximum Area Inscribing Ellipse within a Convex Quadrangle

Within this chapter, two previously existing solutions will be presented to the problem of defining the maximum area inscribing ellipse within a convex quadrangle. Though both solutions can, in some circumstances, produce the maximum area inscribing ellipse, they are both inherently limited. The first, found in [10], uses a projective collineation which maps the unit square onto the quadrangle in question; the inverse of this collineation is then used to map the parametric unit circle equation onto the ellipse contained within the quadrangle. The second, found in [2], utilises the geometry of the quadrangle in order to construct a non-orthogonal coordinate system, upon which the parametric equation for the family of ellipses inscribing the quadrangle can be created, and maximised with respect to the area of the ellipse.

3.1 Projective Transformation Method

In order to determine the maximum area inscribing ellipse contained within any given convex quadrangle, it is beneficial to analyze the special case quadrangle of the unit circle, i.e. with a radius of one unit, inscribing the corresponding "unit" square. In Section 2.3 it is evident that the centre for all of the ellipses present within the square lies at the mid point of the diagonals, where they each intersect, and thus in the exact centre of the square. Each pole point of the conic must lie tangent to the mid points of the parallel sides of the quadrangle; seeing as a square consists of two pairs of perpendicular parallel lines, this means that the maximum area inscribing ellipse must lie tangent to the midpoints of every side, and is thus the unit circle. Figure 3.1 shows the square and the inscribing unit circle.

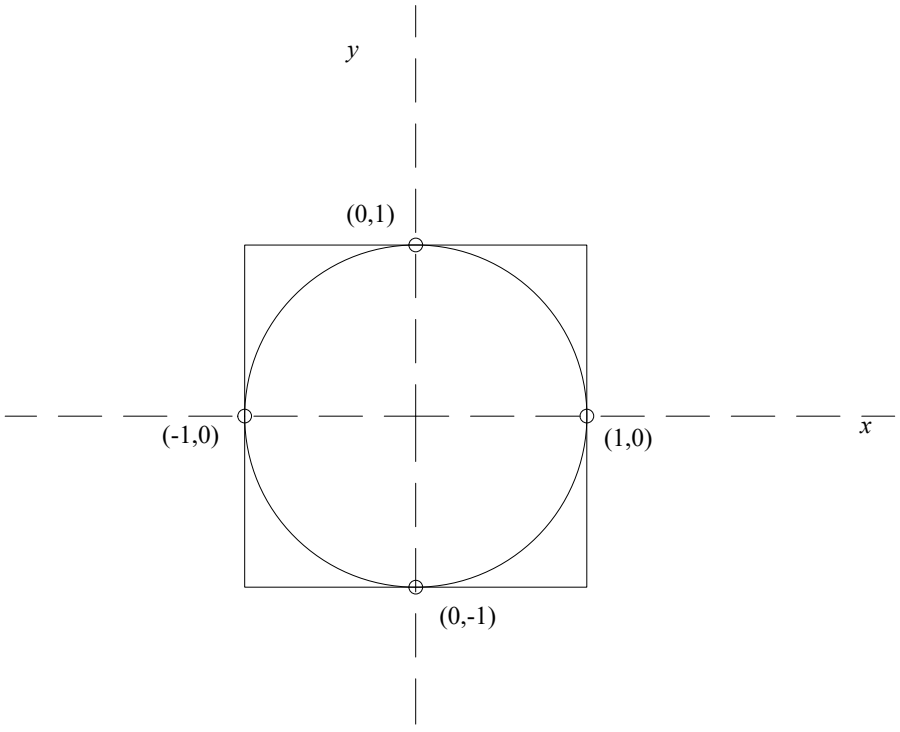


Figure 3.1: The square with its largest area inscribing ellipse, the unit circle, and all four tangent points.

3.1.1 Projective Mapping Generation

Considering that a projective collineation can be used to map points from an existing coordinate system onto different points within the same coordinate system while maintaining the collinear nature of the points, a square may be mapped onto any given quadrangle, at any location within the same space, through the employment of a collineation.

Two distinct sets of four points within the projective plane uniquely determine a projective collineation, provided that no three of the points are collinear. Given a distinct quadrangle, the points representing the vertices of the quadrangle can be used to produce a projective collineation which maps the square onto any general quadrangle.

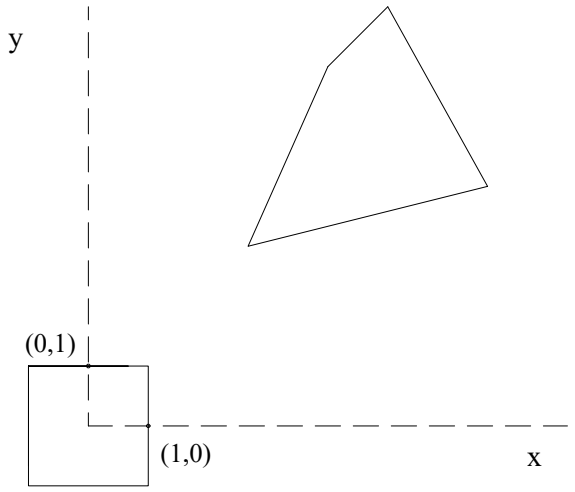


Figure 3.2: The square and the desired, mapped quadrangle.

Consider a set of four points which represent an arbitrary convex quadrangle; $W(W_0 : W_1 : W_2)$, $X(X_0 : X_1 : X_2)$, $Y(Y_0 : Y_1 : Y_2)$, $Z(Z_0 : Z_1 : Z_2)$, represented using their homogenous coordinate point triples. These four points will be mapped onto a square, centred at the origin, represented by the homogeneous coordinate

point triples; $w(w_0 : w_1 : w_2)$, $x(x_0 : x_1 : x_2)$, $y(y_0 : y_1 : y_2)$, $z(z_0 : z_1 : z_2)$.

Specifically, this collineation may be represented through the following vector-algebraic relationship,

$$\lambda \begin{bmatrix} W_0 \\ W_1 \\ W_2 \end{bmatrix} = \mu \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix}, \tag{3.1}$$

where λ and μ represent arbitrary scaling factors; only the ratio of these factors arising from the use of homogeneous coordinates is important. Without any loss in generality, one may use this ratio and set $\rho = \lambda/\mu$, thus:

$$\rho \mathbf{W} = \mathbf{T}\mathbf{w}. \tag{3.2}$$

The elements of matrix $\mathbf{T}$, t_{ij}, are dependent solely upon the points upon which the mapping is being performed. Due to the conditions present in order to construct a general projective collineation, namely that no property aside from the cross ratio between points is preserved under this type of mapping, the matrix $\mathbf{T}$ has no orthogonality conditions present within any of its rows or columns; the t_{ij} can take on any numerical value, based solely on the conditions of the mapping.

This general transformation matrix $\mathbf{T}$ contains nine unknowns. However, as a result of the use of homogeneous coordinates, at most eight of these numbers are independent. In order to completely characterise this mapping, the scaling factor ρ must be taken into account due to the imposition of the Cartesian coordinate system imposed onto the Euclidean plane into which the points are projected;

because this mapping is not taking place directly within the projective plane, merely the projective extension of the Euclidean plane, the scale cannot be neglected. Specifically, this means that the scaling factor for each point pair represented by ρ_i, i.e. 0, 1, 2, 3, must be incorporated. For example, ρ_0 results from mapping W to w, while ρ_3 is obtained through mapping Z to z.

Within the specific case of the mapping required to generate the maximum area inscribing ellipse, points *W, X, Y,* and Z represent the vertices of the general convex quadrangle, where points *w, x, y,* and z represent the vertices of the unit square, and are the images of the vertices of the general quadrangle. Expanding equation 3.2 and including the terms ρ_i for each point and its image, the following series of equations can be obtained:

$$\begin{aligned}
t_{11}w_0 + t_{12}w_1 + t_{13}w_2 - \rho_0 W_0;\\
t_{21}w_0 + t_{22}w_1 + t_{23}w_2 - \rho_0 W_1;\\
t_{31}w_0 + t_{32}w_1 + t_{33}w_2 - \rho_0 W_2;\\
t_{11}x_0 + t_{12}x_1 + t_{13}x_2 - \rho_1 X_0;\\
\vdots\\
t_{31}z_0 + t_{32}w_1 + t_{33}z_2 - \rho_3 Z_2.
\end{aligned} \tag{3.3}$$

Equations (3.3), when fully defined, represent 12 equations comprised of 13 unknowns; the elements t_{ij}, which define the transformation itself, as well as the four scaling factors, ρ_i. Due to the fact that only eight of the nine elements contained within matrix $\mathbf{T}$ are independent, its elements can be normalised by dividing $\mathbf{T}$ by t_{11} provided that $t_{11} \neq 0$, thus simplifying the system of equations to twelve equations

with an equal number of unknowns.

3.1.2 Mapping the Unit Circle

Given the matrix, $\mathbf{T}$, determined above, it is possible to transform the maximum area inscribing ellipse of what we call the unit square, the unit circle, to the maximum area inscribing ellipse of the general quadrangle. While it is simple to think of the problem in terms of mapping the unit circle to the maximum area inscribing ellipse of the quadrangle, the projective collineation determined above was posed with the idea that the unit square was the image of the vertices of the quadrangle, and therefore the inverse transformation of $\mathbf{T}$ must be used in order to effectively map the unit circle to the maximum area inscribing ellipse.

Given the parametric equation for the circle,

$$c = \begin{bmatrix} 1 \\ \cos(\theta) \\ \sin(\theta) \end{bmatrix}, \tag{3.4}$$

the inscribing ellipse, $\mathbf{e}$, may be computed through pre-multiplication with the inverse of the transformation matrix, $\mathbf{T}$,

$$\mathbf{e} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ \cos(\theta) \\ \sin(\theta) \end{bmatrix}. \tag{3.5}$$

3.1.3 Applicability and Suitability of the Solution

In general, Equation (3.5) represents a projective mapping of the unit circle inscribing the "unit" square onto the ellipse inscribing the convex quadrangle. However, this image does not necessarily represent the largest area inscribing ellipse within the quadrangle. Ideally, a transformation of this sort will preserve the relative size of the objects; area ratios would remain the same, and therefore the image of the unit circle should be the largest area inscribing ellipse within the quadrangle. However, this is not generally the case.

During the mapping to a general convex quadrangle, it is possible for the transformation to satisfy the conditions of more specific mappings on more tightly conditioned spaces. Within a general projective transformation, the only geometric property which is preserved is the cross ratio between four distinct collinear points [15], whereas in an affine transformation, parallelism is preserved (through preservation of the line which lies at infinity), alongside the ratios of the sizes of an object [15]. Figure 3.3 shows a quadrangle, ABCD, used to continue this discussion.

The quadrangle ABCD is, in this case, a trapezoid. The open line segment, TS, between lines AC and DB is created by joining the midpoints of each of these line segments, and represents the locus of centres of the all inscribing ellipses contained within the quadrangle with the diagonals bounding the pencil. However, more importantly, the parallel edges AB and DC are also marked with points E and F, respectively. These points represent the pole points of the maximum area inscribing ellipse within the trapezoid, and also coincide with the midpoints of line segments AB and DC, respectively. The line segment EF also passes through the centre of the line segment TS.

From Figure 3.3, it is clear that three properties of what we call the unit square have been preserved through the transformation into the trapezoid; the polar points

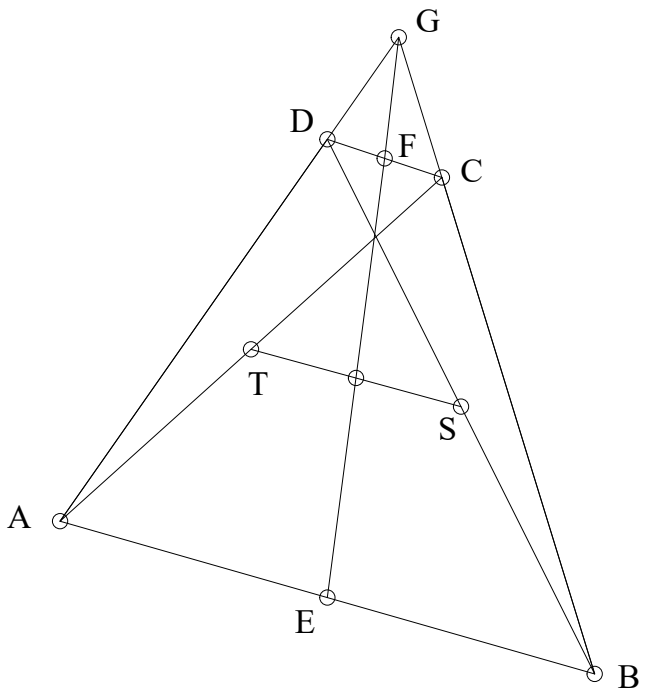


Figure 3.3: A trapezoid ABCD, and its relevant properties.

belonging to the parallel sides of the quadrangle are incident with the midpoints of those lines, and the centre of the maximum area inscribing ellipse lies equidistant from points T and S, or is directly in the centre of the line segment. Furthermore, the parallelism between sides AB and DC is also preserved. Specifically, the projective collineation used to map the trapezoid onto the unit square within this problem must preserve the cross ratio between points; four points on a line under a projective transformation have a cross ratio of exactly -1, when one of these points is taken as an ideal point, and are referred to as a harmonic sequence of points [21]. Consider four distinct and collinear points (A, B, C, D), the cross ratio of the points $CR(\text{AB};\text{CD}) = -1$ if D lies on the line at infinity.

Specifically, consider two adjacent vertices of a square, A and B, and the midpoint between their vertices, C, while point D is the ideal point of the line supporting the edge AB, located on the line at infinity with homogeneous coordinates of $(0 : x_1)$. Suppose then that the distance on the line between points A and B is p. Table 3.1

shows the coordinates of these four points.

Table 3.1: Coordinates used to showcase the special case cross ratio value for a harmonic sequence of points.

Point	x_0	x_1
A	1	0
B	1	$\frac{p}{2}$
C	1	p
D	0	x_1

Given the standard definition of the cross ratio of four points, and considering that because point D lies on the line at infinity, and its cross ratio with respect to points B and C are therefore not computed, the following relationship may be obtained for the harmonic sequence of points (A, B, C, D),

$$CR(\mathrm{AB};\mathrm{CD}) = \frac{\begin{vmatrix} a_0 & a_1 \\ c_0 & c_1 \end{vmatrix}}{\begin{vmatrix} b_0 & b_1 \\ c_0 & c_1 \end{vmatrix}} = \frac{\begin{vmatrix} 1 & 0 \\ 1 & \frac{p}{2} \end{vmatrix}}{\begin{vmatrix} 1 & p \\ 1 & \frac{p}{2} \end{vmatrix}}, \tag{3.6}$$

which upon expanding the determinants yields,

$$CR(\mathrm{AB};\mathrm{CD}) = \frac{\frac{\mathrm{p}}{2}}{\frac{-\mathrm{p}}{2}} = -1. \tag{3.7}$$

Given that this cross ratio is preserved under any value of p, the projective mapping of the trapezoid onto the "unit" square, and subsequent mapping of the unit circle onto the ellipse inscribing the trapezoid must preserve the polar points of the maximum area inscribing ellipse at the mid points of the two parallel edges, as they

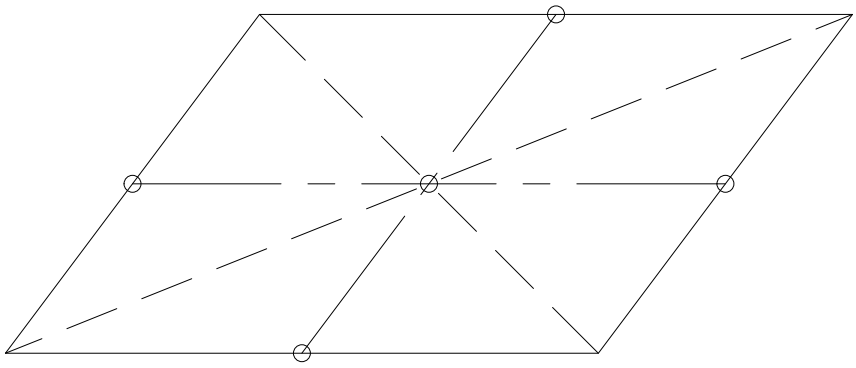


Figure 3.4: The parallelogram, and its mapped properties; diagonals in dotted lines with adjacent dotted lines connecting the midpoints of each opposing edge, while the centre of the pencil of inscribing ellipses, alongside the tangent points of the maximum area inscribing ellipse are circled.

share an ideal point on the line at infinity and form a harmonic ratio with their pole points.

Figure 3.4 illustrates a parallelogram, as well as some key properties. While the parallelogram differs from the mapping of the trapezoid in that the locus of the centres for all ellipses contained within the quadrangle is a double-mapped point, represented by the midpoints of the diagonals of the quadrangle, several other properties are also preserved; the polar points for the area maximizing inscribing ellipse lie incident with the midpoints of each side, and the centre for the area maximizing inscribing ellipse lies at the intersection of the diagonals and the lines connecting the midpoints of each set of parallel sides.

Though these two cases facilitate a projective mapping between the unit circle and the unit circles image within the convex quadrangle, the mapping will only generate

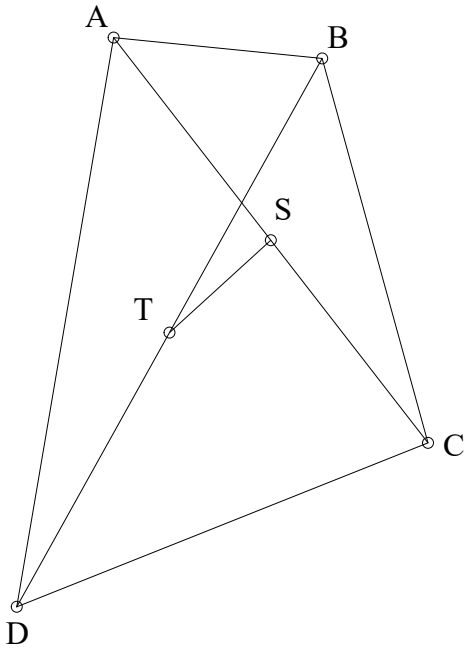


Figure 3.5: A general convex quadrangle, with no sides parallel.

the maximum area inscribing ellipse in the event that it also satisfies the conditions of an affine mapping for the parallelogram, and not just those of a projective transformation. For the trapezoid, the harmonic sequence of points formed by the vertices and polar points of the parallel edges is preserved. In Section 2.1 several properties of different transformation types were expounded upon. Recall that a projective mapping will only preserve the cross ratio between four points, but an affine transformation will preserve parallelism between lines, as well as the area ratios of distinct objects. It is by this preservation of relative size that the projective mapping generates the maximum area inscribing ellipse when the operation is performed on a parallelogram or a trapezoid; their parallel sides and the fact that the maximum area inscribing ellipse must intersect the midpoints of these sides provides a set of constraints to the mapping. This entails that, within the general case (such as that depicted within Figure 3.5) the projective mapping will not generally produce the maximum area inscribing ellipse.

It is obvious that the geometric properties of the quadrangle are decidedly different from those of a parallelogram in Figure 3.4; no two sides of the quadrangle in Figure 3.5 lie parallel to each other, nor do the midpoints of the diagonals intersect with each other such as in Figure 3.4. Furthermore, line ts is not parallel to any side of the quadrangle, indicating that there are no properties of an affine transformation preserved under this most general mapping, nor is there a harmonic ratio of parallel edges to be preserved, indicating that this transformation is purely projective and lacking the capacity to preserve a geometric constraint required to identify the maximum area inscribing ellipse.

As a result of the fact that the mapping between the quadrangle illustrated in Figure 3.5 and the unit square in Figure 3.1 does not preserve the line at infinity, the mapping will be incapable of preserving the ratios of areas required to generate the maximum area inscribing ellipse within the general quadrangle.

3.2 Area Optimisation through the Construction of a Non-Metric Basis Coordinate System

In Section 3.1 constructing the maximum area inscribing ellipse was attempted through the construction of a generalised projective mapping, treating the unit square and the unit circle as the image of the convex quadrangle and its maximum area inscribing ellipse. Problematically, this approach is limited to certain classes of convex quadrangle which allow their respective image transformations to satisfy the specific subset of cross ratios required for the relative area proportions to be preserved. However, in [2], Gfrerrer utilises non-orthogonal coordinate bases developed from the specific geometric properties of the quadrangle at hand [2], subsequently laying the foundation for the development of a one parameter pencil of ellipses whose area can

then be maximised within the bounds of the quadrangle.

3.2.1 Case Declarations

Given a convex quadrangle, comprising vertices A, B, C, D, Figure 3.6 can be constructed. Showcased within this figure is the quadrangle ABCD, alongside its diagonals, AC, and BD. Line TS represents the locus of centres for all conics contained within the quadrangle, and is constructed through the connection of the midpoints of the diagonals, points t and s. Point O represents the intersection of the two diagonals, while points E and F represent the intersections of the lines which serve to extend the sides of the quadrangle ABCD. Specifically, point F represents the intersection of the extension of sides AD and BC, while point E represents the intersection of the extensions of lines AB and CD.

Connecting points E and F facilitates the construction of line g, and the creation of point Q upon its intersection with line f, the extension of the diagonal BD. Point P is the intersection of line e, the extension of diagonal AC, and line g. Line m is the extension of line TS, the locus of centres for all inscribing conics within the quadrangle, and intersects line g at point R.

