

Where is the Cone?

Drew Armstrong
www.math.miami.edu/~armstrong

August 16, 2017

Introduction

A *real quadric curve* is defined by an equation of the form

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

for some real numbers $a, b, c, d, e, f \in \mathbb{R}$. One often sees real quadric curves described as “conic sections.” This terminology suggests that we can view this curve as the intersection of the x, y -plane with a circular cone living in x, y, z -space. However, it is very rare to see the details of this spelled out. The purpose of the present note is to answer the question: **where is the cone?** More specifically, we will answer the following three questions:

- From which points in space does a given quadric curve look like a circle?
(This is the apex of the cone.)
- In which direction should we look to see this circle?
(This is the axis of symmetry of the cone.)
- How big is the circle?
(This is the angle of aperture of the cone.)

The answers to these questions were well known to geometers in the late nineteenth and early twentieth centuries. However, based on my several years of internet searching, it seems that the answers are not well known today. Eventually I found a clue in footnote 4 on page 24 of Hilbert and Cohn-Vossen’s *Geometry and the Imagination* (1932).¹ This led me to a genre of textbooks on the “analytic geometry of three dimensions” that were written around the same time. I found the text of D.M.Y. Sommerville (1934) particularly helpful.

Without further ado, here is the main theorem describing cones over quadric curves.

¹For the precise statement see Section 5 below.

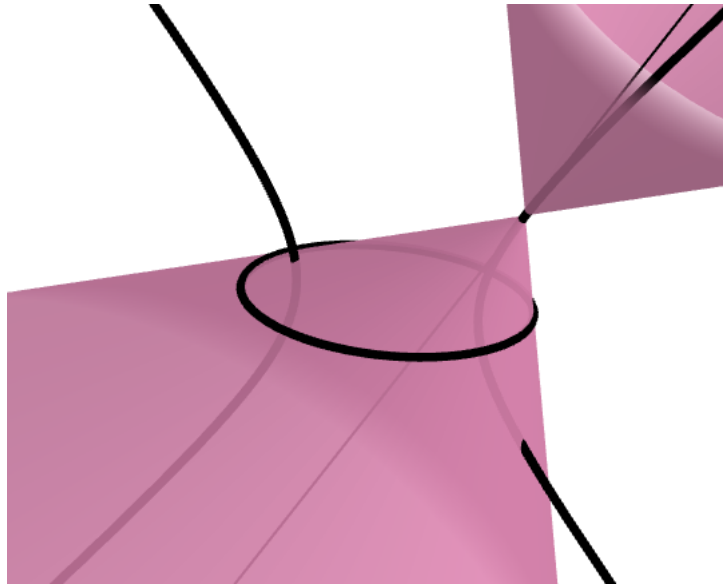


Figure 1: Realizing an ellipse as a conic section.

Main Theorem. Let $a, b, c \in \mathbb{R}$ be distinct real numbers and consider the following space curves in the three principal coordinate planes:

$$\begin{aligned} x^2/(a-c) + y^2/(b-c) = 1 \quad \text{and} \quad z = 0, \\ x^2/(a-b) + z^2/(c-b) = 1 \quad \text{and} \quad y = 0, \\ y^2/(b-a) + z^2/(c-a) = 1 \quad \text{and} \quad x = 0. \end{aligned}$$

In general, two of these curves are real and the third is imaginary. If \mathbf{u} is any point on one of the real curves then the cone from \mathbf{u} to the other real curve is circular. Furthermore, the axis of symmetry of this cone is the tangent line to the first curve at \mathbf{u} .² The cone from any **other** point in space to any one of the real curves is **not** circular. ///

For example, if $c < b < a$ then we have an ellipse in the x, y -plane and a hyperbola in the x, z -plane. Figure 1 shows a typical circular cone over the ellipse and Figure 2 shows a typical circular cone over the hyperbola.³

This theorem provides a completely satisfying answer to the question “where is the cone?” and I am surprised by how difficult it was to track down. After reading the result in Hilbert and Cohn-Vossen I found it stated in several textbooks of the period.⁴ However, I still have not seen a proof written down anywhere. Proving it turned out to be trickier than I expected but eventually I found a slick proof via matrix arithmetic.