In order to simplify the analysis presented herein, it is possible to break the problem down into several cases, which can then be approached in increasingly specific ways in an effort to simplify the creation of the maximum area inscribing ellipse. Each of these cases will make reference to the variables described in Figure 3.6. Given that the family of quadrangles expressed by ABCD are strictly convex, point O lies within the quadrangle, while points E, F, P, Q always lie outside of the quadrangle.

- **Parallelogram**, all four of the points E, F, P, Q are improper points, meaning that line g is coincident with the line at infinity.

- **Convex Quadrangle**, or the general case, where points E, F, P, Q are all

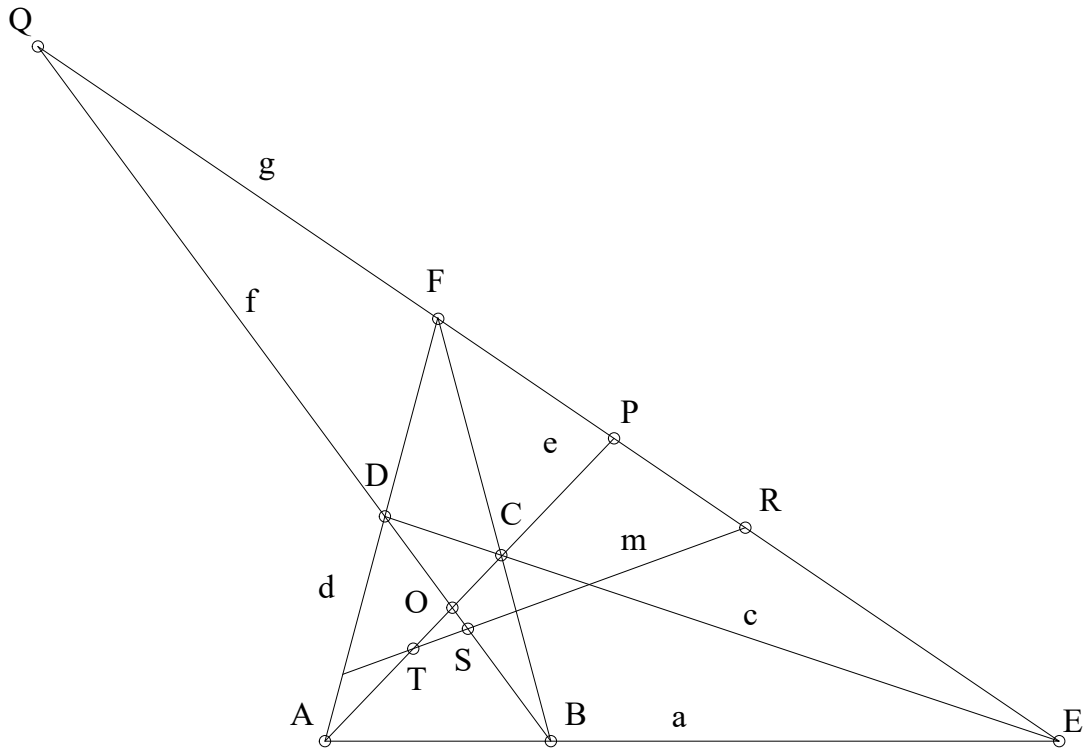


Figure 3.6: A general convex quadrangle ABCD, and its diagonal trilateral efg, and the midpoints of its diagonals; T, and S.

proper points.

- **Trapezoid**, points P and Q are proper points, while one of E or F is an improper point.

- **Kite**, contrary to the trapezoid case, here points E and F are proper points, while only one of P or Q is a proper point.

In general, the case of the parallelogram can be viewed as a simple affine correlation, whereas the cases of the general convex quadrangle, trapezoid, and kite require a more robust development procedure in order to fully characterise their maximum area inscribing ellipse.

3.2.2 General Quadrangle Squared Area Function Derivation

Characterization of the maximum area inscribing ellipse within a general quadrangle, in this context, must first begin with the construction of a non-orthogonal coordinate system based upon the geometry of the quadrangle at hand. While this construction facilitates a simplified derivation for the maximum area inscribing ellipse, it should be noted that it is unique to the quadrangle at hand; the direction of each affine basis vector will change with a manipulation of the geometry in question.

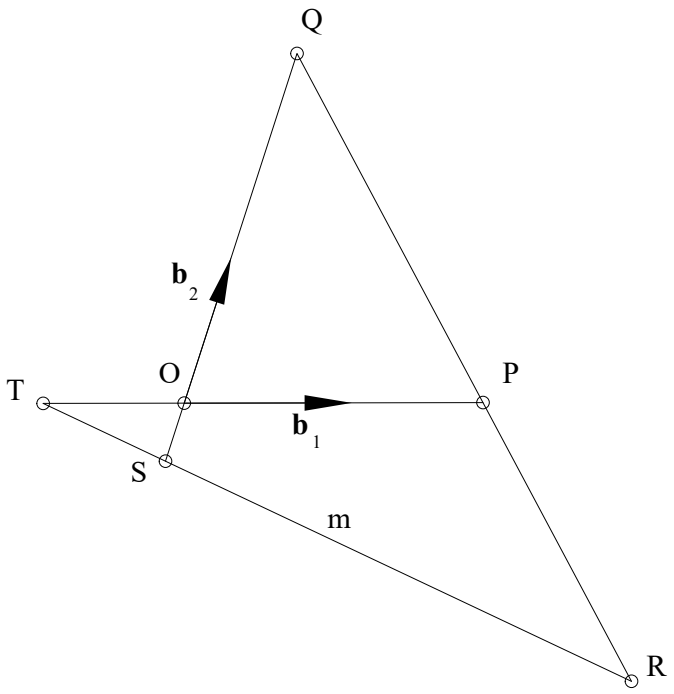


Figure 3.7: Non-orthogonal basis vectors for the coordinate system [2].

Figure 3.7 showcases an appropriate choice of affine basis vectors, b_1 and b_2. It should be noted that this diagram's variables coincide exactly with those of Figure 3.6. Due to the fact that the basis vectors for this coordinate system are generated based upon the geometry of the quadrangle at hand, they must be selected on a case by case basis. Furthermore, there is no guarantee that this selection will generate an orthogonal coordinate system, thus the angle between b_1 and b_2 can vary dramatically.

Due to the construction of this problem, specifically within the requirements to maintain an expression that is inclusive of improper points, all subsequent equations and line descriptions will be carried out using homogeneous point triplets, $[b_0, b_1, b_2]$, representative of the coordinate system selected from Figure 3.7. Therefore, it is possible to determine the locations of points O, P, Q, S, and T in terms of this homogeneous point triplet:

$$\mathrm{O} = [1, 0, 0]^T; \tag{3.8}$$

$$\mathrm{P} = [1, p, 0]^T; \tag{3.9}$$

$$\mathrm{Q} = [1, 0, q]^T; \tag{3.10}$$

$$\mathrm{S} = [1, 0, s]^T; \tag{3.11}$$

$$\mathrm{T} = [1, t, 0]^T. \tag{3.12}$$

Furthermore, because point R lies at the intersection of lines PQ and ST, its coordinates can be expressed as,

$$\mathrm{R} = [\mathrm{qt} - \mathrm{ps}, \mathrm{pt}(\mathrm{q} - \mathrm{s}), \mathrm{qs}(\mathrm{t} - \mathrm{p})]^{\mathrm{T}}. \tag{3.13}$$

Within the case of a general quadrangle, point R is a proper point which describes the midpoint of the line segment EF from Figure 3.6. From this basis vector set, the aim of the problem is to parameterise the pencil of conics inscribing the quadrangle ABCD in terms of a single variable. Conveniently, the locus of centres lying on line **m** provides a suitable choice for the parameterisation;

$$\mathbf{m}(u) = \begin{bmatrix} 1 \\ (1-u)t \\ us \end{bmatrix}, \tag{3.14}$$

such that the parameter $u \in \Re$. Points T, S, and R can be obtained with this expression through substitution of $u = 0, 1, \frac{q(t-p)}{qt-ps}$, respectively, into Equation (3.14). Now, while keeping this parameterisation in mind, it is necessary to compute the point equation of the conic section, in its quadratic form. Recall that it is possible to represent this pencil of inscribing ellipses in matrix-vector form,

$$k := \mathbf{b}^T \mathbf{A} \mathbf{b} = 0, \tag{3.15}$$

such that the vector $\mathbf{b}$ is a homogeneous point triplet, (b_0, b_1, b_2), and using the polar lines of the points O, P, and Q:

$$PQ \quad : \quad pqx_0 - qx_1 - px_2 = 0; \tag{3.16}$$

$$OQ \quad : \quad x_1 = 0; \tag{3.17}$$

$$OP \quad : \quad x_2 = 0; \tag{3.18}$$

Substituting Equations (3.17) through (3.18) into (3.15), the conic shape coefficient factors a_{00}, a_{01}, a_{02} and a_{12}, can be solved in terms of a_{00}:

$$a_{00} = pq; \tag{3.19}$$

$$a_{01} = -qa_{00}; \tag{3.20}$$

$$a_{02} = -pa_{00}; \tag{3.21}$$

$$a_{12} = 1. \tag{3.22}$$

This simplifies the conic shape coefficient matrix, $\mathbf{A}$, to,

$$\mathbf{A} = \begin{bmatrix} pqa_{00} & -qa_{00} & -pa_{00} \\ -qa_{00} & a_{11} & a_{00} \\ -pa_{00} & a_{00} & a_{22} \end{bmatrix}. \tag{3.23}$$

From Equation (3.23), and utilizing the polar point of the parameterised line $\mathbf{m}(u)$ as its ideal point on the line at infinity, such that $b_0 = 0$, and substituting the subsequent expression into the place of the basis vectors multiplying Equation (3.15), the following two expressions may be obtained,

$$(-q + s + u)a_{00} + t(1 - u)a_{11} = 0; \tag{3.24}$$

$$(-p + t(1 - u))a_{00} + sua_{22} = 0. \tag{3.25}$$

These can then be solved for the remaining point conic shape coefficient variables, a_{00}, a_{11}, a_{22}:

$$a_{00} = stu(1-u); \tag{3.26}$$

$$a_{11} = su(q-su); \tag{3.27}$$

$$a_{22} = t(1-u)(p-t(1-u)). \tag{3.28}$$

After substitution of Equations (3.26) through (3.28) into Equation (3.23), the following parameterised expression for the conic shape coefficient matrix, $\mathbf{A}$, is obtained;

$$\mathbf{A} = \begin{bmatrix} pqstu(1-u) & -qstu(1-u) & -pstu(1-u) \\ -qstu(1-u) & su(q-su) & stu(1-u) \\ -pstu(1-u) & stu(1-u) & t(1-u)(p-t(1-u)) \end{bmatrix}. \tag{3.29}$$

Thus, the point conic equation of the pencil of conics inscribing the ellipse can be described as:

$$k(u) : pqstu(1-u)b_0^2 + su(q-us)b_1^2 + t(1-u)(p-t(1-u))b_2^2 - 2qstu(1-u)b_0b_1 - 2pstu(1-u)b_0b_2 + 2stu(1-u)b_1b_2 = 0, \tag{3.30}$$

Which represents the parameterised point equations of the pencil of conics inscribing the quadrangle ABCD. Variables b_0, b_1, and b_2 are coordinates which lie on the constructed basis coordinate system of the quadrangle. From this point equation for the pencil of conics inscribing the quadrangle ABCD, the squared area function, α, can be developed in terms of the parameterised variable, u,

$$\alpha(u) = \frac{\det^2 \mathbf{A(u)}}{(\Delta_0(u))^3}. \tag{3.31}$$

After substitution of Equation (3.29) into (3.31), and taking $\Delta(u)_0$ as the discriminant of $\mathbf{A}(u)$, the following expression is obtained, representing the squared area function of the parameterised pencil of conics inscribing the quadrangle ABCD,

$$\alpha(u) = stu(u-1)((ps-qt)u - q(p-t)). \tag{3.32}$$

In order the represent the solutions to Equation 3.32, two variable substitutions, related directly to the quadrangle cases being examined are proposed. Namely, in order to simplify the expression, substitute,

$$\mu = ps - qt, \tag{3.33}$$

$$\nu = q(p-t), \tag{3.34}$$

in Equation (3.32), which yields the squared area function:

$$\alpha(u) = stu(u-1)(\mu u - \nu). \tag{3.35}$$

Henceforth, Equation (3.35) will be used to describe each maximum area inscribing ellipse within a convex quadrangle, ABCD. Specifically, this equation represents the area value of any conic section belonging to the pencil of sections inscribing the quadrangle, parameterised in terms of the variable u, contained on line segment m; the squared nature of the function facilitates the description of negative areas, arising from degenerate conic sections contained within the pencil of conics.

3.2.3 Convex Quadrangle Cases

In the previous section, the geometric constraints presented within a general convex quadrangle were used to derive a fully general parameterisation for any given convex quadrangle and its pencil of inscribing conics. While this function can be used to effectively describe the maximum area inscribing ellipse within any convex quadrangle, each of the cases presented in Section 3.2.1 can be treated independently, in order to provide simplified solutions, as well as generalisations about the behaviour of the pencil of conics within each type of convex quadrangle.

General Convex Quadrangle

Assuming that $\mu = ps - qt \neq 0$, which constrains the quadrangle ABCD to not be a trapezoid, Equation 3.32 simplifies to a third order polynomial with three real zeroes; 0, 1, and $\frac{\nu}{\mu}$, which represents the three singular conics within the pencil of inscribing conics within ABCD. As limiting cases, these three conics represent lines AC, BD, EF, seen in Figure 3.6. Due to the construction of the quadrangle ABCD, specifically because T and S are the only points which lie inside the quadrangle, it stands to reason that the area maximising inscribing ellipse of the quadrangle ABCD will fall in the range,

$$0 < \frac{\nu}{\mu} < 1. \tag{3.36}$$

From Equation (3.35), it is possible to compute the first and second derivatives of the squared conic area function, while using Equation (3.36) to check the value of u which is obtained therein:

$$\frac{d\alpha}{du} = 3\mu u^2 - 2(\mu + \nu)u + \nu; \tag{3.37}$$

$$\frac{d^2\alpha}{du^2} = 6\mu u - 2(\mu + \nu). \tag{3.38}$$

Upon evaluation, Equation (3.37) yields two real and distinct zeroes:

$$u_- := \frac{(\mu + \nu) - \sqrt{\mu^2 - \mu\nu + \nu^2}}{3\mu}; \tag{3.39}$$

$$u_+ := \frac{(\mu + \nu) + \sqrt{\mu^2 - \mu\nu + \nu^2}}{3\mu}. \tag{3.40}$$

Substituting each of these zeroes into Equation (3.38) shows that the location of the first zero supplies a positive concavity, while the location of the second zero generates a negative concavity; this is turns means that the first zero, u_- is concave down, and the positive zero of the squared area function. Equation (3.35) is plotted within the convex quadrangle ABCD in Figure 3.8, alongside the maximum area inscribing ellipse, generated by using the first zero of Equation (3.37).

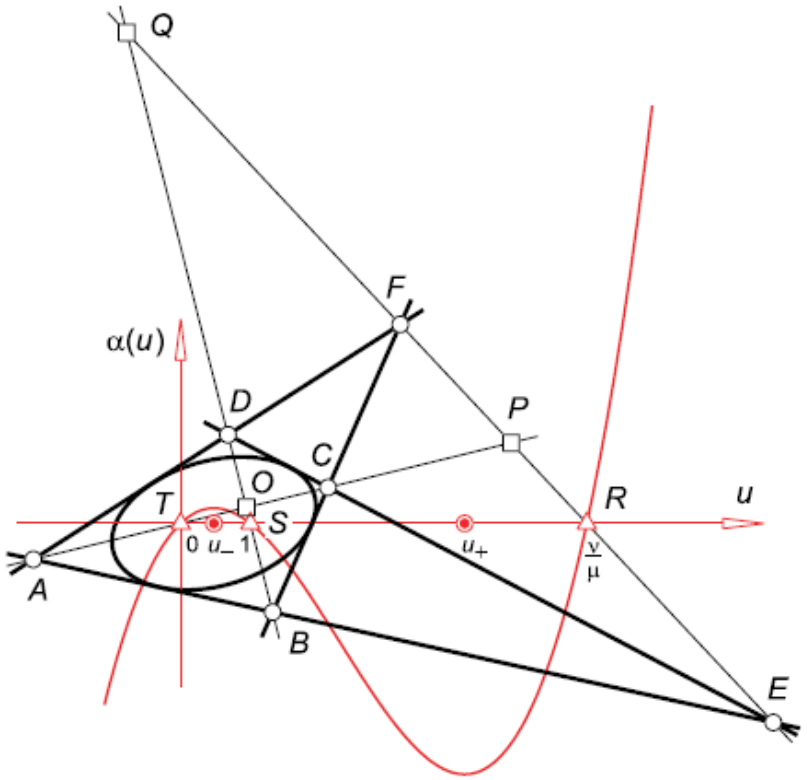


Figure 3.8: The squared area function, α(u), along line TS, for a general convex quadrangle.

Due to the variability within the construction and geometry of a general convex quadrangle, it is entirely possible that the value of $u_+ < 0$, however, u_- will always generate the maximum area inscribing ellipse within a general convex quadrangle. [1,2]

Trapezoid

When ABCD is a trapezoid with sides AB and CD lying parallel to each other, $\mu = ps - qt = 0$, and Equation (3.32) reduces to the polynomial,

$$\alpha(u) = qst(t-p)u(u-1). \tag{3.41}$$

Equation (3.41) simplifies to a simple parabolic function with respect to the parameterised variable u. Its first derivative is therefore a linear function:

$$\frac{d\alpha}{du} = qst(t-p)(2u-1). \tag{3.42}$$

Therefore, the only point at which this function will return a maxima is at $u = \frac{1}{2}$, the midpoint of the line TS. Figure 3.9 depicts the corresponding squared area function, which is itself a parabola in this simplified case, plotted alongside an example trapezoid ABCD, and its maximum area inscribing ellipse, and clearly illustrates the location for the centre of the maximum area inscribing ellipse.

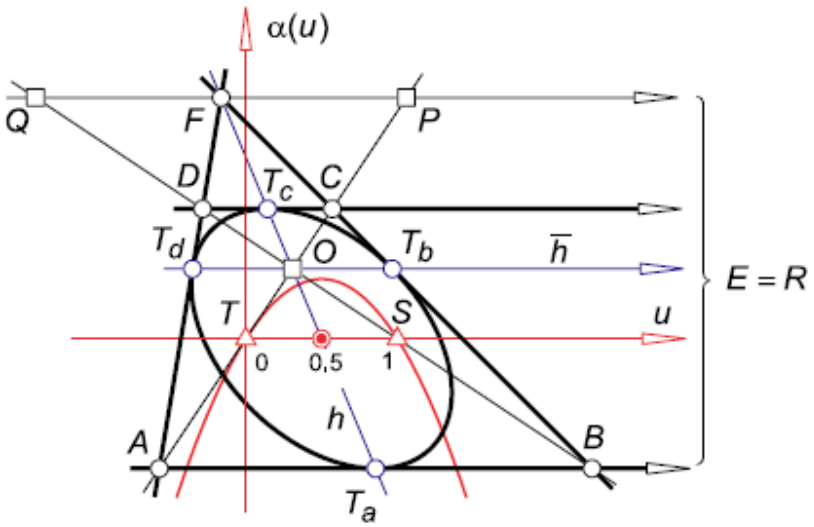


Figure 3.9: The squared area function, α(u), along line TS, for a trapezoid [2].

Convex Skew Kite

Finally, the last remaining special case to be considered within this discussion is that of a skew kite. Skew kites are a special case of kites which possess a projected line of symmetry along a non-orthogonal vector. Specifically, a skew kite such that the vertex P is an improper point, thus forcing point O to be incident with point T, the midpoint of diagonal AC. Moreover, this incidence of points T and O means that the

line m lies incident with the opposing diagonal, BD.

From the general case, presented in Equation (3.30), the skew kite case can be obtained by setting $pt = -1$, and subsequently taking the limit as t tends towards zero. After this procedure, the following expression for the pencil of inscribing conics present within a skew kite can be expressed as:

$$k(u) : qsu(u-1)b_0^2 - su(su-q)b_1^2 + (u-1)b_2^2 - 2su(u-1)b_0b_2 = 0. \tag{3.43}$$

Moreover, the squared area function Equation (3.32) becomes:

$$\alpha(u) = su(u-1)(q-su), \tag{3.44}$$

for t approaching zero. This third degree monovariate polynomial in μ has zeros present at points 0, 1, and q/s, yielding points T, S, and R, respectively. After taking the first derivative of the squared area function, $\alpha(u)$, the following two local extrema can be defined,

$$u_- = \frac{(s-q) - \sqrt{s^2 - sq + q^2}}{3s}, \tag{3.45}$$

$$u_+ = \frac{(s-q) + \sqrt{s^2 - sq + q^2}}{3s}, \tag{3.46}$$

where, analogous to the previous cases, u_- lies on m between points T (in this case, point T is also incident with point O), and S, while u_+ lies outside of the bounds of the kite, specifically between points B and R, or Q. Figure 3.10 shows this squared area function distribution, within an example skew kite, alongside the maximum area inscribing ellipse within the example quadrangle, generated from Equation (3.43),

such that $u = u_{-}$.

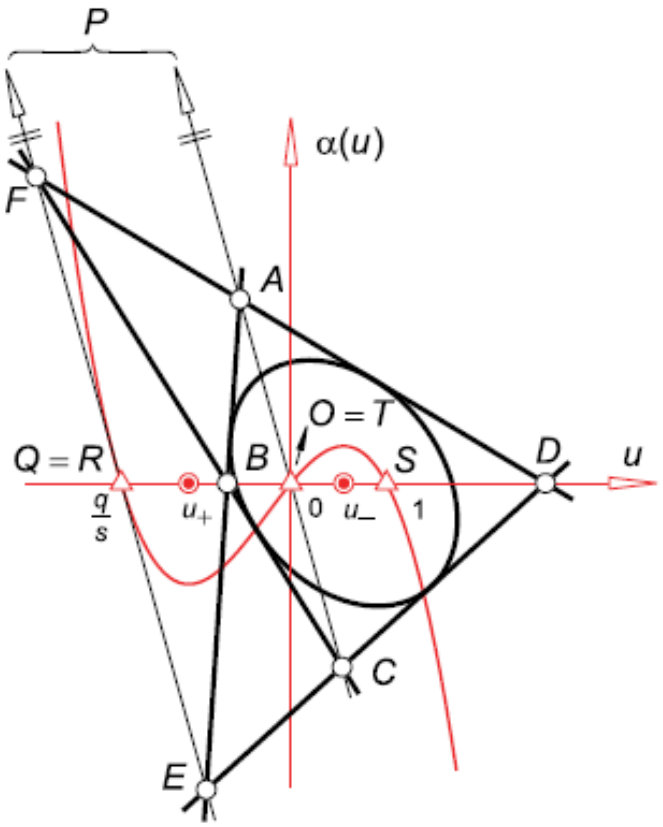


Figure 3.10: The squared area function, α(u), along line TS, for an asymmetric kite.

3.2.4 Applicability and Suitability of the Solution

While completely general, from a mathematical standpoint, generating the maximum area inscribing ellipse within a convex quadrangle employing non-orthogonal bases is cumbersome and impossible to automate using computational algorithms given the cases for which diagonals and/or quadrangle vertices are also vertices of the diagonal triangle on the line at infinity, implying a divide by zero condition during computation. Moreover, each and every unique quadrangle will necessarily produce different

requirements based upon its own geometry; each case and their subsequent derivations contained within Sections 3.2.3 through 3.2.3 are evidence of the computational limitations of this solution.

Furthermore, since the basis coordinate set used for this problem is non-orthogonal, the ellipse presented using this solution is defined in such a way so as to have a potentially very cumbersome metric associated to it. Both the area measure and point conic shape coefficients defined and evaluated in the non orthogonal coordinate system need to be transformed back to the original orthogonal system, potentially adding significant computational expense and allowing for concatenating computational errors.

What is required is a robust method to identify the maximum area inscribing ellipse directly within the original orthogonal coordinate system used to define the quadrangle.

Chapter 4

Generalised Solution Using Orthogonal Bases

In this chapter a novel generalised solution using orthogonal bases whereby the maximum area inscribing ellipse of any given convex quadrangle can be generated is presented. Throughout the course of the development of the generalised solution, homogeneous coordinates will be employed in keeping with the European standard; that is to say that the initial coordinate presented represents the homogenising coordinate, x_0, while x_1 and x_2 represent the x and y coordinates, respectively, when x_1 and x_2 are divided by x_0.

4.1 Generation of the Pencil of Ellipses

As shown within previous sections, it is possible to generate a pencil of ellipses given a system of five linear constraints, allowing one variable to be used in order to alter the shape of an ellipse. Through the lens of this problem, the sides of the convex quadrangle will be used in order to generate four of the five linear constraints required to define this pencil. In order to begin this process, a general convex quadrangle, such as the one in Figure 4.1 will be translated so as to force one side of the quadrangle

to lie coincident to the x_1 axis. After normalising the homogeneous coordinates, the quadrangle illustrated in Figure 4.1 has vertices located at the coordinate triples listed in Table 4.1.

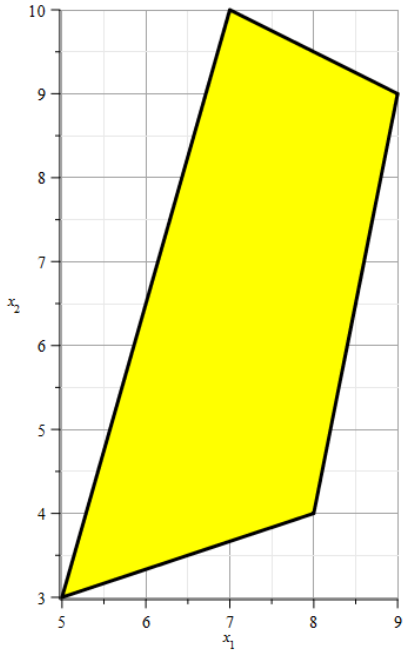


Figure 4.1: The initial quadrangle.

Table 4.1: Coordinates of the vertices of the example general convex quadrangle.

Vertex	x_0	x_1	x_2
A	1	5	3
B	1	8	4
C	1	9	9
D	1	7	10

Euclidean transformation matrices in the plane take the form of,

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 \\ a & \cos\theta & -\sin\theta \\ b & \sin\theta & \cos\theta \end{bmatrix}. \tag{4.1}$$

Where a and b represent the x_1 and x_2 displacements of vertex A, respectively. We define the angle θ to be the orientation of the edge possessing vertices with the two lowest x_2 coordinates, in this case $(1:5:3)$ and $(1:8:4)$. The angle θ is identified as,

$$\theta = \arctan\left(\frac{\mathrm{B}_{\mathrm{x}_2} - \mathrm{A}_{\mathrm{x}_2}}{\mathrm{B}_{\mathrm{x}_1} - \mathrm{A}_{\mathrm{x}_1}}\right), \tag{4.2}$$

while A and B represent the x_1 and x_2 coordinates of the vertex with the least values for both x_1 and x_2, respectively. However, it is important to note that the transformation matrix in Equation (4.1) needs to be inverted in order to transform the quadrangle from its displaced and rotated pose to the origin. This inversion yields:

$$\mathbf{T}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -a\cos\theta - b\sin\theta & \cos\theta & \sin\theta \\ a\sin\theta - b\cos\theta & -\sin\theta & \cos\theta \end{bmatrix}. \tag{4.3}$$

Now, through matrix multiplication with each of the vertices which define the initial convex quadrangle, the translated quadrangle is generated, and illustrated in Figure 4.2.

From this stage, the pencil of inscribing ellipses will be defined. The line coordinate triples used to define the pencil of ellipses are defined by:

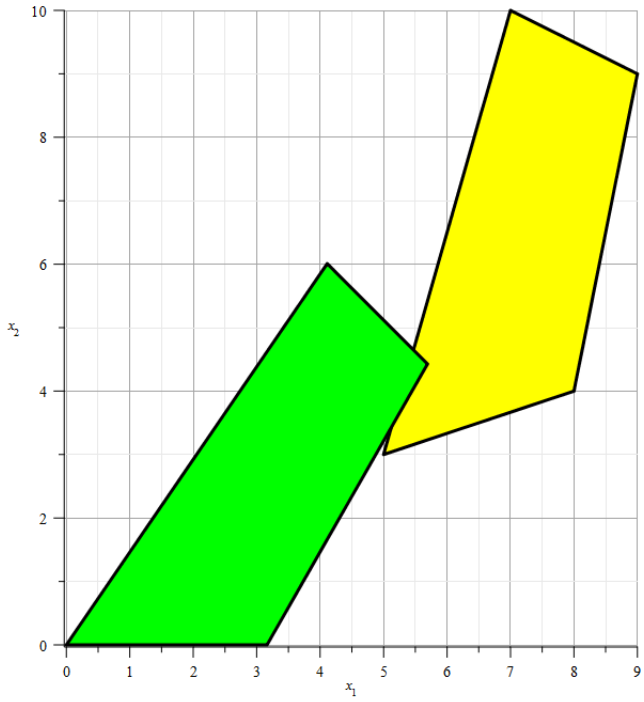


Figure 4.2: Comparison between the specified and transformed quadrangle.

$$\begin{bmatrix} X_0 \\ X_1 \\ X_2 \end{bmatrix} . \tag{4.4}$$

The line conic shape coefficient matrix is the matrix of coefficients which provide the shape constraints for the general second order curve. It is defined below, as the inverse of the symmetric matrix $\mathbf{A}$,

$$\mathbf{A}_L = \mathbf{A}^{-1} = \begin{bmatrix} A_{00} & A_{01} & A_{02} \\ A_{01} & A_{11} & A_{12} \\ A_{02} & A_{12} & A_{22} \end{bmatrix} . \tag{4.5}$$

The line conic implicit equation is obtained through multiplication of the transposed line coordinate triple with $\mathbf{A}_L$, followed by multiplication by the original column vector. Equation (2.16) is replicated below, for convenience,

$$k := A_{00}X_0^2 + 2A_{01}X_0X_1 + 2A_{02}X_0X_2 + A_{11}X_1^2 + 2A_{12}X_1X_2 + A_{22}X_2^2 = 0. \quad (4.6)$$

Equation (4.6) is the general projective line conic. With this definition in mind, it is important to define constraints for the pencil of second order curves represented by this equation. Four of the five constraints for this inscribing line conic will be created through the line coordinates of each side of the quadrangle. For example, the line coordinate which corresponds to the line coordinate of the first side is defined as the Grassmannian expansion of the point coordinates of points A and B, respectively,

$$L_{AB} = \begin{vmatrix} X_0 & X_1 & X_2 \\ 1 & \mathrm{A_{x_1}} & \mathrm{A_{x_2}} \\ 1 & \mathrm{B_{x_1}} & \mathrm{B_{x_2}} \end{vmatrix}. \quad (4.7)$$

Similarly, the line coordinates which correspond to the remaining edges of the quadrangle are defined. In order to provide the final constraint which yields a parametric expression for the pencil of inscribing ellipses, a fifth and final constraint is provided through the inclusion of a pole point, specifically the pole point of the ellipse on the edge of the quadrangle transformed to the x_1 axis.

Now, it is obvious that any inscribing ellipse within the quadrangle will necessarily be tangent to all four of the sides. This is analogous to the special case of a unit circle lying within the unit square; the unit circle lies incident and tangent to the unit square at the midpoint of each side of the square. Given the proper definition,

it is possible to create a fifth line which the ellipse would be tangent to, as well. Problematically, this line would need to be parameterised in such a way so as to ensure that it can describe a pencil of lines, separate from those already present, which always lie tangent to this family of ellipses. However, defining such a line is an unnecessarily difficult undertaking.

It is possible through the construction of this problem to generate a fifth linear constraint, without specifying an additional unique line in the system. Namely, this fifth constraint will be expressed as the location of the pole point on the x_1 axis, point a_x. As a_x varies from 0 to the coordinate value of the next vertex on the x_1 axis, the pole point of the inscribing ellipse on that quadrangle edge is implied. The location of a_x is used as a variable in the line coordinates of a distinct linear constraint from the x_1 axis. Subsequently, this places the shape coefficients of the general projective line conic equation of the inscribing ellipse in terms of a_x. The value of a_x that maximises the area of the inscribing ellipse can then be determined. This constraint is made possible due to the geometric transformation which forced the quadrangle to be located at the origin, with one side incident to the x_1 axis.

The pole point, a_x, lies on line g whose line coordinates are:

$$g := \begin{bmatrix} G_0 : & G_1 : & G_2 \end{bmatrix}. \tag{4.8}$$

Once this vector of line coordinates is multiplied with the line conic shape coefficient matrix, it will yield the pole point coordinates of an ellipse with line g. In other words, the tangent point of line g and the ellipse. The pole point on any line g with any ellipse is the array in Equation (4.9):

$$\begin{bmatrix} A_{00}G_0 + A_{01}G_1 + A_{02}G_2 & A_{01}G_0 + A_{11}G_1 + A_{12}G_2 & A_{02}G_0 + A_{12}G_1 + A_{22}G_2 \end{bmatrix}.$$

Due to homogeneity, each element in this vector can be divided by the first element, thereby projecting the line coordinates into the Euclidean plane. Since the pole point, a_x, lies on the x_1 axis, this implies that the x_2 term is equal to zero, while the second term, which corresponds to the x_1 axis, is equal to a_x,

$$a_x = \frac{A_{01}G_0 + A_{11}G_1 + A_{12}G_2}{A_{00}G_0 + A_{01}G_1 + A_{02}G_2}. \tag{4.9}$$

However, point a_x lie on the x_1 axis, which possesses the line coordinates,

$$\begin{bmatrix} G_0 : & G_1 : & G_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}. \tag{4.10}$$

Substituting these values into Equation (4.9) yields,

$$a_x = \frac{A_{12}}{A_{02}}, \tag{4.11}$$

or, in normal form,

$$A_{02}a_x - A_{12} = 0. \tag{4.12}$$

This line conic shape coefficient form, representing a_x, is a truly general form. Provided that the quadrangle had been located so as to allow one side to lie incident with the x_1 axis, this form will be true for any given convex quadrangle. In the event that the quadrangle has not been located in such a way, the same parameterisation process may be undertaken in order to express this constraint in terms of the line conic shape coefficients, however, it becomes significantly more complicated due to the fact that the line coordinate values of G_1 and G_2 are not, in general, equal to zero. While line g lies incident with the x_1 axis, however, the following matrix can be used to determine the values of A_{ij} through Grassmannian expansion [21],

$$\text{Grassmanian}(\mathbf{A}_{\text{L}}) = \begin{bmatrix} A_{00} & A_{01} & A_{02} & A_{11} & A_{12} & A_{22} \\ 0 & 0 & a_x & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & R_6 \\ S_1 & S_2 & S_3 & S_4 & S_5 & S_6 \\ T_1 & T_2 & T_3 & T_4 & T_5 & T_6 \\ U_1 & U_2 & U_3 & U_4 & U_5 & U_6 \end{bmatrix} \tag{4.13}$$

Such that R_i, S_i, T_i, and U_i, all represent the line conic coefficients derived from substituting each set of line coordinates which represent the sides of the quadrangle into Equation (4.6). In this instance, R_i only contains one non-zero element, due to its incidence with the x_1 axis. Due to the size of this matrix, the general solutions for A_{ij} obtained through Grassmannian expansion will not be presented, though the numerical result (corresponding to the aforementioned demonstration quadrangle) will be presented below.

Once the line conic shape coefficient matrix has been determined in terms of the parameter a_x, it is important to convert this line coefficient matrix to its dual point conic coefficient matrix in order to plot the conics themselves. While line coordinates are exceedingly useful in representing linear constraints within this system, each line coordinate triple represents a line that lies tangent to a point on the ellipse. The line conic equation represents the family of lines tangent to the ellipse, and is therefore not suitable for visualisation.

It can be shown [27] that the point conic shape coefficient matrix, $\mathbf{A}$ is proportional to the inverse of the line conic shape coefficient matrix:

$$\mathbf{A} = \mathbf{A}_L^{-1} \propto \text{adj}\mathbf{A}_\text{L} \tag{4.14}$$

However, since matrix $\mathbf{A}_L$ is never rank deficient for a proper ellipse, its determinant is never identically equal to zero. We find that simpler expressions are obtained using the Laplacian expansion theorem [28] to evaluate $\mathbf{A}^{-1}$ by evaluating the individual minor determinants but not dividing the resulting adjoint matrix by the determinant of $\mathbf{A}_L$. Following this process, the initial value of the matrix $\mathbf{A}$ is normalized by the numerical value preceding the a_{00} term, and the following point conic shape coefficient matrix for the example quadrangle depicted earlier is obtained:

$$\mathbf{A} = \begin{bmatrix} a_x^2 & -a_x & \frac{1}{133}\sqrt{10}a_x^2 + \frac{11}{133}a_x \\ -a_x & 1 & -\frac{5}{133}a_x^2 + \frac{20}{133}\sqrt{10}a_x - \frac{9}{7} \\ \frac{1}{133}\sqrt{10}a_x^2 + \frac{11}{133}a_x & -\frac{5}{133}a_x^2 + \frac{20}{133}\sqrt{10}a_x - \frac{9}{7} & \frac{920}{17689}a_x^2 - \frac{200}{931}\sqrt{10}a_x + \frac{81}{49} \end{bmatrix}. \tag{4.15}$$

From this matrix, it is evident that the parameterisation was successful in defining a pencil of conics characterized solely by the pole point, a_x. Special care should be taken in order to ensure that this equation makes use of rational numbers in order to avoid concatenating computational errors resulting from numerical truncation. Upon pre and post multiplication with the point $(x_0 : x_1 : x_2)$, the following general point conic function is generated; its full form being too cumbersome to replicate in the text, thus Equation (2.18) is replicated in its place,

$$a_{00}x_0^2 + 2a_{01}x_0x_1 + 2a_{02}x_0x_2 + a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = 0 \tag{4.16}$$

Figure 4.3 shows that, within the bounds of the bottom edge of the quadrangle,

substituting different values of a_x into Equation (4.16) yields ellipses contained entirely within the quadrangle. Furthermore, as a_x tends from the origin towards the bounding vertex of the quadrangle, it is clear that the ellipse generated by Equation (4.16) starts as a degenerate ellipse at $a_x = 0$, tends towards some maximum value, and finally generates an additional degenerate ellipse at the opposing vertex located on the x_1 axis.

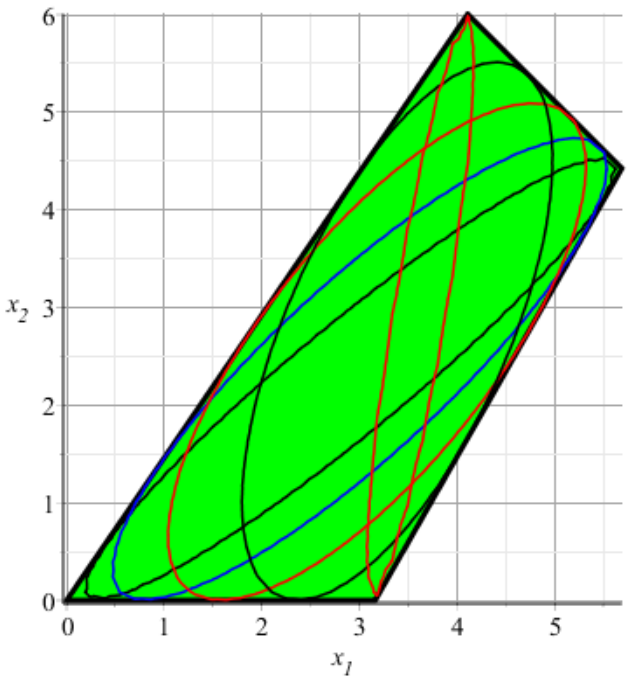


Figure 4.3: A sample of the pencil of ellipses contained within the example quadrangle.

4.2 Elliptical Area Maximisation

Due to the tendency of this pencil of ellipses towards a maximum at some value for a_x within the quadrangle, it is possible through the use of this parameterisation to determine the exact solution for the area maximising inscribing ellipse. However, as discussed in Chapter 2.5 the area of an ellipse generated in this fashion is inherently

nontrivial. Hence, the generalised conic area equation is replicated below as Equation (4.17), as the area function of the pencil of conics, k,

$$Area(k) = \left| \frac{\pi \det(\mathbf{A})}{(\sqrt{\Delta_0})^3} \right|. \tag{4.17}$$

In general, this equation can be used to define the area of any given ellipse which is expressed in its point conic coefficient matrix form. Specifically, for the purposes of the problem at hand, the point conic shape coefficient matrix $\mathbf{A}$ is parameterised in terms of the pole point, a_x. Combining these two facts means that equation (4.17) yields an expression which places the area of this pencil of conics in terms of this pole point.

Since this area function remains continuous within the bounds of the quadrangle, the derivative of the area function can be used to define the local maximum area, $\partial\mathbf{A}/\partial a_x$ within the bounds of the quadrangle for a_x, subsequently using the value of a_x which lies within the two vertices of the quadrangle on the x_1 axis. Through evaluation of multiple different pencils of conics generated through their respective quadrangle, it was found that Equation (4.17) generates a minimum of two global zeroes. Due to the nature of this function, it will generate multiple distinct zeroes; however, only one of these zeroes corresponds to a physically meaningful value. Specifically, there will be one value of a_x which lies between the vertices of the quadrangle on the x_1 axis, corresponding to the maximum area inscribing ellipse. For the example quadrangle, the first derivative of the area function evaluates to,

$$\frac{\partial \mathbf{A_p}}{\partial a_x} = \frac{P_1}{P_2} \frac{\pi(T-K)}{FH}, \tag{4.18}$$

where,

$$P_1 = 266;$$

$$P_2 = 5;$$

$$T = 44875\sqrt{5}a_x^8 + 18254440\sqrt{5}a_x^6 + 977614884\sqrt{5}a_x^4 + 7830778880\sqrt{5}a_x^2 + 4731699200\sqrt{5};$$

$$K = 1125a_x^9 + 1914950\sqrt{2}a_x^7 + 266060640\sqrt{2}a_x^5 + 5631652320\sqrt{2}a_x^3 + 14903296000\sqrt{2}a_x;$$

$$F = (40\sqrt{5}\sqrt{2}a_x^2 - 5a_x^3 + 608\sqrt{5}\sqrt{2} - 958a_x)^3;$$

$$H = \sqrt{a_x(40\sqrt{5}\sqrt{2}a_x^2 - 5a_x^3 + 608\sqrt{5}\sqrt{2} - 958a_x)}.$$

Once this function has been obtained, Maple can be used to evaluate the zeroes and provide a value of the pole point, a_x, at which the area is maximised. Figure 4.4 illustrates the maximised area inscribing ellipse contained within the example quadrangle.