²and from this one can easily compute the angle of aperture

³These pictures were made with GeoGebra.

⁴It even appears as Exercise 6 on page 100 of Barry Spain’s *Analytic Quadrics* (1960).

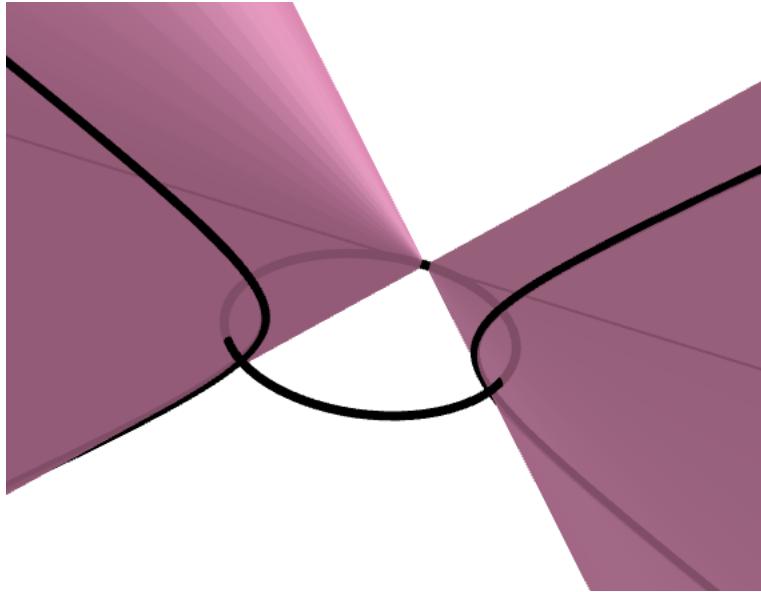


Figure 2: Realizing a hyperbola as a conic section.

Note that the Main Theorem stated above applies to the case of *non-degenerate* and *central* quadric curves (ellipses and hyperbolas). Analogous results for non-central quadric curves (parabolas) and degenerate quadric curves (line pairs) can be obtained from limiting arguments that are not so interesting, so I will omit them. My goal in this note is to provide the slickest possible explanation for the statement in Hilbert and Cohn-Vossen's footnote by employing standard linear algebra that was not standard in 1932. The reader can judge if I have been successful. Here is the outline:

Contents

1	The Equation of a Circular Cone	4
2	Quadric Cones in General	5
3	Tangent Cones Over Quadric Surfaces	9
4	Confocal Quadric Surfaces	11
5	Symmetries of a Tangent Cone	15
6	Proof of the Main Theorem	17
7	Conclusion	21

1 The Equation of a Circular Cone

To give an analytic treatment of conic sections we must first give an analytic treatment of cones. A *circular cone* in \mathbb{R}^3 is specified by the following data:

- A point in space $\mathbf{u} \in \mathbb{R}^3$ defining the apex,
- A unit vector $\mathbf{r} \in \mathbb{R}^3$ defining the axis of symmetry,
- An angle $\theta \in [0, \pi/2]$ defining the aperture of the cone.⁵

Geometrically, the cone consists of all points $\mathbf{x} \in \mathbb{R}^3$ such that the line connecting \mathbf{x} and \mathbf{u} makes an angle of θ (or $\pi - \theta$) with the axis of symmetry. Algebraically, we can express this situation with the dot product. Since \mathbf{r} is a unit vector we have

$$(*) \quad (\mathbf{x} - \mathbf{u})^T \mathbf{r} = \|\mathbf{x} - \mathbf{u}\| \cdot \|\mathbf{r}\| \cdot \cos \theta = \|\mathbf{x} - \mathbf{u}\| \cdot \cos \theta.$$

Furthermore, since every 1×1 matrix is symmetric we have

$$(\mathbf{x} - \mathbf{u})^T \mathbf{r} = [(\mathbf{x} - \mathbf{u})^T \mathbf{r}]^T = \mathbf{r}^T (\mathbf{x} - \mathbf{u}).$$

Then by squaring both sides of (*) we obtain the following equation for the circular cone:

$$\begin{aligned} [(\mathbf{x} - \mathbf{u})^T \mathbf{r}]^2 &= \|\mathbf{x} - \mathbf{u}\|^2 \cdot \cos^2 \theta \\ [(\mathbf{x} - \mathbf{u})^T \mathbf{r}] \cdot [\mathbf{r}^T (\mathbf{x} - \mathbf{u})] &= (\mathbf{x} - \mathbf{u})^T (\mathbf{x} - \mathbf{u}) \cdot \cos^2 \theta \\ (\mathbf{x} - \mathbf{u})^T (\mathbf{r} \mathbf{r}^T) (\mathbf{x} - \mathbf{u}) &= (\mathbf{x} - \mathbf{u})^T ((\cos^2 \theta) I) (\mathbf{x} - \mathbf{u}) \\ (\mathbf{x} - \mathbf{u})^T (\mathbf{r} \mathbf{r}^T - (\cos^2 \theta) I) (\mathbf{x} - \mathbf{u}) &= 0. \end{aligned}$$

In summary, for any unit vector $\mathbf{r} = (r, s, t)$ and angle θ we define the matrix

$$C_{\mathbf{r}, \theta} := \mathbf{r} \mathbf{r}^T - (\cos^2 \theta) I = \begin{pmatrix} r^2 - \cos^2 \theta & rs & rt \\ rs & s^2 - \cos^2 \theta & st \\ rt & st & t^2 - \cos^2 \theta \end{pmatrix}.$$

Then the circular cone with apex \mathbf{u} , axis \mathbf{r} and aperture θ is defined by the equation

$$\boxed{(\mathbf{x} - \mathbf{u})^T C_{\mathbf{r}, \theta} (\mathbf{x} - \mathbf{u}) = 0.}$$

Let us make a few observations about the *cone matrix* $C_{\mathbf{r}, \theta}$. Since $\mathbf{r}^T \mathbf{r} = \|\mathbf{r}\|^2 = 1$ we observe that \mathbf{r} is an eigenvector of $C_{\mathbf{r}, \theta}$ with eigenvalue $1 - \cos^2 \theta$:

$$C_{\mathbf{r}, \theta} \cdot \mathbf{r} = (\mathbf{r} \mathbf{r}^T - (\cos^2 \theta) I) \mathbf{r} = \mathbf{r} (\mathbf{r}^T \mathbf{r}) - (\cos^2 \theta) \mathbf{r} = (1 - \cos^2 \theta) \mathbf{r}.$$

⁵The extreme values $\theta = 0$ and $\theta = \pi/2$ correspond to a line and a plane, respectively.

And if \mathbf{r}' is any vector perpendicular to \mathbf{r} (i.e., if $\mathbf{r}^T \mathbf{r}' = 0$) then we observe that \mathbf{r}' is an eigenvector of $C_{\mathbf{r},\theta}$ with eigenvalue $-\cos^2 \theta$:

$$C_{\mathbf{r},\theta} \cdot \mathbf{r}' = (\mathbf{r}\mathbf{r}^T - (\cos^2 \theta)I) \mathbf{r}' = \mathbf{r}(\mathbf{r}^T \mathbf{r}') - (\cos^2 \theta)\mathbf{r}' = -(\cos^2 \theta)\mathbf{r}'.$$