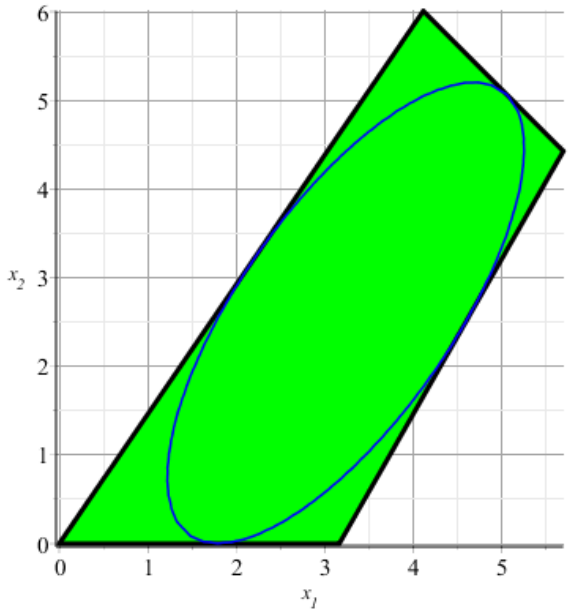


Figure 4.4: The maximum area inscribing ellipse within the example quadrangle.

4.3 Test Cases

In order to demonstrate the efficacy of the solution, several test cases are proposed in an effort to show compliance with solutions that are expected based upon their affine equivalence of arbitrary convex quadrangles to the unit square. First, a non-unit square is proposed, following which a rhombus of the same dimensions is defined. While the maximum area inscribing ellipse within the rhombus is non-circular in nature, it possesses the same area as the circle which inscribes the square of the same side lengths, as well as having tangent points lying in the exact centres of each side of the quadrangle. Secondly, an irregular trapezoid is presented, following which a general asymmetric convex quadrangle with no parallel sides has its maximum area inscribing ellipse computed.

4.3.1 Square

It is clear that the largest area inscribing ellipse within a square is a circle which touches the square at the midpoint of each side, and whose centre lies on the intersection of each corner diagonal of the square. Though trivial in nature, this solution supplies a simple method whereby the optimisation of the area function may be demonstrated for a quadrangle whose solution is known. For example, a square ABCD, with vertices; A(5,5), B(10,5), C(10,10), D(5,10), shown in Figure 4.5.

In order to perform the parameterisation properly through the assignment of the pole point, a_x, it is necessary to relocate the square such that point A lies incident with the coordinate origin. Figure 4.5 shows the initial square, alongside its translated counterpart.

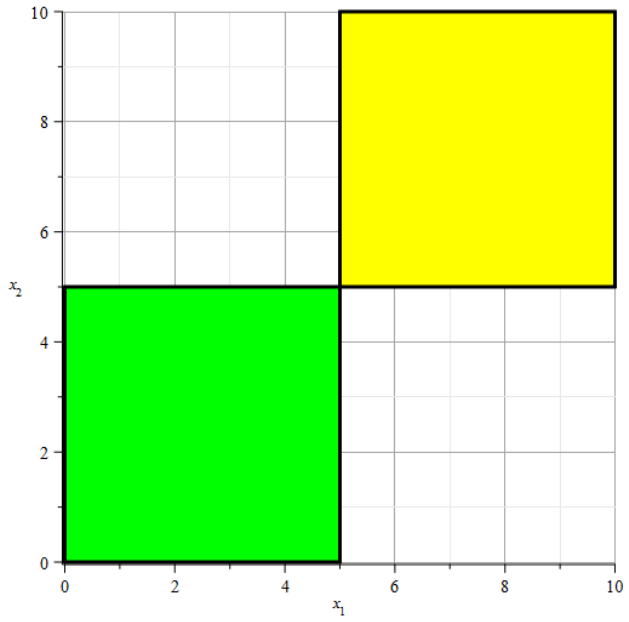


Figure 4.5: Translated square used in the parameterisation.

Parametrisation of the pencil of conics contained within the square is most easily facilitated through the use of line coordinates; specifically those of the lines which govern the boundaries of the square. Equations (4.19)-(4.22) depict the line coordinates

of each side of the square, in counter clockwise order, from the origin:

$$\begin{aligned} \text{AB} &= X_2 = 0; && (4.19)\\ \text{BC} &= 5X_0 - X_1 = 0; && (4.20)\\ \text{CD} &= 5X_0 - X_2 = 0; && (4.21)\\ \text{DA} &= X_1 = 0. && (4.22) \end{aligned}$$

From the line coordinates of the bounding lines which define the square, substitution into the line conic parametric equation will define a series of equations that can then be used to constrain a pencil of conics. Equations (4.23)-(4.26) result:

$$\begin{aligned} A_{22} &= 0; && (4.23)\\ 25A_{00} - 10A_{01} + A_{11} &= 0; && (4.24)\\ 25A_{00} - 10A_{02} + A_{22} &= 0; && (4.25)\\ A_{11} &= 0. && (4.26) \end{aligned}$$

Defining the line conic shape coefficients themselves requires Grassmanian expansion of equations (4.23)-(4.26) together with the parametric variable, a_x, along the x_1 axis. Equation 4.27 showcases the matrix used for the Grassmanian expansion; each columns' values are representative of the coefficients immediately preceding the corresponding variables within Equation (4.68).

$$\text{Grassmanian}(\mathbf{A}_{\text{L}}) = \begin{bmatrix} A_{00} & A_{01} & A_{02} & A_{11} & A_{12} & A_{22} \\ 0 & 0 & a_x & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 25 \\ 25 & -10 & 0 & 1 & 0 & 0 \\ 25 & 0 & -20 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \tag{4.27}$$

Upon completion of the Grassmanian expansion of the matrix in Equation (4.27), the line conic shape coefficient matrix in Equation (4.28) is obtained. While this matrix does, in a mathematically robust fashion, represent the pencil of conics contained within the square, it is not necessarily useful from a visualisation standpoint. Thus, Laplacian expansion is used in order to express the dual of the line within the plane, and develop a parametric solution based upon point coordinates.

$$\text{adj}(\mathbf{A}_{\text{L}}) = \begin{bmatrix} \frac{-2}{5} & -1 & -1 \\ -1 & 0 & -a_x \\ -1 & -a_x & 0 \end{bmatrix}. \tag{4.28}$$

The result is Equation (4.29), the matrix expression of the pencil of conics which inscribe the square:

$$\mathbf{A} = \begin{bmatrix} a_x^2 & -a_x & -a_x \\ -a_x & 1 & \frac{2}{5}a_x - 1 \\ -a_x & \frac{2}{5}a_x - 1 & 1 \end{bmatrix}. \tag{4.29}$$

Pre and post multiplication of the point conic shape coefficient matrix in Equation (4.29) with the homogeneous point triple yields the general expression for the second order projective curve which describes the pencil of point conics inscribing the square:

$$k := a_x^2 x_0^2 - 2a_x x_1 x_0 - 2a_x x_2 x_0 + x_1^2 + (\frac{4}{5}a_x - 2)x_2 x_1 + x_2^2 = 0. \tag{4.30}$$

In order to provide a Cartesian depiction of Equation (4.30) the projectivity of $x_0 = 1$ is taken. Furthermore, a small subset of the conics inscribing the square are plotted, in Figure 4.6. Each conic plotted was generated through assigning a value to the variable a_x with respect to the length of the side which lies along the x_1 axis; specifically 0.1, 0.25, 0.5, 0.75, and finally 0.99 times the length of the side which lies incident with the x_1 axis.

Figure 4.6 clearly shows that for a range of values constrained to be within the edge of the square side length along the x_1 axis. Now, the area of the pencil of conics can be computed as a function of a_x. Problematically, due to the nature of the second order projective curve definitions, it is possible that, for a given value of a_x, Equation (4.30) will define a hyperbola, or a parabola; a global optimisation of the area values generated by this pencil of conics will necessarily always yield an infinite area. Therefore, Equation (4.31), below, is the simplified first derivative of the area function for the square,

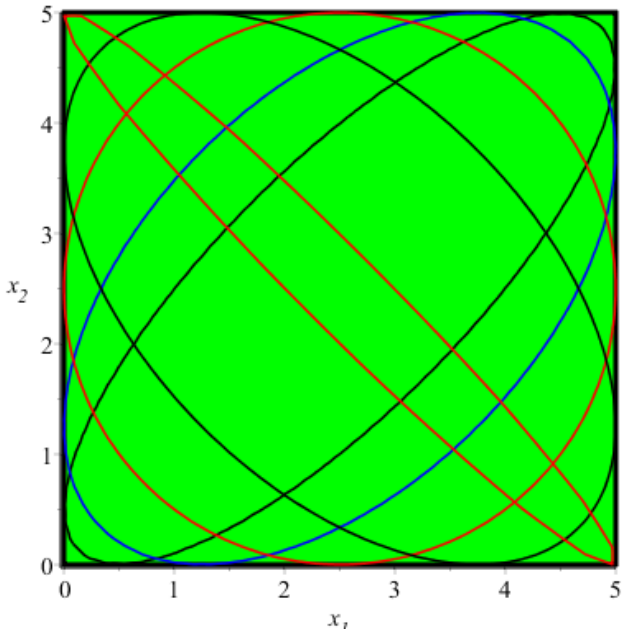


Figure 4.6: Members of the pencil of ellipses inscribing the square.

$$\frac{\partial \alpha}{\partial a_x} = -\frac{5}{2}\frac{(a_x - 5)a_x\pi}{\sqrt{-a_x(a_x - 5)}}. \tag{4.31}$$

Due to the construction of this problem, the area maximum inscribing ellipse of a square is always going to be a circle whose tangent points lie in exactly the midpoints of each line. In keeping with this fact, the only global zero of the area derivative function is located at 2.5 units along the x_1 axis; the exact middle of the edge of the square which lies incident with the x_1 axis. Equation (4.32) represents the result of substituting this value for a_x into Equation (4.30), alongside a Euclidean projectivity of $x_0 = 1$,

$$k_{max} := \frac{25}{4} - 5x_1 - 5x_2 + x_1^2 + x_2^2 = 0. \tag{4.32}$$

Equation (4.32) represents the maximum area inscribing conic section present within the square. Note that the factor preceding x_1x_2 is absent, thus determining a circle that is offset from the origin by five units in the x_1 and x_2 directions. Finally,

Figure 4.7 shows the maximum area inscribing ellipse within the square.

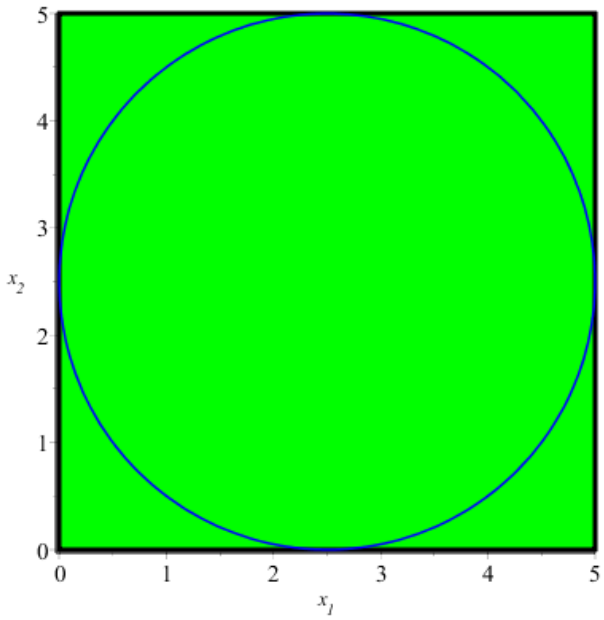


Figure 4.7: The maximum area inscribing ellipse within a square.

4.3.2 Parallelogram

Much like the aforementioned case of the square, it is clear that the maximum area inscribing ellipse within a parallelogram is the ellipse whose centre lies on the intersection of the internal diagonals, and whose tangents points lie exactly at the midpoint of each side of the parallelogram. Furthermore, due to the affine equivalence of the square and parallelogram, given a rhombus of side lengths equal to the square presented in Section 4.3.1, the maximum area inscribing ellipse will posses the same area as that of the circle which inscribes the square. For example, a rhombus ABCD, with vertices; A(5,5), B(10,5), C(12.5,10), D(7.5,10), is shown in Figure 4.8.

While the edge lengths of the rhombus remain identical to that of the aforementioned square, the angle at which the rhombus is deflected is entirely arbitrary. However, in keeping with the procedure required to generate the parameterisation,

the rhombus must be relocated such that one side lies incident with the x_1 axis, shown in Figure 4.8.

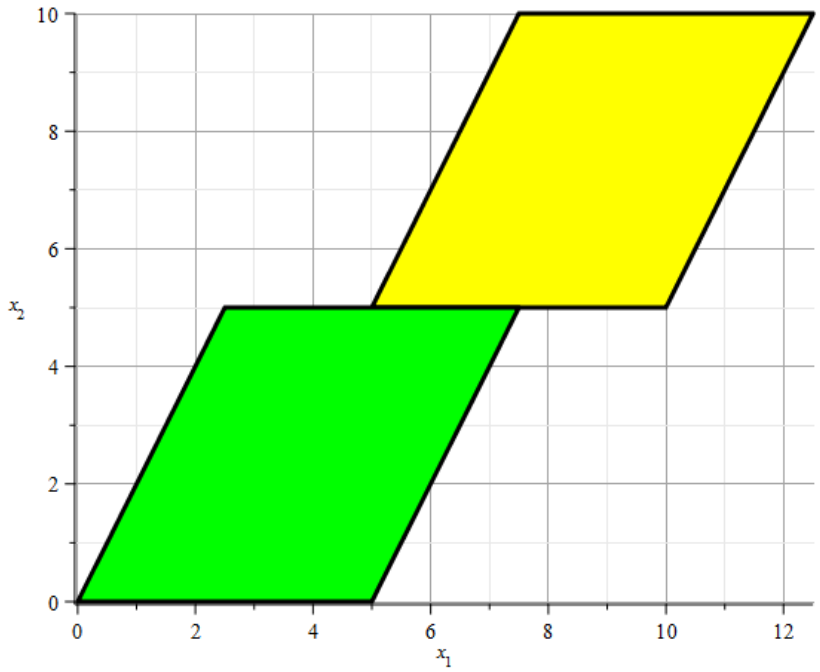


Figure 4.8: Translated rhombus used for parameterisation.

From here, the line coordinates which describe each side of the parallelogram must be determined. Within the case of the rhombus, the line coordinates, specifically those of lines BC and CD are not symmetrical. However, line AB possesses the exact same value for the rhombus as was determined during the computation of the maximum area inscribing ellipse within the square. The resulting line equations are:

$$\begin{aligned} \text{AB} &= X_2; && (4.33)\\ \text{BC} &= 5X_0 - X_1 + \frac{1}{2}X_2; && (4.34)\\ \text{CD} &= 5X_0 - X_2; && (4.35)\\ \text{DA} &= X_1 - \frac{1}{2}X_2. && (4.36) \end{aligned}$$

From the line coordinates of the bounding lines which define the parallelogram, substitution into the line conic parametric equation will define a series of equations that can then be used to constrain a pencil of conics. The inclusion of an additional term within lines BC and DA complicates the conic boundary functions required to define the pencil of conics present within the rhombus:

$$A_{22}; \quad (4.37)$$

$$25A_{00} - 10A_{01} + 5A_{02} + A_{11} - A_{12} + \tfrac{1}{4}A_{22}; \quad (4.38)$$

$$25A_{00} - 10A_{02} + A_{22}; \quad (4.39)$$

$$A_{11} - A_{12} + \tfrac{1}{4}A_{22}. \quad (4.40)$$

Defining the line conic shape coefficients themselves requires Grassmanian expansion of Equation (4.41) alongside the parametric variable, a_x, along the x_1 axis. The process of Grassmanian expansion in order to determine the line conic shape coefficients remains unchanged from the case of the square, while the parametric variable a_x, and subsequently its line equation, remains the exact same, despite the changes within the boundary constraints of the parallelogram. The matrix used to articulate this expansion is,

$$\text{Grassmanian}(\mathbf{A}_{\text{L}}) = \begin{bmatrix} A_{00} & A_{01} & A_{02} & A_{11} & A_{12} & A_{22} \\ 0 & 0 & a_x & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 25 & -10 & 5 & 1 & -1 & \frac{1}{4} \\ 25 & 0 & -10 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & \frac{1}{4} \end{bmatrix}. \tag{4.41}$$

Upon completion of the Grassmanian expansion, the line conic shape coefficient matrix in Equation (4.41) is determined. If pre and post-multiplied by homogeneous line triplets, this matrix would determine the family of lines which lies tangent to the pencil of conics contained within the parallelogram:

$$\mathbf{A}_L = \begin{bmatrix} \frac{-2}{5} & \frac{-3}{2} & -1 \\ \frac{-3}{2} & -a_x & -a_x \\ -1 & -a_x & 0 \end{bmatrix} \tag{4.42}$$

Importantly, the coefficient A_{11} is no longer equal to zero; this necessarily means that the corresponding line conic would have a non-zero factor multiplying its X_1X_2 term, which then forces every inscribing conic section to be non-circular in nature. Equation (4.43) is the adjoint of matrix $\mathbf{A}$ yielding the point form of this pencil of conics:

$$\mathrm{adj}(\mathbf{A}_{\mathrm{L}}) = \begin{bmatrix} a_x^2 & -a_x & -\frac{1}{2}a_x \\ -a_x & 1 & \frac{2}{5}a_x - \frac{3}{2} \\ -a_x & \frac{2}{5}a_x - \frac{3}{2} & -\frac{2}{5}a_x - \frac{9}{4} \end{bmatrix}. \tag{4.43}$$

Equation (4.44) results from pre and post multiplication of Equation (4.43) with a homogeneous point triplet, and describes the pencil of inscribing conic sections contained within the rhombus:

$$k := a_x^2 x_0^2 - 2a_x x_1 x_0 - a_x x_2 x_0 + x_1^2 + (\frac{4}{5}a_x - 2)x_2 x_1 + x_2^2(-\frac{2}{5}a_x + \frac{9}{4}). \tag{4.44}$$

From this conic pencil, several conics can be plotted alongside the parallelogram which they inscribe. Figure 4.9 illustrates several inscribing ellipses corresponding to the same relative a_x values as for the square.

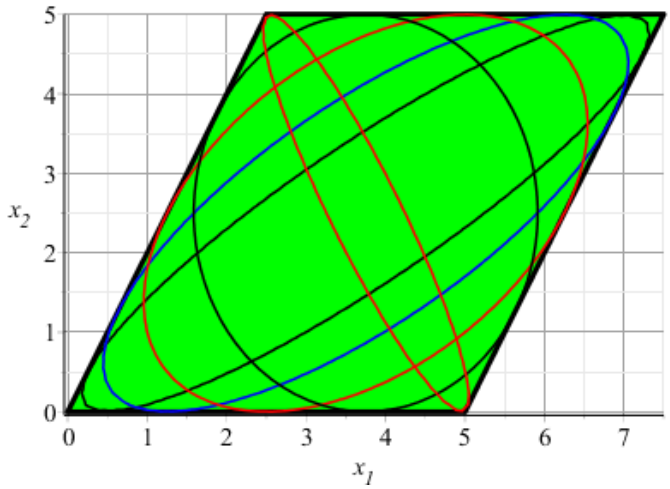


Figure 4.9: Several inscribing ellipses within the rhombus.

Given that the rhombus presented within this example can be mapped to the aforementioned square through the use of an affine collineation, the maximum area

inscribing ellipse contained within it will share its tangent points with those of the square; each tangent point will lie at the exact middle point of edge, and thus the maximum area inscribing ellipse will occur at $a_x = 2.5$. Equation (4.45) is the area derivative function used to find this local area maximum:

$$\frac{\partial \alpha}{\partial a_x} = -\frac{5}{2}\frac{(a_x - 5)a_x\pi}{\sqrt{-a_x(a_x - 5)}}. \tag{4.45}$$

Not only does the maximum area inscribing ellipse occur at the same a_x as that of the square, but through affine equivalence and the preservation of area ratios under an affine transformation, the circle, mapped to its corresponding ellipse within the rhombus, which is a skewed square, possess the exact same area. Thus, Equations (4.45) and (4.31) are identical, even though the point conic shape coefficient matrices and their corresponding point conic equations are not.

$$k_{max} := \frac{25}{4} - 5x_1 - \frac{5}{2}x_2 + x_1^2 - x_1x_2 + \frac{5}{4}x_2^2 = 0. \tag{4.46}$$

Upon substitution of $a_x = 2.5$ into Equation (4.44), we obtain Equation (4.46) which describes the maximum area inscribing ellipse contained within the rhombus, and is shown in Figure 4.10. This ellipse has an area identical to that of the circle inscribing the square shown above, 6.25 square units.