We conclude that the cone matrix $C_{\mathbf{r},\theta}$ has eigenvalues

$$1 - \cos^2 \theta, \quad -\cos^2 \theta, \quad -\cos^2 \theta.$$

Conversely, let $C^T = C$ be **any** real symmetric matrix with the same eigenvalues. In this case I claim that $C = C_{\mathbf{r},\theta}$ for some unit vector $\mathbf{r} \in \mathbb{R}^3$. Indeed, since any real symmetric matrix is orthogonally diagonalizable (by the Principal Axis Theorem), we know that there exists a real orthogonal matrix $Q^T = Q^{-1}$ such that

$$Q^T C Q = \begin{pmatrix} 1 - \cos^2 \theta & 0 & 0 \\ 0 & -\cos^2 \theta & 0 \\ 0 & 0 & -\cos^2 \theta \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - (\cos^2 \theta)I.$$

Now let \mathbf{r} be the first column of the matrix Q , i.e., the eigenvector of C corresponding to eigenvalue $1 - \cos^2 \theta$. Since $Q Q^T = I$ we see that \mathbf{r} is a unit vector and we observe that

$$C = Q \left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - (\cos^2 \theta)I \right] Q^T = Q \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} Q^T - (\cos^2 \theta)Q Q^T = \mathbf{r}\mathbf{r}^T - (\cos^2 \theta)I$$

as desired.

2 Quadric Cones in General

More generally, for any real symmetric matrix $C^T = C$ we consider the quadratic equation

$$(\mathbf{x} - \mathbf{u})^T C (\mathbf{x} - \mathbf{u}) = 0.$$

Clearly the point $\mathbf{x} = \mathbf{u}$ is a solution. If the (necessarily real) eigenvalues of C are either all negative or all positive then $\mathbf{x} = \mathbf{u}$ is the **only (real) solution**. We are interested in the case when the eigenvalues of C are nonzero and not all of the same sign. In this case we say that C is *indefinite* and we say that the equation $(\mathbf{x} - \mathbf{u})^T C (\mathbf{x} - \mathbf{u}) = 0$ defines a *quadric cone*.

The following theorem says that the reflection symmetries of a quadric cone are the same as the eigenvectors of its matrix. In fact it is true that **every** symmetry of the cone is a product of reflection symmetries,⁶ but we won't need that result.

Theorem. Let $C^T = C$ be an indefinite real symmetric matrix (i.e., whose eigenvalues are nonzero and not all of the same sign). For any vector $\mathbf{p} \in \mathbb{R}^3$ the following are equivalent:

⁶This is the famous Cartan-Dieudonne Theorem.

- We have $C\mathbf{p} = c\mathbf{p}$ for some scalar $c \in \mathbb{R}$.
- The quadric cone $(\mathbf{x} - \mathbf{u})^T C(\mathbf{x} - \mathbf{u}) = 0$ is symmetric about the plane $\mathbf{u} + \mathbf{p}^\perp$ which is perpendicular to \mathbf{p} and passes through the apex \mathbf{u} .

///

To prove this we require a general lemma on quadratic forms, which is an easy consequence of Hilbert's Nullstellensatz. To keep the treatment self-contained I will present an elementary proof that I learned from John H. Elton (2009).

Lemma. Let $A^T = A$ and $B^T = B$ be real symmetric matrices of the same size, with A indefinite. If the cone of A is contained in the cone of B , i.e., if for all vectors \mathbf{x} we have

$$\mathbf{x}^T A \mathbf{x} = 0 \implies \mathbf{x}^T B \mathbf{x} = 0,$$

then it follows that $B = \lambda A$ for some real scalar $\lambda \neq 0$.

///

Proof of the Lemma. We will prove the result for 3×3 matrices. (See Elton for the general case.) By the Principal Axis Theorem there exists an orthogonal matrix $Q^T = Q^{-1}$ such that $Q^T A Q$ is diagonal:

$$Q^T A Q = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix}.$$

Since Q is invertible, it is enough to show that $Q^T B Q$ is a scalar multiple of $Q^T A Q$. Furthermore, since A is indefinite we can assume without loss of generality (replacing A by $-A$ if necessary) that the eigenvalues of A satisfy $\lambda_1 > 0 > \lambda_2, \lambda_3$. Let us define the positive real numbers

$$\ell_1 := 1/\sqrt{\lambda_1}, \quad \ell_2 := 1/\sqrt{-\lambda_2} \quad \text{and} \quad \ell_3 := 1/\sqrt{-\lambda_3}.$$

We observe that the following two equations hold:

$$0 = (\ell_1 \quad \pm\ell_2 \quad 0) \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix} \begin{pmatrix} \ell_1 \\ \pm\ell_2 \\ 0 \end{pmatrix} = (\ell_1 \quad \pm\ell_2 \quad 0) Q^T A Q \begin{pmatrix} \ell_1 \\ \pm\ell_2 \\ 0 \end{pmatrix}.$$

Then the implication $(\mathbf{x}^T A \mathbf{x} = 0) \implies (\mathbf{x}^T B \mathbf{x} = 0)$ with $\mathbf{x}^T = (\ell_1 \quad \pm\ell_2 \quad 0) Q^T$ tells us that

$$0 = (\ell_1 \quad \pm\ell_2 \quad 0) Q^T B Q \begin{pmatrix} \ell_1 \\ \pm\ell_2 \\ 0 \end{pmatrix}.$$

By writing b_{ij} for the (i, j) -the entry of $Q^T BQ$ these two equations become

$$\begin{cases} b_{11}/\lambda_1 - b_{22}/\lambda_2 + 2b_{12}\ell_1\ell_2 = 0, \\ b_{11}/\lambda_1 - b_{22}/\lambda_2 - 2b_{12}\ell_1\ell_2 = 0. \end{cases}$$

By adding these equations we find that $b_{22} = \lambda_2(b_{11}/\lambda_1)$ and hence $b_{12} = 0$. Then a similar argument shows that

$$0 = (\ell_1 \ 0 \ \pm\ell_3) Q^T BQ \begin{pmatrix} \ell_1 \\ 0 \\ \pm\ell_3 \end{pmatrix},$$

which implies that $b_{33} = \lambda_3(b_{11}/\lambda_1)$ and $b_{13} = 0$. At this point we know that

$$Q^T BQ - \left(\frac{b_{11}}{\lambda_1}\right) Q^T A Q = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & b_{23} \\ 0 & b_{23} & 0 \end{pmatrix}$$

and it remains only to show that $b_{23} = 0$. To do this we fix any real numbers (α, β, γ) with the properties $\alpha^2 = \beta^2 + \gamma^2$ and $\beta\gamma \neq 0$.⁷ Then we obtain

$$(\alpha\ell_1 \ \beta\ell_2 \ \gamma\ell_3) Q^T A Q \begin{pmatrix} \alpha\ell_1 \\ \beta\ell_2 \\ \gamma\ell_3 \end{pmatrix} = \alpha^2 - \beta^2 - \gamma^2 = 0 \quad \Rightarrow \quad (\alpha\ell_1 \ \beta\ell_2 \ \gamma\ell_3) Q^T BQ \begin{pmatrix} \alpha\ell_1 \\ \beta\ell_2 \\ \gamma\ell_3 \end{pmatrix} = 0,$$

which implies that

$$2\beta\gamma\ell_2\ell_3b_{23} = (\alpha\ell_1 \ \beta\ell_2 \ \gamma\ell_3) \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & b_{23} \\ 0 & b_{23} & 0 \end{pmatrix} \begin{pmatrix} \alpha\ell_1 \\ \beta\ell_2 \\ \gamma\ell_3 \end{pmatrix} = 0.$$

Finally, since $\beta\gamma \neq 0$ we conclude that $b_{23} = 0$. This completes the proof. \square

Proof of the Theorem. Recall that $C^T = C$ is an indefinite real symmetric matrix. It is sufficient to prove the theorem in the case when \mathbf{p} is a unit vector and the apex is at the origin: $\mathbf{u} = \mathbf{0}$. Since \mathbf{p} is a unit vector (i.e., $\mathbf{p}^T \mathbf{p} = 1$) we observe that $P = \mathbf{p}\mathbf{p}^T$ is the matrix that projects orthogonally onto the line $\mathbb{R}\mathbf{p}$ and that

$$R = I - 2P = I - 2\mathbf{p}\mathbf{p}^T$$

is the matrix that reflects orthogonally across the plane \mathbf{p}^\perp .