4.3.3 Trapezoid

Both cases of the square and the parallelogram can be seen as affine equivalences of each other; their solutions will only differ by the scaling of the edge lengths between the two shapes, due to the preservation of the line at infinity when transforming a square into a rombus, or any arbitrary parallelogram. However, the remaining convex quadrangles will either be a trapezoid, or some general convex quadrangle without symmetry present; kites may be treated as a general convex quadrangle, as their

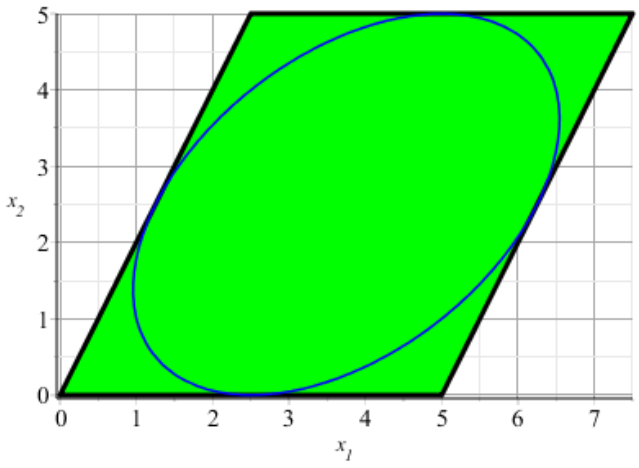


Figure 4.10: The maximum area inscribing ellipse within the rhombus.

geometry supports no special or unique properties of the maximum area inscribing ellipse whereby its properties may be predicted.

Figure 4.11 describes the trapezoid ABCD with vertices, A(5,5), B(9,5), C(8,7), and D(7,7). In order to parameterise the problem in terms of a_x, the trapezoid must then be translated to lie incident with the x_1 axis. Figure 4.11 also shows the translated trapezoid such that point A lies incident with the coordinate origin, while line AB lies incident with the x_1 axis.

Equations (4.47)-(4.50) are the line equations that define each edge of the trapezoid for this example:

$$\text{AB} = X_2; \tag{4.47}$$

$$\text{BC} = 2X_0 - \frac{1}{2}X_1 - \frac{1}{4}X_2; \tag{4.48}$$

$$\text{CD} = \frac{1}{2}X_0 - \frac{1}{4}X_2; \tag{4.49}$$

$$\text{DA} = \frac{1}{2}X_1 - \frac{1}{2}X_2. \tag{4.50}$$

Equations (4.51) through (4.54) showcase the equations used to identify the line

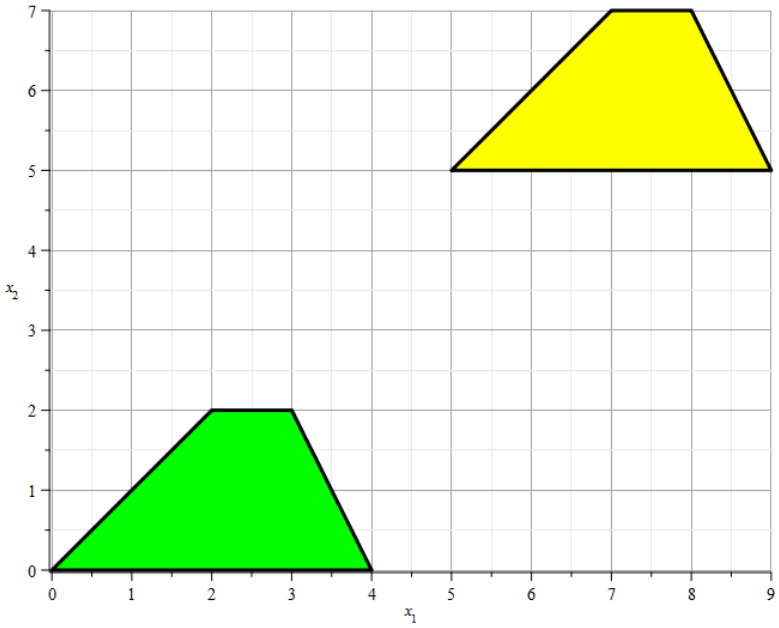


Figure 4.11: Translated trapezoid used to parametrise the pencil of inscribing ellipses.

conic shape coefficients, in addition to the x_1 axis, parameterised with a_x:

$$A_{22}; \tag{4.51}$$

$$4A_{00} - 2A_{01} - A_{02} + \tfrac{1}{4}A_{11} + \tfrac{1}{4}A_{12} + \tfrac{1}{16}A_{22}; \tag{4.52}$$

$$\tfrac{1}{4}A_{00} - \tfrac{1}{4}A_{02} + \tfrac{1}{16}A_{22}; \tag{4.53}$$

$$\tfrac{1}{4}A_{11} - \tfrac{1}{2}A_{12} + \tfrac{1}{4}A_{22}. \tag{4.54}$$

Grassmanian expansion is then performed on the matrix in Equation (4.55), in order to identify the line conic shape coefficients.

$$\text{Grassmanian}(\mathbf{A}_{\text{L}}) = \begin{bmatrix} A_{00} & A_{01} & A_{02} & A_{11} & A_{12} & A_{22} \\ 0 & 0 & a_x & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 16 \\ 4 & -2 & -1 & \frac{1}{4} & \frac{1}{4} & \frac{1}{16} \\ \frac{1}{4} & 0 & -\frac{1}{4} & 0 & 0 & \frac{1}{16} \\ 0 & 0 & 0 & \frac{1}{4} & -\frac{1}{2} & \frac{1}{4} \end{bmatrix}. \tag{4.55}$$

After computing the adjoint of the line conic shape coefficient matrix, matrix $\mathbf{A}$, the point conic shape coefficient matrix, Equation (4.56), is obtained:

$$\mathbf{A} = \begin{bmatrix} a_x^2 & -a_x & -\frac{3}{8}a_x^2 + \frac{1}{2}a_x \\ -a_x & 1 & \frac{5}{8}a_x - \frac{3}{2} \\ -\frac{3}{8}a_x^2 + \frac{1}{2}a_x & \frac{5}{8}a_x - \frac{3}{2} & \frac{9}{64}a_x^2 - \frac{7}{8}a_x + \frac{9}{4} \end{bmatrix}. \tag{4.56}$$

Upon pre and post multiplication of the point conic shape coefficient matrix with a homogeneous point triplet, the parameterised point conic coefficient is obtained as Equation (4.57):

$$k := a_x^2x_0^2 - 2a_xx_1x_0 - (-\frac{3}{4}a_x^2 + a_x)x_2x_0 + x_1^2 + (\frac{5}{4}a_x - 3)x_1x_2 + x_2^2(\frac{9}{64}a_x^2 - \frac{7}{8}a_x + \frac{9}{4}) = 0. \tag{4.57}$$

Figure 4.12 illustrates members of the pencil of inscribing ellipses. Although the parallelogram and square shared a uniform affine equivalence, this is not true for the

case of the trapezoid. However, due to the preservation of parallelism of the top and bottom faces of the trapezoid, the maximum area inscribing ellipse within the trapezoid will have its tangent points located at the exact center of each of these edges, as a result of the harmonic sequence of points generated by their placement. However, this is not the case for the two non-parallel edges.

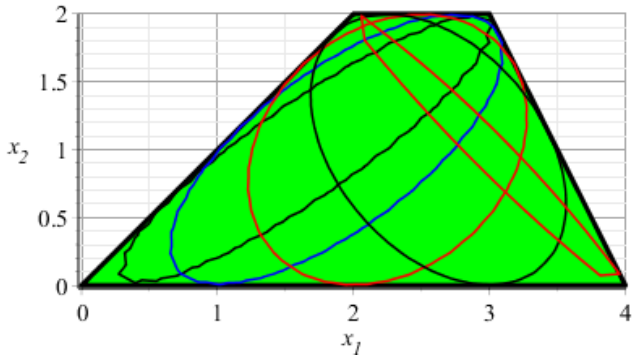


Figure 4.12: Example functions from the pencils of ellipses inscribing the trapezoid.

In order to solve for the maximum area inscribing ellipse within the trapezoid, the first derivative of the area function is used to identify the local maximum area with respect to the pole point, a_x. Equation (4.58) describes the first derivative area function of the trapezoid:

$$\frac{\partial \alpha}{\partial a_x} = -\frac{1}{2}\frac{(a_x - 4)a_x\pi}{\sqrt{-a_x(a_x - 4)}}. \tag{4.58}$$

Naturally, the predicted maximum area conic section inscribing the trapezoid occurs at a value of $a_x = 2$, the mid point of the line lying incident with the x_1 axis. Interestingly enough, the predicted maximum area inscribing ellipse in this case, has an area of exactly pi units squared. Upon substitution of $a_x = 2$ into Equation (4.57), Equation (4.59), the maximum area inscribing ellipse, is obtained:

$$k_{(max)} := 4 - 4x_1 - x_2 + x_1^2 - \frac{1}{2}x_1x_2 + \frac{17}{16}x_2^2 = 0. \tag{4.59}$$

In graphical form, the maximum area inscribing ellipse of this trapezoid is included in Figure 4.13.

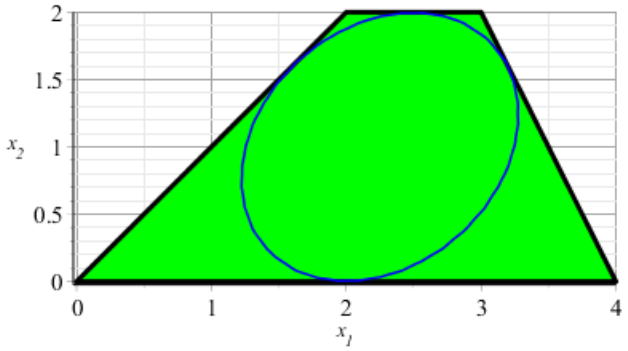


Figure 4.13: The maximum area inscribing ellipse within the example trapezoid.

4.3.4 General Convex Quadrangle

Determining the equation for the maximum area inscribing ellipse within the square, parallelogram, and trapezoid can be viewed as simplified cases of the overall problem of determining the equation of the maximum area inscribing ellipse within an arbitrary asymmetric convex quadrangle. Primarily, the square and parallelogram examples are simplified due to the affine equivalence between the unit square and parallelogram, while the trapezoid case is simplified through preservation of the cross ratios of the pole points of the maximum area inscribing ellipse on each parallel side. For this example, a general asymmetric convex quadrangle possessing none of the geometric constraints required to simplify its maximum area inscribing ellipse is presented. The quadrangle ABCD is illustrated in Figure 4.14, is comprised of vertices A(5,5), B(9,7), C(13,12), and D(6,8). Moreover, this convex quadrangle was selected so as to ensure that it is asymmetric.

Due to the translated and rotated nature of this initial quadrangle, a geometric translation is used to locate point A at the origin of the coordinate system, while the

rotation ensures that edge AB lies incident with the x_1 axis. Figure 4.14 illustrates the prescribed and translated quadrangle, ABCD, which now lies in the correct position and orientation for parameterisation with its pole point, a_x.

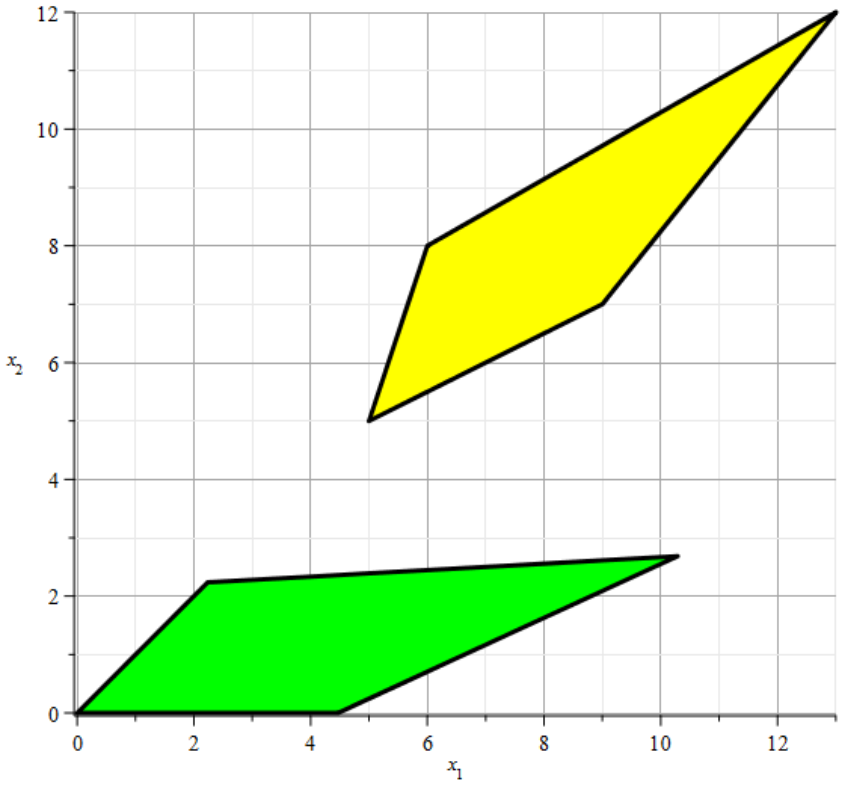


Figure 4.14: Prescribed and transformed convex quadrilaterals used for the parameterisation.

Equations (4.60)-(4.63) are the line equations of the edges of the transformed quadrangle:

$$\text{AB} = X_2 = 0; \tag{4.60}$$

$$\text{BC} = \frac{6}{\sqrt{5}}X_0 - \frac{3}{5}X_1 + \frac{13}{10}X_2 = 0; \tag{4.61}$$

$$\text{CD} = \frac{17}{2\sqrt{5}}X_0 + \frac{1}{10}X_1 - \frac{15}{10}X_2 = 0; \tag{4.62}$$

$$\text{DA} = \frac{1}{2}X_1 - \frac{1}{2}X_2 = 0. \tag{4.63}$$

Comparing Equations (4.64) through (4.67)with those of the earlier examples, it is to be seen that the earlier line equations contained no expressions with more than three terms. However, this is not the case for the general convex quadrangle:

$$20A_{22}; \tag{4.64}$$

$$144A_{00} - \tfrac{144}{5}\sqrt{5}A_{01} + \tfrac{312}{5}\sqrt{5}A_{02} + \tfrac{36}{5}A_{11} - \tfrac{156}{5}A_{12} + \tfrac{169}{5}A_{22}; \tag{4.65}$$

$$289A_{00} + \tfrac{34}{5}\sqrt{5}A_{01} - \tfrac{612}{5}\sqrt{5}A_{02} + \tfrac{1}{5}A_{11} - \tfrac{36}{5}A_{12} + \tfrac{324}{5}A_{22}; \tag{4.66}$$

$$5A_{11} - 10A_{12} - 5A_{22}. \tag{4.67}$$

Equation (4.68) is the matrix resulting from substituting the bounding line conics, as well as the parameterised pole point equation, into Equation (4.13).

$$\text{Grassmanian}(\mathbf{A}_{\text{L}}) = \begin{bmatrix} A_{00} & A_{01} & A_{02} & A_{11} & A_{12} & A_{22} \\ 0 & 0 & a_x & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 20 \\ 144 & -\frac{144}{5}\sqrt{5} & \frac{312}{5}\sqrt{5} & \frac{36}{5} & -\frac{156}{5} & \frac{169}{5} \\ 289 & \frac{34}{5}\sqrt{5} & -\frac{612}{5}\sqrt{5} & \frac{1}{5} & -\frac{36}{5} & \frac{324}{5} \\ 0 & 0 & 0 & 5 & -10 & 5 \end{bmatrix}. \tag{4.68}$$

Upon Grassmanian expansion of the matrix in Equation (4.68), the line conic shape coefficient matrix is obtained, see Equation (4.69). While the maximum order of each line conic shape coefficient within Equation (4.69) is limited to 1, it is important to note that the coefficient A_{00} is no longer independent of the parameterised pole

point, a_x.

$$\mathbf{A}_L = \begin{bmatrix} -1550400 - 31008\sqrt{5}a_x & 387600a_x - 3565920\sqrt{5} & -930240\sqrt{5} \\ 387600a_x - 3565920\sqrt{5} & -1860480\sqrt{5}a_x & -930240\sqrt{5}a_x \\ -930240\sqrt{5} & -930240\sqrt{5}a_x & 0 \end{bmatrix} \tag{4.69}$$

Equation (4.70) is the adjoint of the matrix in Equation (4.69). Unlike previous cases, entries a_{02}, a_{12}, and a_{22} are dependent upon a_x^2 for the general convex quadrangle, indicating that the pencil of conics inscribing the convex quadrangle has a highly nonlinear behaviour with respect to its pole point.

$$\mathbf{A} = \begin{bmatrix} a_x^2 & -a_x & \frac{1}{12}\sqrt{5}a_x^2 - \frac{11}{6}a_x \\ -a_x & 1 & \frac{1}{30}a_x^2 + \frac{5}{12}\sqrt{5}a_x - \frac{23}{6} \\ \frac{1}{12}\sqrt{5}a_x^2 - \frac{11}{6}a_x & \frac{1}{30}a_x^2 + \frac{5}{12}\sqrt{5}a_x - \frac{23}{6} & -\frac{23}{720}a_x^2 - \frac{47}{36}\sqrt{5}a_x - \frac{529}{36} \end{bmatrix}. \tag{4.70}$$

Pre and post multiplication of Equation (4.70) with the homogeneous point triplet yields Equation (4.71), the pencil of point conic sections which, while a_x varies along the x_1 axis within the vertices of the quadrangle, will always generate a closed conic section:

$$\begin{aligned} k := a_x^2x_0^2 - 2a_xx_1x_0 + (\frac{1}{6}\sqrt{5}a_x^2 - \frac{11}{3}a_x)x_2x_0 + x_1^2 + \\ (\frac{1}{15}a_x^2 + \frac{5}{6}a_x - \frac{23}{3})x_2x_1 + x_2^2(-\frac{23}{720}a_x^2 - \frac{47}{36}\sqrt{5}a_x - \frac{529}{36}) = 0. \end{aligned} \tag{4.71}$$

Figure 4.15 shows that varying the pole point, a_x along the x_1 axis within the bounds of the quadrangle forces Equation (4.71) to generate closed conic sections within the bounds of the quadrangle.

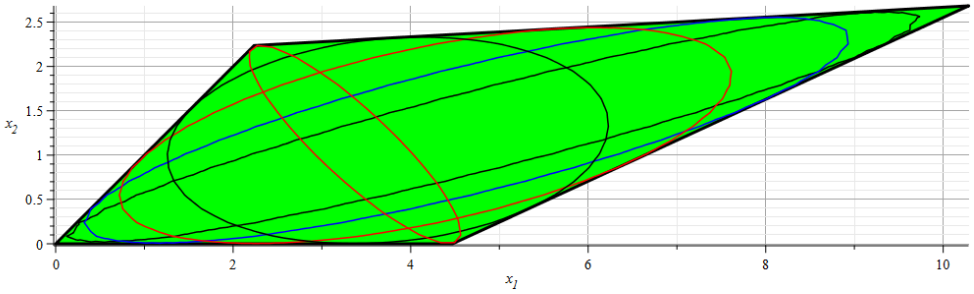


Figure 4.15: Examples from the pencil of ellipses inscribing the convex quadrangle.

The explicit first derivative of Equation (4.71) with respect to a_x is too cumbersome to reproduce here. However, Equation (4.72) is the area function itself. Appendix A lists the Maple code used to the generate this example and the expression for the first derivative for this example is to be found there.

$$\alpha = \frac{30\pi a_x(a_x^4 + 30\sqrt{5}a_x^3 + 785a_x^2 - 5100\sqrt{5}a_x + 28900)}{(a_x^3 + 25\sqrt{5}a_x^2 + 580a_x - 1700\sqrt{5})\sqrt{-a_x(a_x^3 + 25\sqrt{5}a_x^2 + 580a_x - 1700\sqrt{5})}}. \tag{4.72}$$

Unlike previous cases, this pole point position does not coincide with the mid point of the line lying incident with the x_1 axis, at a point of approximately 1.9933 units, thus any projective transformation between the quadrangle and unit square would ultimately fail in mapping the maximum area circle inscribing the square to that of the ellipse within the convex quadrangle. Specifically, the maximum area inscribing ellipse of this general asymmetric convex quadrangle is,

$$k_{max} := (\frac{368}{15}\sqrt{5} - \frac{4}{15}\sqrt{5}\sqrt{9739})x_1 + (\frac{42478}{135}\sqrt{5} - \frac{434}{135}\sqrt{5}\sqrt{9739})x_2 + (\frac{33137}{675} - \frac{361}{675}\sqrt{9739})x_1x_2 + x_1^2 + (\frac{87239}{2025} - \frac{1409}{4050}\sqrt{9739})x_2^2 + \frac{72812}{45} - \frac{736}{45}\sqrt{5} = 0. \tag{4.73}$$

Equation (4.73), obtained through maximisation of Equation (4.72) is shown in Figure 4.16 alongside the asymmetric quadrangle used to define it.

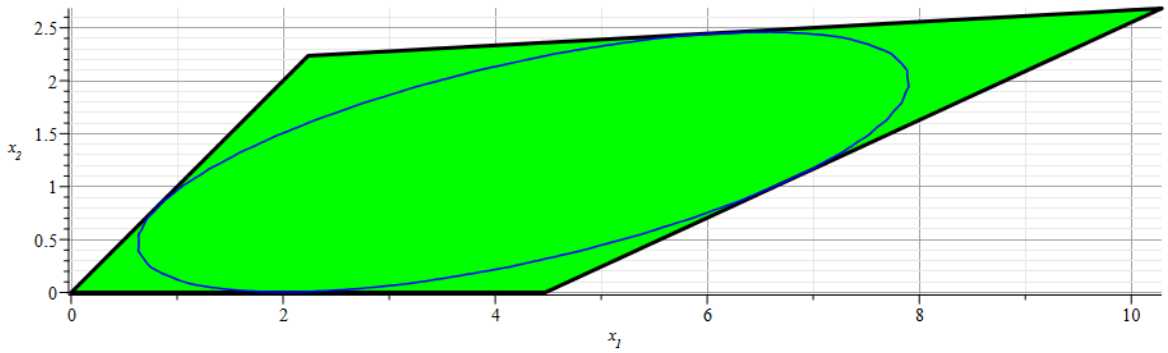


Figure 4.16: The maximum area ellipse inscribing the example asymmetric convex quadrangle.