First let us suppose that \mathbf{p} is an eigenvector of C , say $C\mathbf{p} = c\mathbf{p}$. To prove that the reflection R leaves the cone invariant we must show for any vector \mathbf{x} that

$$\mathbf{x}^T C \mathbf{x} = 0 \quad \Rightarrow \quad (R\mathbf{x})^T C (R\mathbf{x}) = \mathbf{x}^T R C R \mathbf{x} = 0.$$

⁷Elton chooses $(\alpha, \beta, \gamma) = (5, 4, 3)$.

By observing that

$$\begin{aligned}
RCR &= (I - 2\mathbf{p}\mathbf{p}^T)C(I - 2\mathbf{p}\mathbf{p}^T) \\
&= C - 2\mathbf{p}(\mathbf{p}^T C) - 2(C\mathbf{p})\mathbf{p}^T + 4\mathbf{p}\mathbf{p}^T(C\mathbf{p})\mathbf{p}^T \\
&= C - 2\mathbf{p}(c\mathbf{p}^T) - 2(c\mathbf{p})\mathbf{p}^T + 4\mathbf{p}\mathbf{p}^T(c\mathbf{p})\mathbf{p}^T \\
&= C - 2c\mathbf{p}\mathbf{p}^T - 2c\mathbf{p}\mathbf{p}^T + 4c\mathbf{p}(\mathbf{p}^T \mathbf{p})\mathbf{p}^T \\
&= C - 2c\mathbf{p}\mathbf{p}^T - 2c\mathbf{p}\mathbf{p}^T + 4c\mathbf{p}\mathbf{p}^T \\
&= C,
\end{aligned}$$

we see that the statement is vacuously true. Conversely, let us suppose that

$$\mathbf{x}^T C \mathbf{x} = 0 \quad \Rightarrow \quad \mathbf{x}^T R C R \mathbf{x} = 0$$

for all vectors \mathbf{x} . Since C is indefinite, the Lemma tells us that $RCR = \lambda C$ for some scalar $\lambda \in \mathbb{R}$. Since $\det(C) \neq 0$ and $\det(R) = -1$ this implies that

$$\begin{aligned}
\det(\lambda C) &= \det(RCR) \\
\lambda^3 \det(C) &= \det(R)^2 \det(C) \\
\lambda^3 \det(C) &= \det(C) \\
\lambda^3 &= 1,
\end{aligned}$$

which implies that $\lambda = 1$ since λ is real. Thus we conclude that $RCR = C$. Finally, since $R^2 = I$ (indeed, R is a reflection) we observe that

$$\begin{aligned}
RC &= CR \\
(I - 2\mathbf{p}\mathbf{p}^T)C &= C(I - 2\mathbf{p}\mathbf{p}^T) \\
C - 2\mathbf{p}(C\mathbf{p})^T &= C - 2(C\mathbf{p})\mathbf{p}^T \\
\mathbf{p}(C\mathbf{p})^T &= (C\mathbf{p})\mathbf{p}^T.
\end{aligned}$$

Let \mathbf{p}_i and $(C\mathbf{p})_i$ denote the i -th entries of the vectors \mathbf{p} and $C\mathbf{p}$, respectively. By comparing the (i, j) -th entry on both sides of the previous equation we find that $\mathbf{p}_i(C\mathbf{p})_j = (C\mathbf{p})_i \mathbf{p}_j$ for all i and j . Since $\mathbf{p} \neq \mathbf{0}$ there exists some index i such that $\mathbf{p}_i \neq 0$; let us define $c := (C\mathbf{p})_i / \mathbf{p}_i$. Then for all indices j we have $(C\mathbf{p})_j = c\mathbf{p}_j$ and it follows that $C\mathbf{p} = c\mathbf{p}$ as desired. \square

The following theorem summarizes the results of Sections 1 and 2.