4.3.5 Applicability and Suitability of the Solution

Through implementation of the projective extension of the Euclidean plane, it is possible to generate a pencil of conics lying within any given quadrangle parameterised by the location of a pole point, a_x, on one of its edges. Through this pencil of conics, it is possible to determine the maximum area ellipse contained within an arbitrary convex quadrangle using the first derivative of the area function with respect to the pole point, a_x, corresponding to the point conic shape coefficient matrix $\mathbf{A}$. In order to facilitate this parameterisation, the quadrangle is first transformed so as to force one edge to be incident with the x_1 axis. While this step alters the position and

orientation of the ellipse generated through the procedure, it does not alter its shape in any way. The resulting ellipse is transformed back to the original quadrangle using the inverse transformation matrix.

Aside from the transformation of the quadrangle in question, the aforementioned procedure functions in a completely general fashion. Given any parallelogram, the procedure presented within this chapter identifies an ellipse which possesses a pole point at the mid point of each side, as well as preserving the ratio of areas between these shapes; given edge lengths, it will always generate an ellipse lying inside the parallelogram with the same area as the circle within the correspondingly dimensioned square. Given a trapezoid, the same code will generate an ellipse which possesses a tangent point at the mid point of each parallel edge, preserving the harmonic ratio between four appropriate points on each parallel line, namely the first vertex coupled with the second vertex, alongside the midpoint of the edge and the point at infinity, in that order. Furthermore, the ellipse generated by this code remains metric, and thus its size is immediately calculable, requiring no further metric transforms, or mappings to orthogonal bases once it has been identified.

Most importantly, however, is the fact that this code requires no differentiation between the possible cases; the same code will generate the maximum area inscribing ellipse regardless of the case it is presented with. In contrast to existing solutions, this process is streamlined, and completely general.

Chapter 5

Conclusions

5.1 Summary of Previous Solutions

While it is well accepted that there exists a maximum area inscribing ellipse within a convex quadrangle [7], only two solution methodologies for determining the exact equation of this ellipse have been presented previously; one relies upon the implementation of a projective mapping which applies onto to special quadrangles, while the other relies on the derivation of a non-orthogonal coordinate system to generate the equation of the maximum area inscribing ellipse within the same non-orthogonal coordinate system.

Using a projective mapping to treat the unit square and its maximum area inscribing ellipse as an image of a quadrangle possessing two or four parallel edges, and its maximum area inscribing ellipse in turn, facilitates an accurate presentation of this maximum area inscribing ellipse, however, it relies on specific configurations of quadrangles in order to be successful. Provided that the quadrangle is either a parallelogram or a trapezoid, this projective mapping will successfully identify the maximum area inscribing ellipse. However, if the quadrangle is not a parallelogram or a trapezoid, the polar points of the ellipse are no longer preserved through the

preservation of a harmonic sequence of points on the parallel edges, and this projective transformation ultimately fails to produce the maximum area inscribing ellipse of the quadrangle.

Through the employ of a non-orthogonal coordinate system generated by the specific geometry of any given asymmetric convex quadrangle, it is possible to compute the pencil of conics contained within the quadrangle parameterised by a single variable. This pencil of conics is then optimised through its squared area function derivative in order to supply the maximum area inscribing ellipse of the quadrangle. While mathematically robust, this procedure fails to supply a metric result at its outcome, and applying a metric transformation to this solution leads to unnecessary computational expense and inaccuracy.

5.2 Summary and Comparison of Newly Developed Solution

Upon the definition of four points, no three of which are collinear, the vertices of a convex quadrangle ABCD may be defined, and transformed such that one vertex is incident with the origin, while the other vertex on that same edge lies on the x_1 axis. Through use of the line coordinates that define its edges, alongside a polar point a_x, a pencil of inscribing line conics may be computed, provided that a_x remains between the quadrangle vertices which are incident with the x_1 axis. From this point, the matrix of line conic shape coefficients may be inverted through the computation of its adjoint matrix, yielding the pencil of inscribing point conics, after which the area derivative function is locally maximized and provides the equation for the maximum area inscribing ellipse, in terms of the polar point variable, a_x. The resulting ellipse must then be transformed back to the original quadrangle, which will affect neither

its area nor shape, only its position and orientation.

Within each of the previously proposed solutions, it is impossible to provide a metric solution which supplies the maximum area inscribing ellipse of a general asymmetric quadrangle. While projective transformations may be used to map ellipses which satisfy a harmonic sequence of points, this harmonic sequence is not preserved upon mapping a square to a general asymmetric convex quadrangle. Implementation of a non-orthogonal basis coordinate system ultimately leads to a solution which is not metric, and is not immediately useful for identifying maximum covariance, or maximum isotropic workspace, for example.

5.3 Trapezoid Behaviour - A Brief Note

Throughout the course of this investigation, the aforementioned Maple code was used in order to check that the process generated solutions which remained consistent with proven geometric truths. Specifically: the maximum area inscribing ellipse of a square is a circle which possesses polar points at the midpoint of each side; and through affine equivalence that the maximum area inscribing ellipse within a parallelogram also includes polar points at the midpoints of the sides of the parallelogram; and furthermore, that the area of said ellipse is the same as the circle inscribing the square of equivalent edge lengths.

While testing to ensure that the maximum area inscribing ellipse within the trapezoid lies tangent to the midpoints of the parallel sides, however, interesting behaviour was observed with respect to the value of the area described by this maximum area inscribing ellipse. Through a pure concatenation of coincidence, the test case used to initially prove that the Maple file would generate the proper ellipse within the trapezoid generated an ellipse with an area of exactly π units squared. Through

investigation of this phenomenon, it was found that the maximum area inscribing ellipse of any trapezoid in question was, at most, π scaled by a factor of one algebraic number, and one rational number.

For example, given an increase in the height of the initial trapezoid to a total of three units in height scales the maximum area inscribing ellipse within the trapezoid by a factor of $\frac{3}{2}$, shown in Figure 5.1.

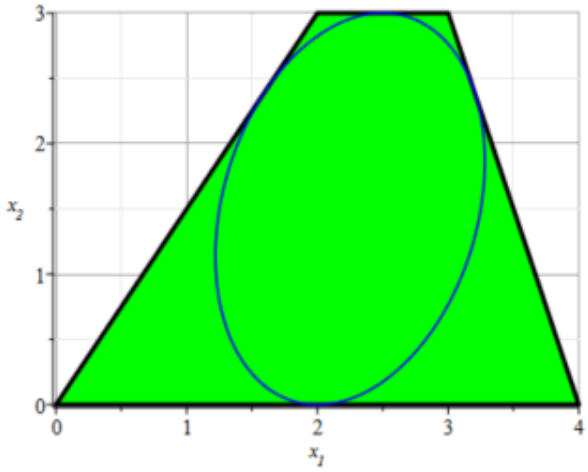


Figure 5.1: The maximum area inscribing ellipse within a trapezoid scaled by one unit vertically, with an area of exactly $\frac{3}{2}\pi$.

Additionally, scaling the top face of the initial trapezoid by one unit in length scales the maximum area inscribing ellipse within by a factor of $\sqrt{2}$, shown in Figure 5.2.

Furthermore, it was found upon iteration of test cases, that this scaling was predictable and consistent within the scope of the values through which the testing was conducted. Two separate test cases are examined briefly, so as to demonstrate the scalability of this solution. First, given a scaling of $\frac{3}{2}$ square units per unit increase in height, one would expect a trapezoid with a height of six units to possess a maximum area inscribing ellipse of 3π square units. Figure 5.3 shows this trapezoid,

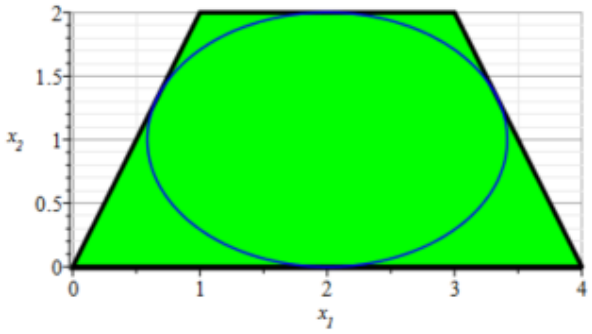


Figure 5.2: The maximum area inscribing ellipse within a trapezoid scaled by one unit horizontally on its top face, with an area of exactly $\sqrt{2}\pi$.

with its maximum area inscribing ellipse of 3π square units.

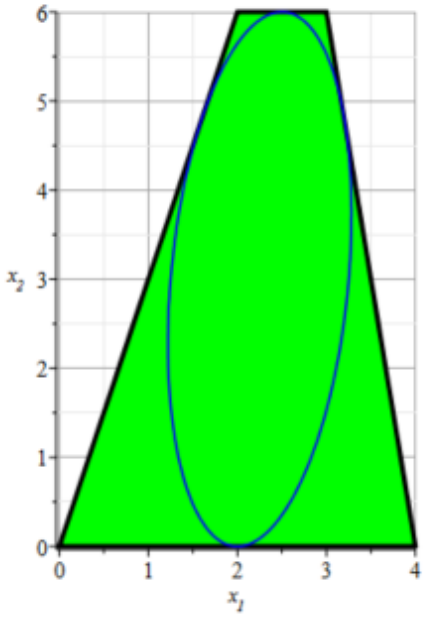


Figure 5.3: The maximum area inscribing ellipse within a trapezoid scaled to a total height of six units, with an area of exactly 3π.

Additionally, if this exact same trapezoid has its top side length increased in length by one unit, the resulting maximum area inscribing ellipse should possess an area scaled by $\sqrt{2}$. Figure 5.4 shows this trapezoid, and the scaled maximum area

inscribing ellipse of $3\sqrt{2}\pi$.

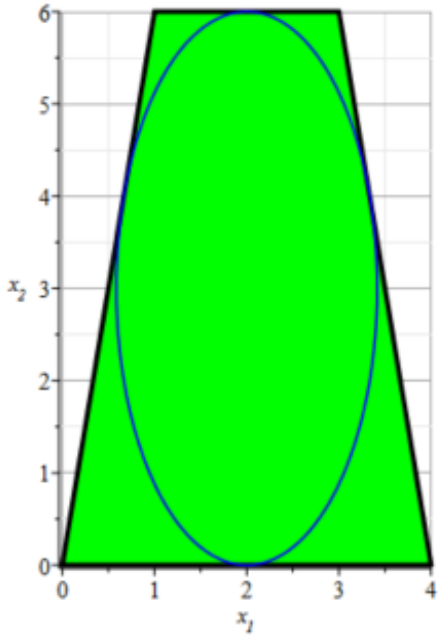


Figure 5.4: The maximum area inscribing ellipse within a trapezoid scaled to a total height of six units, with one additional unit added to the length of its top side, with an area of exactly $3\sqrt{2}\pi$.

Not only is the scaling in either direction consistent and predictable, but each axial scaling factor is independent of one another; scaling the the x_2 direction solely yields an increase in the fractional value present, while scaling in the x_1 direction only affects the preceding value under the square root. Specifically, an increase in height of one unit forces a scaling of the maximum area inscribing ellipse by $\frac{3}{2}$ square units, while an increase in the length of the top parallel side increases the area of the maximum area inscribing ellipse by $\sqrt{2}$ units. Increasing the height of the quadrangle will not affect the square root coefficient, and increasing the length of either parallel side will not affect the rational coefficient.

Although the exact cause of this correlation is not entirely known, it is possible

that the nature of the harmonic ratio between the parallel sides forces this algebraically direct scaling of solutions. Obviously, further investigation into the nature of these maximum area inscribing ellipses is necessary to uncover the true causality of this uncoupled correlation. Once this causality is uncovered, it is possible that an algorithm could be developed to automatically calculate the maximum area inscribing ellipse contained within a trapezoid while bypassing the majority of the computations required for the general solution.

References

[1] K Strubecker. *Einfuhrung in die hohere Mathematik: Grundlagen*. Oldenbourg, 1956.

[2] A Gfrerrer. The area maximizing inellipse of a convex quadrangle. private communication, December 19, 2002.

[3] T Yoshikawa. Manipulability of robotic mechanisms. *The international journal of Robotics Research*, 4(2):3–9, 1985.

[4] F Marquet and S Krut. archi, a redundant mechanism for machining with unlimited rotation capacities. In *Proc. of ICAR*, 2001.

[5] S Krut, F Pierrot, et al. Velocity performance indices for parallel mechanisms with actuation redundancy. *Robotica*, 22(02):129–139, 2004.

[6] WV Parker and JE Pryor. Polygons of greatest area inscribed in an ellipse. *The American Mathematical Monthly*, 51(4):205–209, 1944.

[7] A Horwitz. Finding ellipses and hyperbolas tangent to two, three, or four given lines. *Southwest Journal of Pure and Applied Mathematics*, 1(1), 2002.

[8] A Horwitz. The locus of centers of ellipses inscribed in quadrilaterals. Technical report, math/0312403, 2003.

[9] MJD Hayes, PJ Zsombor-Murray, and A Gfrerrer. Largest ellipse inscribing an arbitrary polygon. *Proceedings of the 19th Canadian Congress on Applied Mechanics (CANCAM 2003)*, 3:1, 2003.

[10] MJD Hayes. Maximum area ellipses inscribing specific quadrilaterals. In *Proceedings of The Canadian Society for Mechanical Engineering International Congress 2016*, 2016 CCToMM M3 Symposium. CCToMM, 2016.

[11] Euclid. *The thirteen books of Euclid's Elements*. Courier Corporation, 1956.

[12] D Hilbert and S Cohn-Vossen. *Geometry and the Imagination*, volume 87. American Mathematical Soc., 1952.

[13] DMY Sommerville. *The elements of non-Euclidean geometry*. Courier Corporation, 2012.

[14] HSM Coxeter. *Non-euclidean geometry*. Cambridge University Press, 1998.

[15] FC Klein et al. A comparative review of recent researches in geometry. *Bulletin of the New York Mathematical Society*, 2(10):215–249, 1893.

[16] FC Klein. *Vergleichende Betrachtungen über neuere geometrische Forschungen: Programm zum Eintritt in die philosophische Facultät und den Senat der k. Friedrich-Alexanders-Universität zu Erlangen*. Deichert, 1872.

[17] WT Fishback. *Projective and Euclidean geometry*. 1969.

[18] V Chari and AN Pressley. *A guide to quantum groups*. Cambridge university press, 1995.

[19] JJ Craig. *Introduction to robotics: mechanics and control*, volume 3. Pearson Prentice Hall Upper Saddle River, 2005.

[20] D Hilbert and B Sturmfels. *Theory of algebraic invariants*. Cambridge university press, 1993.

[21] FC Klein. Elementary mathematics from an advanced standpoint, geometry (1908). *Reprinted Dover, New York*, 1939.

[22] HSM Coxeter and WOJ Moser. *Generators and relations for discrete groups*, volume 14. Springer Science & Business Media, 2013.

[23] JW Rutter. *Geometry of curves*. CRC press, 2000.

[24] A Gfrerrer. Euclidean classifications of conic sections. private communication, August 12, 2016.

[25] DC Lay. *Linear algebra and its applications*. Pearson, 2005.

[26] HG Funkhouser. A short account of the history of symmetric functions of roots of equations. *The American mathematical monthly*, 37(7):357–365, 1930.

[27] G Glaeser, H Stachel, and B Odehnal. *The Universe of Conics: From the ancient Greeks to 21st century developments*. Springer, 2016.

[28] L Mirsky. *An introduction to linear algebra*. Courier Corporation, 2012.

Appendix A

Sample Algorithm, Maple Code

Maximum Area Inscribing Ellipse within a Convex Quadrilateral that requires Transformation to the x__1 axis

This code will automatically and without user intervention generate the maximum area inscribing ellipse contained within any convex quadrilateral.

A user will input their quadrilateral, as point coordinates of the four vertices, in the PointN places, and run the code.

Once the initial user defined quadrilateral has been transformed so as to place one edge in concurrence with the x__1 axis, the maximum area inscribing ellipse will be created by using the line coordinates of each edge of the quadrilateral to define the line coordinate ellipse shape matrix, whereby a point coordinate ellipse may be generated through Grassmanian expansion. From this point ellipse, a variable a__x will be used to search for the maximum area inscribing ellipse with one tangent, or polar point, along the x__1 axis.

Several interstitial mathematical steps are taken in an effort to showcase the effectiveness of the solution, most notably a step wherein several inscribing ellipses are generated in parametric form, and then plotted within the transformed quadrangle. This was taken solely to demonstrate that the pencil of ellipses translated properly between line and point coordinates, and furthermore, that the form of all of those ellipses is also contained completely within the transformed quadrilateral.

restart;

$with(LinearAlgebra) : with(linalg) : with(plots) : with(geometry) : with(tensor) :$

First, the matrix used for the geometric translation of the quadrangle will be defined. Many forms of this matrix exist, but in keeping with the European standard, the first element in the matrix will be used as the homogenizing coordinate, and the shape of the remaining matrix will be dependent upon that positioning.

$GeoTranMat := Matrix(3, 3, [1, 0, 0, a, \cos(\theta), -\sin(\theta), b, \sin(\theta), \cos(\theta)])$

$$\begin{bmatrix} 1 & 0 & 0 \\ a & \cos(\theta) & -\sin(\theta) \\ b & \sin(\theta) & \cos(\theta) \end{bmatrix} \quad \textbf{(1)}$$

$InverseTransMat := simplify(MatrixInverse(\textbf{(1)}))$

$$\begin{bmatrix} 1 & 0 & 0 \\ -a\cos(\theta) - \sin(\theta) b & \cos(\theta) & \sin(\theta) \\ a\sin(\theta) - \cos(\theta) b & -\sin(\theta) & \cos(\theta) \end{bmatrix} \quad \textbf{(2)}$$

Logic would follow that an individual need only transform the ellipse to the origin, as would be defined by the first transformation. However, in order to center the quadrangle on the origin, it is required instead to treat the initially defined quadrangle as the translated verson, and take the inverse transformation to locate it at the origin, with one side incident with the x__1 axis.

The variables a, b, and θ will be defined via the quadrilateral that is entered by the user. Unfortunately,

as it currently stands, the process will only function properly for a quadrilateral that is entered with the bottom left point first, and working around the points in a counter clockwise fashion. Additional programmatical steps will eventually be taken so as to eliminate this dependence upon entrance order. The variables a and b will be defined by the x__1 and x__2 coordinates of the initial point, P1, while angle theta is the inclination angle above parallel between P1 and P2.

For the purposes of this demonstration, an arbitrary convex quadrilateral was selected, which is presented below.

$Point1 := point(P1, [5, 5]) :$

$Point2 := point(P2, [9, 7]) :$

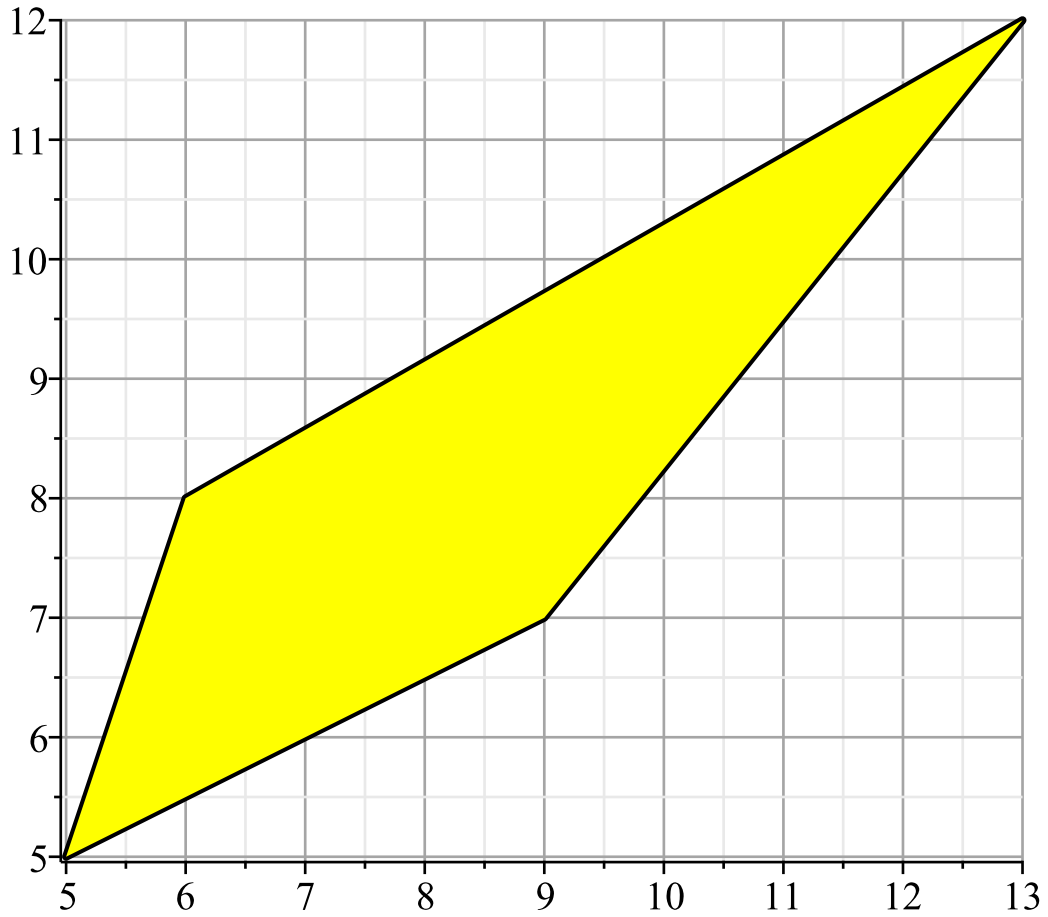
$Point3 := point(P3, [13, 12]) :$

$Point4 := point(P4, [6, 8]) :$

$Quadrilateral := [coordinates(P1), coordinates(P2), coordinates(P3), coordinates(P4)] :$

$Quadrilateral1 := polygonplot(Quadrilateral, scaling = constrained, colour = yellow, thickness = 3, axes = framed, gridlines = true) :$

$display(Quadrilateral1)$



Now that the quadrangle has been defined above, the variables a, b, and θ are defined below, with subsequent subsitution into the inverse tranformation matrix.