Theorem. Let $C^T = C$ be an indefinite real symmetric matrix with (real, nonzero) eigenvalues λ, μ, ν , not all of the same sign. By the Principal Axis Theorem there exists an orthogonal basis of eigenvectors $\mathbf{r}, \mathbf{s}, \mathbf{t}$ corresponding to eigenvalues λ, μ, ν , respectively.

- If λ, μ, ν are distinct then the quadric cone $(\mathbf{x} - \mathbf{u})^T C (\mathbf{x} - \mathbf{u}) = 0$ is symmetric with respect to the three mutually perpendicular planes $\mathbf{u} + \mathbf{r}^\perp, \mathbf{u} + \mathbf{s}^\perp, \mathbf{u} + \mathbf{t}^\perp$, called the *principal planes* of the cone. The cone is not symmetric with respect to any other plane.
- If two eigenvalues collide, say $\mu = \nu$, then the cone $(\mathbf{x} - \mathbf{u})^T C (\mathbf{x} - \mathbf{u}) = 0$ is circular with axis of symmetry $\mathbf{u} + \mathbb{R}\mathbf{r}$ and angle of aperture θ satisfying

$$\cos^2 \theta = \frac{\mu}{\mu - \lambda}.$$

///

Proof. It remains only to prove the second statement. We assume that the eigenvalues of C are λ, μ, μ with corresponding orthogonal eigenbasis $\mathbf{r}, \mathbf{s}, \mathbf{t}$. Since the eigenvalues of \mathbf{s} and \mathbf{t} are equal we see that the cone is symmetric with respect to any plane of the form $\mathbf{u} + (a\mathbf{s} + b\mathbf{t})^\perp$. In other words, the cone has rotational symmetry around the perpendicular axis $\mathbf{u} + \mathbb{R}\mathbf{r}$. Now let us consider the matrix $C' := C/(\lambda - \mu)$ with eigenvalues

$$\frac{\lambda}{\lambda - \mu}, \quad \frac{\mu}{\lambda - \mu}, \quad \frac{\mu}{\lambda - \mu}.$$

Observe that the equations $(\mathbf{x} - \mathbf{u})^T C (\mathbf{x} - \mathbf{u}) = 0$ and $(\mathbf{x} - \mathbf{u})^T C' (\mathbf{x} - \mathbf{u}) = 0$ define the same cone. Since λ and μ have opposite signs we observe that $0 < \mu/(\mu - \lambda), \lambda/(\lambda - \mu) < 1$ and $\mu/(\mu - \lambda) + \lambda/(\lambda - \mu) = 1$, hence there exists a unique angle $\theta \in [0, \pi/2]$ satisfying

$$\frac{\lambda}{\lambda - \mu} = 1 - \cos^2 \theta \quad \text{and} \quad \frac{\mu}{\mu - \lambda} = \cos^2 \theta.$$

It follows from the remarks of Section 1 that $C' = C_{\mathbf{r}, \theta}$ as desired. \square

3 Tangent Cones Over Quadric Surfaces

Our goal is to study the cone from a point \mathbf{u} in x, y, z -space to a quadric curve lying in the x, y -coordinate plane. Surprisingly, it turns out that the best way to do this is to first consider the **tangent cones** from a given point \mathbf{u} to a certain family of **quadric surfaces** in x, y, z -space. Then we will allow these quadric surfaces to degenerate to the desired quadric curve.