$a := coordinates(P1)[1]$

$$5 \quad \textbf{(3)}$$

$b := coordinates(P1)[2]$

$$5 \quad \textbf{(4)}$$

$$\theta := \arctan\left(\frac{(coordinates(P2)[2] - coordinates(P1)[2])}{(coordinates(P2)[1] - coordinates(P1)[1])} \right)$$

$$\arctan\left(\frac{1}{2} \right) \quad \textbf{(5)}$$

InverseTransMat

$$\begin{bmatrix} 1 & 0 & 0 \\ -3\sqrt{5} & \frac{2}{5}\sqrt{5} & \frac{1}{5}\sqrt{5} \\ -\sqrt{5} & -\frac{1}{5}\sqrt{5} & \frac{2}{5}\sqrt{5} \end{bmatrix} \quad \textbf{(6)}$$

From this archetype, any different point addition for P1 and P2 will cascade through and fundamentally alter the transformation matrix presented above. Everything was output solely to ensure functionality of the coding process. From here, the transformed points may be computed.

```
Point1VectorInput := Vector[column]([1, coordinates(P1)[1], coordinates(P1)[2]]) :
Point2VectorInput := Vector[column]([1, coordinates(P2)[1], coordinates(P2)[2]]) :
Point3VectorInput := Vector[column]([1, coordinates(P3)[1], coordinates(P3)[2]]) :
Point4VectorInput := Vector[column]([1, coordinates(P4)[1], coordinates(P4)[2]]) :
```

```
P1TO := multiply(InverseTransMat, Point1VectorInput)
```

$$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \quad \textbf{(7)}$$

```
P2TO := multiply(InverseTransMat, Point2VectorInput)
```

$$\begin{bmatrix} 1 & 2\sqrt{5} & 0 \end{bmatrix} \quad \textbf{(8)}$$

```
P3TO := multiply(InverseTransMat, Point3VectorInput)
```

$$\begin{bmatrix} 1 & \frac{23}{5}\sqrt{5} & \frac{6}{5}\sqrt{5} \end{bmatrix} \quad \textbf{(9)}$$

```
P4TO := multiply(InverseTransMat, Point4VectorInput)
```

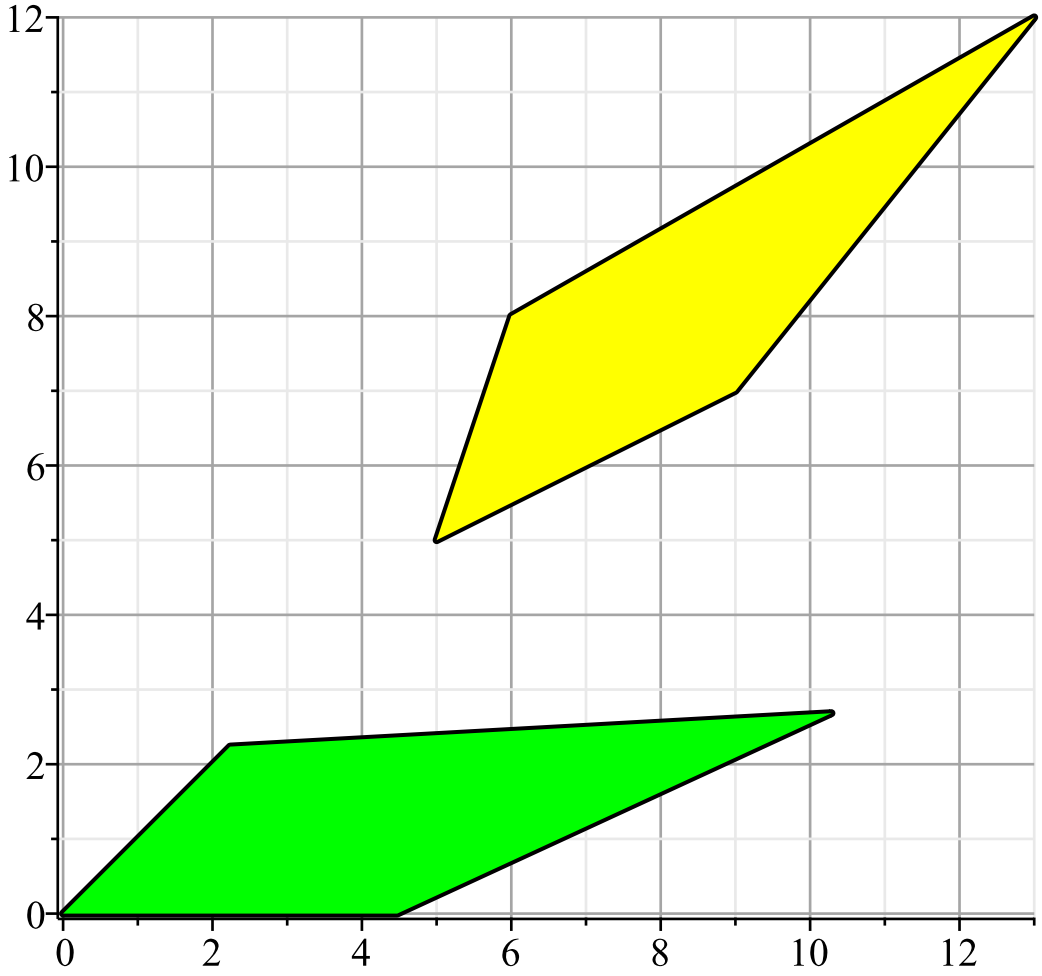
$$\begin{bmatrix} 1 & \sqrt{5} & \sqrt{5} \end{bmatrix} \quad \textbf{(10)}$$

Importantly, one should be able to look at the coordinates post transformation and know that the ellipse was translated properly by consideration of the following; the first point is exactly equal to one, as it is the homogenizing coordinate, while all other points will be strictly positive in nature. This ensures that the ellipse is contained solely within the first quadrant, subsequently removing difficulties within the area calculations that follow.

However, it is also possible to simply plot the quadrangle and view the fact that it is contained solely within the first quadrant, which is done below. Mainly because visual representations of mathematical concepts are easier to understand, and it shows that the transformation did, in fact, work as intended.

```
TranQuad := [[(7)[2], (7)[3]], [(8)[2], (8)[3]], [(8)[2], (8)[3]], [(9)[2], (9)[3]], [(10)[2],
    (10)[3]]] :
TranQuad1 := polygonplot(TranQuad, scaling = constrained, colour = green, thickness = 3, axes
    = framed, gridlines = true) :
```

```
display(TranQuad1, Quadrilateral1)
```



From here, the process of defining the pencil of line conics contained within the quadrangle is intiated.

First, we define XT, and XX such that X__0 is used to define the homogenizing coordinate,

$XT := Vector[row]([X_0, X_1, X_2])$

$$\begin{bmatrix} X_0 & X_1 & X_2 \end{bmatrix} \tag{11}$$

$XX := Vector[column]([X_0, X_1, X_2])$

$$\begin{bmatrix} X_0 \\ X_1 \\ X_2 \end{bmatrix} \tag{12}$$

Now, the line conic shape coefficient matrix will be defined below. Upon characterization of every bounding line segment, as well as the polar point, this matrix will be used in conjunction with the above defined vectors to solve for the line conic shape coefficients. Once this solution has been obtained, it will be used to define the area function of the ellipse, whereby a maximum inscribing ellipse may be found

from the pencil of conics defined in the previous step.

$$LineConicShapeCoefficientMatrix := Matrix\left(3, 3, \left[A_{00}, A_{01}, A_{02}, A_{01}, A_{11}, A_{12}, A_{02}, A_{12}, A_{22}\right]\right)$$

$$\begin{bmatrix} A_{00} & A_{01} & A_{02} \\ A_{01} & A_{11} & A_{12} \\ A_{02} & A_{12} & A_{22} \end{bmatrix} \tag{13}$$

Now, define the line conic function that will be used to characterize each of the above variables,

$$LineConicFunction := simplify(multiply(\mathbf{(11)}, \mathbf{(13)}, \mathbf{(12)}))$$

$$A_{00} X_0^2 + 2 A_{01} X_0 X_1 + 2 A_{02} X_0 X_2 + A_{11} X_1^2 + 2 A_{12} X_1 X_2 + A_{22} X_2^2 \tag{14}$$

Equation 15 represents the general expression for any ellipse, including the pencil of ellipses defined via utilization of line coordinates, namely the vectors XX and XT defined above it. From here, as we are looking for the pencil of conics defined by the previously stated ellipse, we will use the line coordinates that depict each side of the quadrangle to define this equation, and subsequently, the variables contained within the Line Conic Shape Coefficient Matrix.

This will be done via substitution of the points that define each of the coordinates after their transformation into the line coordinate definition matrix shown below.

$$LineCoordinateDefinitionMatrix := Matrix\left(3, 3, \left[X_0, X_1, X_2, 1, Pt1_{x_1}, Pt1_{x_2}, 1, Pt2_{x_1}, Pt2_{x_2}\right]\right)$$

$$\begin{bmatrix} X_0 & X_1 & X_2 \\ 1 & Pt1_{x1} & Pt1_{x2} \\ 1 & Pt2_{x1} & Pt2_{x2} \end{bmatrix} \tag{15}$$

Now, each line will be defined as a triple of homogeneous coordinates, with P1TO being the first point, and P2T0 being the second. Using the determinant of the above matrix, with those points, it will define each of the bounding line coordinates.

$$LineCoordinate1 := Determinant\left(subs\left(Pt1_{x_1} = P1TO[2], Pt1_{x_2} = P1TO[3], Pt2_{x_1} = P2TO[2], Pt2_{x_2} = P2TO[3], LineCoordinateDefinitionMatrix\right)\right)$$

$$2 X_2 \sqrt{5} \tag{16}$$

$$LineCoordinate2 := Determinant\left(subs\left(Pt1_{x_1} = P2TO[2], Pt1_{x_2} = P2TO[3], Pt2_{x_1} = P3TO[2], Pt2_{x_2} = P3TO[3], LineCoordinateDefinitionMatrix\right)\right)$$

$$12 X_0 - \frac{6}{5} X_1 \sqrt{5} + \frac{13}{5} X_2 \sqrt{5} \tag{17}$$

$$LineCoordinate3 := Determinant\left(subs\left(Pt1_{x_1} = P3TO[2], Pt1_{x_2} = P3TO[3], Pt2_{x_1} = P4TO[2], Pt2_{x_2} = P4TO[3], LineCoordinateDefinitionMatrix\right)\right)$$

$$17\,X_0 + \frac{1}{5}\,X_1\sqrt{5} - \frac{18}{5}\,X_2\sqrt{5} \tag{18}$$

$LineCoordinate4 := Determinant\left(subs\left(Pt1_{x_1} = P4TO[2],\ Pt1_{x_2} = P4TO[3],\ Pt2_{x_1} = P1TO[2],\ Pt2_{x_2} = P1TO[3],\ LineCoordinateDefinitionMatrix\right)\right)$

$$X_1\sqrt{5} - X_2\sqrt{5} \tag{19}$$

$$X_1\sqrt{5} - X_2\sqrt{5} \tag{20}$$

The above coordinates are the line coordinates that define each of the bounding edges of the quadrilateral. While the coefficients are represented via placing the square roots second in order, the coefficient operation still extracts the full coefficient term, including the square roots. Seems to simply be a representational quirk from Maple.

From here, it is important to define an additional line that will contain the polar pointmof the pencil of conics, a__x, namely the x__1 axis.

$GT := Vector[row]\left(\left[G_0, G_1, G_2\right]\right)$

$$\begin{bmatrix} G_0 & G_1 & G_2 \end{bmatrix} \tag{21}$$

$GG := Vector[column]\left(\left[G_0, G_1, G_2\right]\right)$

$$\begin{bmatrix} G_0 \\ G_1 \\ G_2 \end{bmatrix} \tag{22}$$

In order to define the equation that will inevitably lead to the definition of the polar point in terms of the line conic shape variables, the vector G must be concatenated with the line conic shape coefficient matrix,

$PolarPointShapeCoefficientVector := multiply(\mathbf{(13)}, \mathbf{(22)})$

$$\begin{bmatrix} A_{00}\,G_0 + A_{01}\,G_1 + A_{02}\,G_2 & A_{01}\,G_0 + A_{11}\,G_1 + A_{12}\,G_2 & A_{02}\,G_0 + A_{12}\,G_1 + A_{22}\,G_2 \end{bmatrix} \tag{23}$$

$SimplifiedPointShapeVector := simplify\left(\frac{\mathbf{(23)}[2]}{\mathbf{(23)}[1]}\right)$

$$\frac{A_{01}\,G_0 + A_{11}\,G_1 + A_{12}\,G_2}{A_{00}\,G_0 + A_{01}\,G_1 + A_{02}\,G_2} \tag{24}$$

Because the only term that has an effect on the definition of the line coordinate which includes the solution point a__x is contained on the x__1 axis, Polar Point Shape Coefficient Vector which defines this point can be simplified to contain solely the second component, and due to homogeneity, this component will be divided by the first term in the initial polar point shape coefficient vector. From here, it is required to define the polar point, a__x, in terms of its line coordinates.

$PolarLineCoordinateVector := subs\left(G_0 = 0,\ G_1 = 0,\ G_2 = 1,\ \mathbf{(24)} = a_x\right)$

$$\frac{A_{12}}{A_{02}} = a_x \tag{25}$$

From the polar line coordinate defined above, the normal depiction of the equation is required, which then places the polar point in terms of the line coordinate shape coefficients, all while equating the entirety of the expression to zero.

$PolarPointShapeVector := normal(\mathbf{(26)})$

$$A_{02}\, a_x - A_{12} \tag{26}$$

Now, in order to solve for the shape coefficients and thus define the pencil of ellipses that inscribe the quadrilateral, Grassmanian expansion utilizing the solved line conic shape coefficient ellipse definitions will be performed. In order to facillitate this expansion, each of the line conic definition functions must be determined,

Substituting the line coordinates for each side of the bounding quadrilateral into the previously defined line conic conic shape coefficient equation, it is possible to create a matrix of equations which can be used to solve for the line conic shape coefficients,

$LineConic1 := subs\left(X_0 = coeff\left(\mathbf{(16)}, X_0\right), X_1 = coeff\left(\mathbf{(16)}, X_1\right), X_2 = coeff\left(\mathbf{(16)}, X_2\right), \mathbf{(14)}\right)$

$$20\, A_{22} \tag{27}$$

$LineConic2 := subs\left(X_0 = coeff\left(\mathbf{(17)}, X_0\right), X_1 = coeff\left(\mathbf{(17)}, X_1\right), X_2 = coeff\left(\mathbf{(17)}, X_2\right), \mathbf{(14)}\right)$

$$\frac{36}{5} A_{11} - \frac{156}{5} A_{12} + \frac{169}{5} A_{22} - \frac{144}{5} A_{01}\sqrt{5} + \frac{312}{5} A_{02}\sqrt{5} + 144\, A_{00} \tag{28}$$

$LineConic3 := subs\left(X_0 = coeff\left(\mathbf{(18)}, X_0\right), X_1 = coeff\left(\mathbf{(18)}, X_1\right), X_2 = coeff\left(\mathbf{(18)}, X_2\right), \mathbf{(14)}\right)$

$$\frac{1}{5} A_{11} - \frac{36}{5} A_{12} + \frac{324}{5} A_{22} + \frac{34}{5} A_{01}\sqrt{5} - \frac{612}{5} A_{02}\sqrt{5} + 289\, A_{00} \tag{29}$$

$LineConic3 := subs\left(X_0 = coeff\left(\mathbf{(19)}, X_0\right), X_1 = coeff\left(\mathbf{(19)}, X_1\right), X_2 = coeff\left(\mathbf{(19)}, X_2\right), \mathbf{(14)}\right)$

$$5\, A_{11} - 10\, A_{12} + 5\, A_{22} \tag{30}$$

Now that the bounding lines have been expressed in terms of functions that involve the line conic shape coefficients, and alongside the previously shown polar point conic shape variable equation, there are five separate equations. A unique conic can be defined once five points on that conic are known, so these five equations will be used in order to completely characterize the conic shape coefficients, Aij.

From here, it is possible to create a matrix such that the equations can be solved for the conic shape variables through Grassmanian expansion. Each column of the matrix will contain one of the line conic shape coefficient variables, Aij, while each row will contain one of the previously defined equations.

$GrassmanianConicMatrix := Matrix\Big(6, 6, \big[A_{00}, A_{01}, A_{02}, A_{11}, A_{12}, A_{22}, coeff\left(\mathbf{(26)}, A_{00}\right), coeff\left(\mathbf{(26)}, A_{01}\right), coeff\left(\mathbf{(26)}, A_{02}\right), coeff\left(\mathbf{(26)}, A_{11}\right), coeff\left(\mathbf{(26)}, A_{12}\right), coeff\left(\mathbf{(26)}, A_{22}\right), coeff\left(\mathbf{(27)}, A_{00}\right), coeff\left(\mathbf{(27)}, A_{01}\right), coeff\left(\mathbf{(27)}, A_{02}\right), coeff\left(\mathbf{(27)}, A_{11}\right), coeff\left(\mathbf{(27)}, A_{12}\right), coeff\left(\mathbf{(27)}, A_{22}\right), coeff\left(\mathbf{(28)}, A_{00}\right), coeff\left(\mathbf{(28)}, A_{01}\right), coeff\left(\mathbf{(28)}, A_{02}\right), coeff\left(\mathbf{(28)}, A_{11}\right), coeff\left(\mathbf{(28)}, A_{12}\right), coeff\left(\mathbf{(28)}, A_{22}\right), coeff\left(\mathbf{(29)}, A_{00}\right), coeff\left(\mathbf{(29)}, A_{01}\right), coeff\left(\mathbf{(29)}, A_{02}\right), coeff\left(\mathbf{(29)}, A_{11}\right), coeff\left(\mathbf{(29)}, A_{12}\right), coeff\left(\mathbf{(29)}, A_{22}\right), coeff\left(\mathbf{(30)}, A_{00}\right), coeff\left(\mathbf{(30)}, A_{01}\right), coeff\left(\mathbf{(30)}, A_{02}\right), coeff\left(\mathbf{(30)}, A_{11}\right), coeff\left(\mathbf{(30)}, A_{12}\right), coeff\left(\mathbf{(30)}, A_{22}\right)\big]\Big)$

$$\begin{bmatrix} A_{00} & A_{01} & A_{02} & A_{11} & A_{12} & A_{22} \\ 0 & 0 & a_x & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 20 \\ 144 & -\frac{144}{5}\sqrt{5} & \frac{312}{5}\sqrt{5} & \frac{36}{5} & -\frac{156}{5} & \frac{169}{5} \\ 289 & \frac{34}{5}\sqrt{5} & -\frac{612}{5}\sqrt{5} & \frac{1}{5} & -\frac{36}{5} & \frac{324}{5} \\ 0 & 0 & 0 & 5 & -10 & 5 \end{bmatrix} \quad \textbf{(31)}$$

Now, from here, each first row minor of the matrix will determine the value corresponding to the variable in the eliminated column of the matrix. In order to increase the brevity of the equations, it is advantageous to collect the resultant equations in terms of the desired variable, which is in this case, a__x.

Also, to avoid overwriting each of the variable values, we will use the matrix Bij as a place holder while conducting the Grassmanian expansion.

$B_{00} := collect(det(minor(\textbf{(31)}, 1, 1)), a_x)$

$$-1550400 - 31008\sqrt{5}\, a_x \quad \textbf{(32)}$$

$B_{01} := collect(-det(minor(\textbf{(31)}, 1, 2)), a_x)$

$$387600\, a_x - 3565920\sqrt{5} \quad \textbf{(33)}$$

$B_{02} := collect(det(minor(\textbf{(31)}, 1, 3)), a_x)$

$$-930240\sqrt{5} \quad \textbf{(34)}$$

$B_{11} := collect(-det(minor(\textbf{(31)}, 1, 4)), a_x)$

$$-1860480\sqrt{5}\, a_x \quad \textbf{(35)}$$

$B_{12} := collect(det(minor(\textbf{(31)}, 1, 5)), a_x)$

$$-930240\sqrt{5}\, a_x \quad \textbf{(36)}$$

$B_{22} := collect(-det(minor(\textbf{(31)}, 1, 6)), a_x)$

$$0 \quad \textbf{(37)}$$

Now that the Grassmanian expansion has been performed, all of the line conic shape coefficients have been determined, and are collected below,

$LineConicShapeMatrixFinal := subs(A_{00} = B_{00}, A_{01} = B_{01}, A_{02} = B_{02}, A_{11} = B_{11}, A_{12} = B_{12}, A_{22} = B_{22},$ **(13)**$)$

$$\begin{bmatrix} -1550400 - 31008\sqrt{5}\,a_x & 387600\,a_x - 3565920\sqrt{5} & -930240\sqrt{5} \\ 387600\,a_x - 3565920\sqrt{5} & -1860480\sqrt{5}\,a_x & -930240\sqrt{5}\,a_x \\ -930240\sqrt{5} & -930240\sqrt{5}\,a_x & 0 \end{bmatrix} \tag{38}$$

Problematically, the line conic shape matrix does not behave in such a way so as to have any meaningful capacity with respect to determining the area of the conic section itself. This, in turn, means that the line conic must be transformed into a point conic shape function in order to extract this data.

$PointConicCoefficientMatrix := Matrix(3, 3, [det(minor(\mathbf{(38)}, 1, 1)), -det(minor(\mathbf{(38)}, 1, 2)), det(minor(\mathbf{(38)}, 1, 3)), -det(minor(\mathbf{(38)}, 2, 1)), det(minor(\mathbf{(38)}, 2, 2)), -det(minor(\mathbf{(38)}, 2, 3)), det(minor(\mathbf{(38)}, 3, 1)), -det(minor(\mathbf{(38)}, 3, 2)), det(minor(\mathbf{(38)}, 3, 3))])$

$$\begin{bmatrix} \begin{bmatrix} -4326732288000\,a_x^2, 4326732288000\,a_x, 7932342528000\,a_x - 360561024000\sqrt{5}\,a_x^2 \end{bmatrix}, \\ \begin{bmatrix} 4326732288000\,a_x, -4326732288000, -144224409600\,a_x^2 + 16585807104000 \\ -1802805120000\sqrt{5}\,a_x \end{bmatrix}, \\ \begin{bmatrix} 7932342528000\,a_x - 360561024000\sqrt{5}\,a_x^2, -144224409600\,a_x^2 + 16585807104000 \\ -1802805120000\sqrt{5}\,a_x, 138215059200\,a_x^2 - 63578927232000 + 5648789376000\sqrt{5}\,a_x \end{bmatrix} \end{bmatrix} \tag{39}$$

For computational efficiency and ease of expression, it is advantageous to represent the above matrix so that the initial coefficient on the a__x^2 term in the [1,1] position is equal to one. This is possible due to the fact that that any scalar multiple of a given function is always exactly equal to that function.

$$\frac{\mathbf{(39)}}{coeff\left(\mathbf{(39)}[1,1], a_x^2\right)}$$

$$\begin{bmatrix} a_x^2 & -a_x & -\frac{11}{6}\,a_x + \frac{1}{12}\sqrt{5}\,a_x^2 \\ -a_x & 1 & \frac{1}{30}\,a_x^2 - \frac{23}{6} + \frac{5}{12}\sqrt{5}\,a_x \\ -\frac{11}{6}\,a_x + \frac{1}{12}\sqrt{5}\,a_x^2 & \frac{1}{30}\,a_x^2 - \frac{23}{6} + \frac{5}{12}\sqrt{5}\,a_x & -\frac{23}{720}\,a_x^2 + \frac{529}{36} - \frac{47}{36}\sqrt{5}\,a_x \end{bmatrix} \tag{40}$$

Now, the above matrix is substantially easier to deal with in terms of computations. From here, it will be expanded and placed in terms of the homogenous point coordinates, [x__0,x__1,x__2], defined below,

$xxT := Vector[row]([x_0, x_1, x_2])$

$$\begin{bmatrix} x_0 & x_1 & x_2 \end{bmatrix} \tag{41}$$

$xx := Vector[column]([x_0, x_1, x_2])$

$$\begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} \quad (42)$$

Now that the defining point coordinate vectors have been established, the expansion of the point conic coefficient matrix yields the following expression,

$HomogenousPointConicEquation := collect(multiply((41), (40), (42)), [x_0, x_1, x_2], distributed)$

$$a_x^2 x_0^2 + \left(-\frac{11}{3} a_x + \frac{1}{6}\sqrt{5}\, a_x^2\right) x_2 x_0 - 2 a_x x_1 x_0 + x_1^2 + \left(\frac{1}{15} a_x^2 - \frac{23}{3} + \frac{5}{6}\sqrt{5}\, a_x\right) x_2 x_1 + x_2^2 \left(-\frac{23}{720} a_x^2 + \frac{529}{36} - \frac{47}{36}\sqrt{5}\, a_x\right) \quad (43)$$

The above equation represents the general equation for the pencil of ellipses contained within the previously defined quadrilateral. Now, we will define values of a__x and subsequently, members of the pencil, and their equations, based upon the side of the quadrilateral which lies coincident with the x__1 axis.

$Ellipsoid1 := subs\left(x_0 = (7)[1], a_x = \left(\frac{1}{10}\right)(8)[2], (43)\right)$

$$\frac{1}{5} - \frac{7}{10}\sqrt{5}\, x_2 - \frac{2}{5} x_1 \sqrt{5} + x_1^2 - \frac{341}{50} x_2 x_1 + \frac{5353}{400} x_2^2 \quad (44)$$

$Ellipsoid2 := subs\left(x_0 = (7)[1], a_x = \left(\frac{1}{4}\right)(8)[2], (43)\right)$

$$\frac{5}{4} - \frac{13}{8}\sqrt{5}\, x_2 - x_1 \sqrt{5} + x_1^2 - \frac{11}{2} x_2 x_1 + \frac{729}{64} x_2^2 \quad (45)$$

$Ellipsoid3 := subs\left(x_0 = (7)[1], a_x = \left(\frac{1}{2}\right)(8)[2], (43)\right)$

$$5 - \frac{17}{6}\sqrt{5}\, x_2 - 2 x_1 \sqrt{5} + x_1^2 - \frac{19}{6} x_2 x_1 + \frac{1153}{144} x_2^2 \quad (46)$$

$Ellipsoid4 := subs\left(x_0 = (7)[1], a_x = \left(\frac{3}{4}\right)(8)[2], (43)\right)$

$$\frac{45}{4} - \frac{29}{8}\sqrt{5}\, x_2 - 3 x_1 \sqrt{5} + x_1^2 - \frac{2}{3} x_2 x_1 + \frac{2617}{576} x_2^2 \quad (47)$$

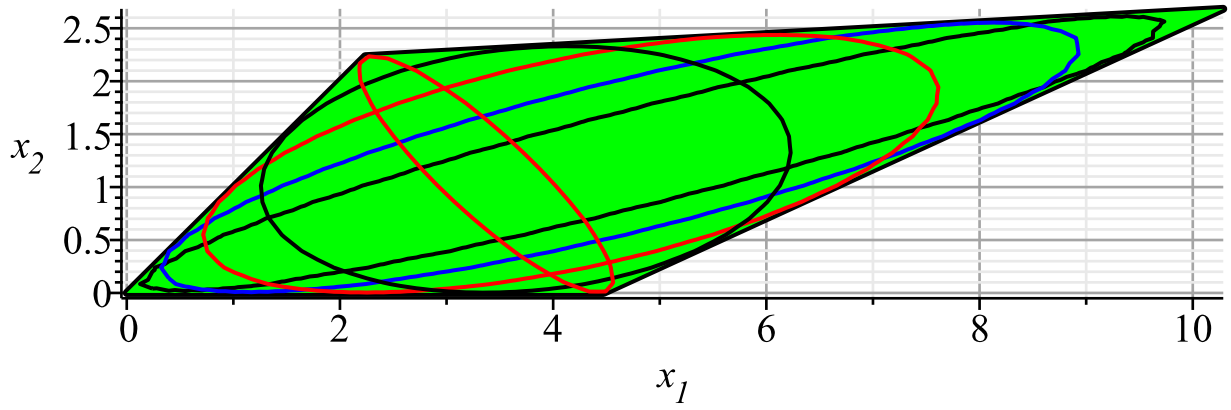
$Ellipsoid5 := subs\left(x_0 = (7)[1], a_x = \left(\frac{99}{100}\right)(8)[2], (43)\right)$

$$\frac{9801}{500} - \frac{3993}{1000}\sqrt{5}\, x_2 - \frac{99}{25} x_1 \sqrt{5} + x_1^2 + \frac{3544}{1875} x_2 x_1 + \frac{411577}{360000} x_2^2 \quad (48)$$

Once the ellipses have been described via the point conic ellipse equations, they can be plotted alongside the initial quadrilateral, as defined by the user. They must first be evaluated, between the bounds of the quadrilateral in question, and then plotted on the same chart.

$InscribingEllipse1 := implicitplot((44), x_1 = -1..10, x_2 = -1..10, scaling = constrained, numpoints = 5000, colour = black) : InscribingEllipse2 := implicitplot((45), x_1 = -1..10, x_2 = -1..10, scaling$

```
= constrained, numpoints = 5000, colour = blue) : InscribingEllipse3 := implicitplot((46), x1 = -1
..10, x2 = -1..10, scaling = constrained, numpoints = 5000, colour = red) : InscribingEllipse4 :=
implicitplot((47), x1 = -1..10, x2 = -1..10, scaling = constrained, numpoints = 5000, colour
= black) : InscribingEllipse5 := implicitplot((48), x1 = -1..10, x2 = -1..10, scaling = constrained,
numpoints = 5000, colour = red) :
display(TranQuad1, InscribingEllipse1, InscribingEllipse2, InscribingEllipse3, InscribingEllipse4,
InscribingEllipse5)
```



From this plot, it is quite clear that the pencil of ellipses defined by way of the line coordinate shape coefficient matrix translated equally well into the point conic shape coefficient matrix. It is easy to see that, as the polar point a__x moves along the x__1 axis, the area tends to increase towards a maximum and then decrease from that maximum, in either the positive or the negative x__1 direction, to become a degenerate ellipse.

In order to define the maximum area inscribing ellipse, one must place the ellipse into a parametric form whereby the area may then be computed. Furthermore, in order for this operation to contain a meaningful answer, the polar point will be used as the parameter which this equation is placed in terms of. Gfrerer proposed an equation that was presented in his paper *The Area Maximizing Inellipse of a Convex Quadrangle,* which is presented below, and used for exactly this purpose.

```
DeterminantPointConicMatrix := collect(det((40)), a_x)
```

$$-\frac{157}{180}\,a_x^{\,4}+\frac{17}{3}\,a_x^{\,3}\sqrt{5}-\frac{1}{900}\,a_x^{\,6}-\frac{1}{30}\,a_x^{\,5}\sqrt{5}-\frac{289}{9}\,a_x^{\,2} \qquad \textbf{(49)}$$

```
MinorDeterminantDelta := collect(det(minor((40), 1, 1)), a_x)
```

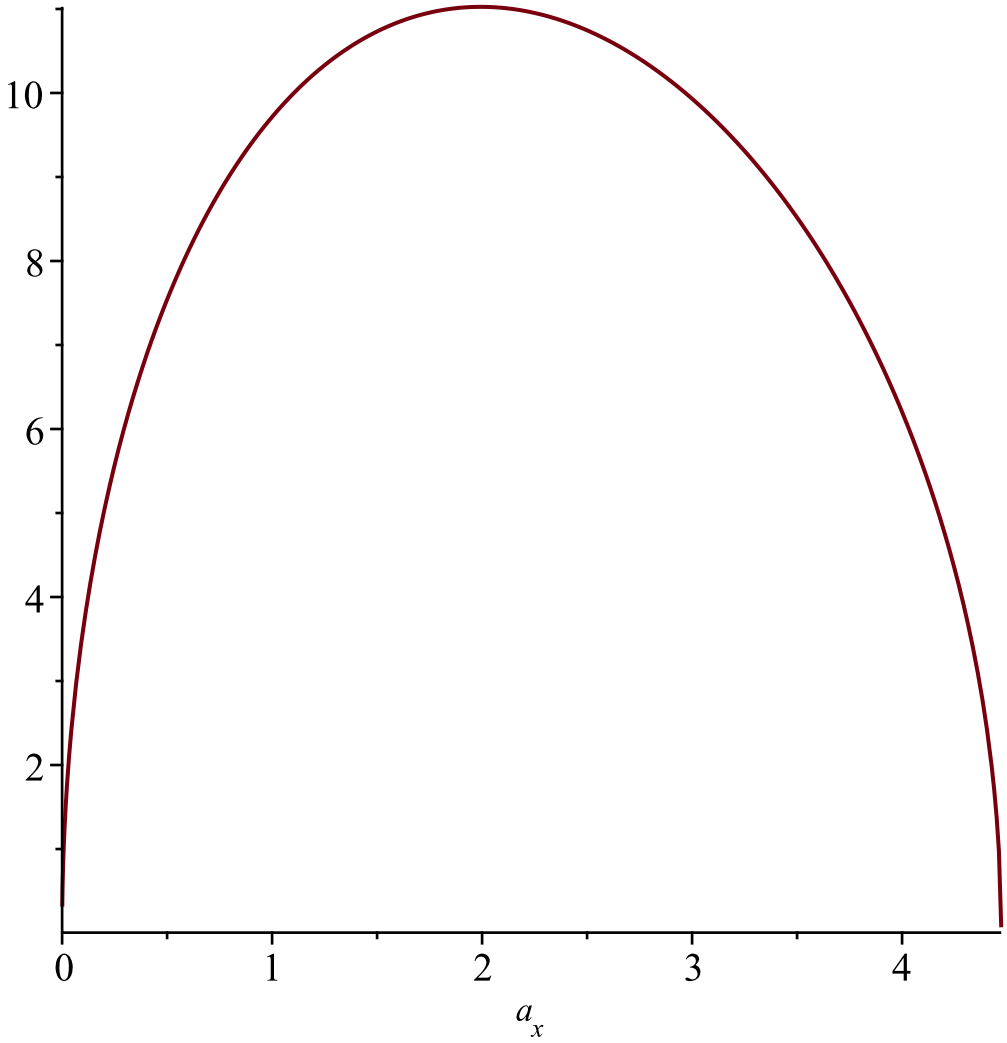
$$-\frac{29}{45}\,a_x^{\,2}+\frac{17}{9}\,\sqrt{5}\,a_x-\frac{1}{900}\,a_x^{\,4}-\frac{1}{36}\,a_x^{\,3}\,\sqrt{5} \tag{50}$$

$$AreaFunctionPointConic := \mathrm{abs}\left(simplify\left(collect\left(-\frac{((49)\cdot \mathrm{Pi})}{(50)^{\frac{1}{2}}},\,a_x\right)\right)\right)$$

$$30\,\pi\left|\left(a_x\left(30\,a_x^{\,3}\sqrt{5}+a_x^{\,4}-5100\sqrt{5}\,a_x+785\,a_x^{\,2}+28900\right)\right)\Big/\left(\left(25\sqrt{5}\,a_x^{\,2}+a_x^{\,3}-1700\sqrt{5}+580\,a_x\right)\sqrt{-a_x\left(25\sqrt{5}\,a_x^{\,2}+a_x^{\,3}-1700\sqrt{5}+580\,a_x\right)}\right)\right| \tag{51}$$

In order to verify the assertion that the area function does indeed contain a maximum, and to better visualize the areas of the ellipses being plotted, it is useful to display the function along the bounded interval which defines the base of the quadrilateral, and is incident with the x__1 axis.

$$plot\left((51),\,a_x=(7)[2]..(8)[2]\right)$$



Due to the nature of the function and the fact that it contains complex values for the ellipses lying on either side of the degenerate ellipse states, there is no easy way to solve for the maximum of this function directly from the area function. Instead, it will be treated much like a general quadratic

function, wherein the derivate is taken, and the zero point will define the maximum of the function.

$$AreaDerivativeFunction := simplify\left(diff\left(\mathbf{(51)}, a_x\right)\right)$$

$$\left(75\,\pi\left(1193\,a_x^{\,8}+3\sqrt{5}\,a_x^{\,9}+2242195\,a_x^{\,6}+35245\sqrt{5}\,a_x^{\,7}-255186900\,a_x^{\,4}\right.\right. \tag{52}$$

$$+8158800\,a_x^{\,5}\sqrt{5}+27453266000\,a_x^{\,2}-994602000\,a_x^{\,3}\sqrt{5}+83521000000$$

$$\left.\left.-38517920000\sqrt{5}\,a_x\right)\operatorname{abs}\left(1,\right.\right.$$

$$\left.\frac{a_x\left(30\,a_x^{\,3}\sqrt{5}+a_x^{\,4}-5100\sqrt{5}\,a_x+785\,a_x^{\,2}+28900\right)}{\left(25\sqrt{5}\,a_x^{\,2}+a_x^{\,3}-1700\sqrt{5}+580\,a_x\right)\sqrt{-a_x\left(25\sqrt{5}\,a_x^{\,2}+a_x^{\,3}-1700\sqrt{5}+580\,a_x\right)}}\right)$$

$$\Bigg/\left(\left(25\sqrt{5}\,a_x^{\,2}+a_x^{\,3}-1700\sqrt{5}\right.\right.$$

$$\left.\left.+580\,a_x\right)^3\sqrt{-a_x\left(25\sqrt{5}\,a_x^{\,2}+a_x^{\,3}-1700\sqrt{5}+580\,a_x\right)}\right)$$

$$AreaDerivativeFunctionRoots := solve\left(AreaDerivativeFunction = 0, a_x\right)$$

$$-\frac{184}{15}\sqrt{5}+\frac{2}{15}\sqrt{48695},\ -\frac{184}{15}\sqrt{5}-\frac{2}{15}\sqrt{48695},\ -10\sqrt{5},\ 2\sqrt{5},\ -17\sqrt{5} \tag{53}$$

$$evalf\left(AreaDerivativeFunctionRoots\right)$$

$$1.99349093,\ -56.85169199,\ -22.36067977,\ 4.472135954,\ -38.01315561 \tag{54}$$

$$MaxArea := simplify\left(subs\left(a_x = \mathbf{(53)}[1], AreaFunctionPointConic\right)\right)$$

$$-\frac{15\,\pi\sqrt{-92+\sqrt{9739}}\left(-170155486+1698983\sqrt{9739}\right)}{\left(-7156\sqrt{9739}+797327\right)^{3/2}} \tag{55}$$

$$evalf\left(\mathbf{(55)}, 6\right)$$

$$11.0254 \tag{56}$$

As can be seen from equation 55, there is far more than one zero point contained within the equation. However, anything that lies outside the bounds of the quadrilateral results in physically meaningless values; either errors resulting from a zero division, or a purely imaginary value in the case of the first zero. In essence, there is one real maximum area, and that is the area defined by the second coordinate. It would also appear, regardless of the initially defined quadrilateral, to always be the second zero of the area derivative function.

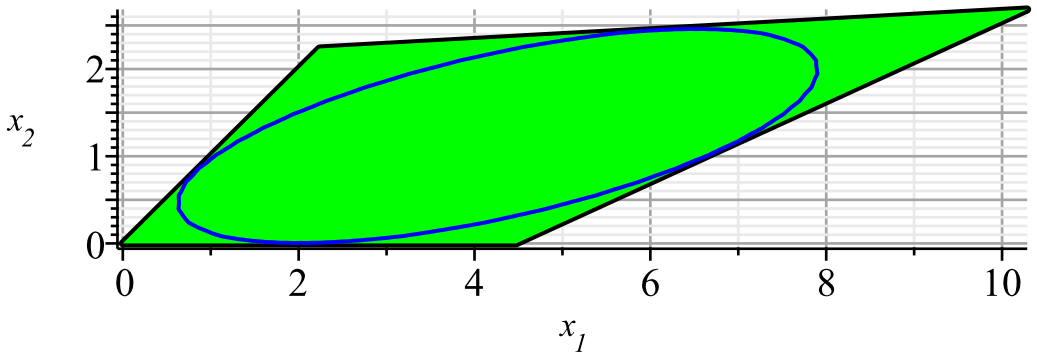
From here, it is quite useful to plot this maximum ellipse.

$$EllipseMax := collect\left(simplify\left(subs\left(x_0 = 1,\ a_x = \textbf{(53)}[1],\ \textbf{(43)}\right)\right),\ [x_1, x_2],\ distributed\right)$$

$$\frac{72812}{45} - \frac{736}{45}\sqrt{9739} + x_1^{\,2} + \left(\frac{87239}{2025} - \frac{1409}{4050}\sqrt{9739}\right)x_2^{\,2} + \left(\frac{42478}{135}\sqrt{5} - \frac{434}{135}\sqrt{5}\sqrt{9739}\right)x_2 + \left(\frac{368}{15}\sqrt{5} - \frac{4}{15}\sqrt{5}\sqrt{9739}\right)x_1 + \left(\frac{33137}{675} - \frac{361}{675}\sqrt{9739}\right)x_2 x_1 \qquad \textbf{(57)}$$

$$InscribingEllipseMax := implicitplot\left(\textbf{(57)},\ x_1 = -1..10,\ x_2 = -1..10,\ scaling = constrained,\ numpoints = 5000,\ colour = blue\right):$$

$$display(TranQuad1,\ InscribingEllipseMax)$$



While it is possible that this code will fail for certain specific orientations of ellipses, specifically those that are translated in such a way that one of the line coordinates is purely negative with respect to X__1, this general approach is a good start for the majority of quadrilaterals.

Obviously, more test cases will be required in the event that this is to be made truly robust, however, the solution method presented herein will always allow for solution of the maximum area inscribing ellipse within a convex quadrilateral.

Further expansion of this code is possible in order to demonstrate that the centre location for all of the ellipses lies along the line that bisects the corner connecting lines of the quadrilateral.