The quadric surfaces we will consider have the form

$$\mathbf{x}^T A \mathbf{x} = 1,$$

where $A^T = A$ is a real symmetric matrix. If \mathbf{u} is any point, then the *tangent cone from \mathbf{u} to the surface* consists of all points \mathbf{x} such that the line $t\mathbf{x} + (1-t)\mathbf{u}$ has double contact with the surface. Observe that the line and surface intersect when

$$\begin{aligned} (t\mathbf{x} + (1-t)\mathbf{u})^T A (t\mathbf{x} + (1-t)\mathbf{u}) &= 1 \\ t^2 \mathbf{x}^T A \mathbf{x} + 2t(1-t)\mathbf{x}^T A \mathbf{u} + (1-t)^2 \mathbf{u}^T A \mathbf{u} &= 1 \\ t^2(\mathbf{x}^T A \mathbf{x} + \mathbf{u}^T A \mathbf{u} - 2\mathbf{x}^T A \mathbf{u}) + t(2\mathbf{x}^T A \mathbf{u} - 2\mathbf{u}^T A \mathbf{u}) + (\mathbf{u}^T A \mathbf{u} - 1) &= 0. \end{aligned}$$

This quadratic equation in t usually has two distinct roots corresponding to two distinct points of contact with the surface. Double contact occurs when the discriminant vanishes, i.e., when

$$\begin{aligned} (2\mathbf{x}^T A \mathbf{u} - 2\mathbf{u}^T A \mathbf{u})^2 - 4(\mathbf{x}^T A \mathbf{x} + \mathbf{u}^T A \mathbf{u} - 2\mathbf{x}^T A \mathbf{u})(\mathbf{u}^T A \mathbf{u} - 1) &= 0 \\ (\mathbf{x}^T A \mathbf{u} - \mathbf{u}^T A \mathbf{u})^2 - (\mathbf{x}^T A \mathbf{x} + \mathbf{u}^T A \mathbf{u} - 2\mathbf{x}^T A \mathbf{u})(\mathbf{u}^T A \mathbf{u} - 1) &= 0 \\ (\mathbf{x}^T A \mathbf{u})^2 - 2(\mathbf{x}^T A \mathbf{u}) - (\mathbf{x}^T A \mathbf{x})(\mathbf{u}^T A \mathbf{u}) + \mathbf{x}^T A \mathbf{x} + \mathbf{u}^T A \mathbf{u} &= 0 \\ (\mathbf{x}^T A \mathbf{u} - 1)^2 - (\mathbf{x}^T A \mathbf{x} - 1)(\mathbf{u}^T A \mathbf{u} - 1) &= 0. \end{aligned}$$

In summary, the tangent cone from \mathbf{u} to the surface $\mathbf{x}^T A \mathbf{x} = 1$ has the equation

$$\boxed{(\mathbf{x}^T A \mathbf{u} - 1)^2 = (\mathbf{x}^T A \mathbf{x} - 1)(\mathbf{u}^T A \mathbf{u} - 1).}$$

However, since this is a quadric cone with apex at \mathbf{u} , we prefer to express it in the form $(\mathbf{x} - \mathbf{u})^T C (\mathbf{x} - \mathbf{u}) = 0$ for some real symmetric matrix $C^T = C$. To find the matrix C we expand the equations in terms of \mathbf{x} to get

$$\begin{aligned} (\mathbf{x} - \mathbf{u})^T C (\mathbf{x} - \mathbf{u}) &= 0 \\ \mathbf{x}^T [C] \mathbf{x} + [-2\mathbf{u}^T C] \mathbf{x} + [\mathbf{u}^T C \mathbf{u}] &= 0 \end{aligned}$$

and

$$\begin{aligned} (\mathbf{x}^T A \mathbf{u} - 1)^2 - (\mathbf{x}^T A \mathbf{x} - 1)(\mathbf{u}^T A \mathbf{u} - 1) &= 0 \\ (\mathbf{x}^T A \mathbf{u})^2 - 2(\mathbf{x}^T A \mathbf{u}) - (\mathbf{x}^T A \mathbf{x})(\mathbf{u}^T A \mathbf{u}) + \mathbf{x}^T A \mathbf{x} + \mathbf{u}^T A \mathbf{u} &= 0 \\ (\mathbf{x}^T A \mathbf{u})(\mathbf{x}^T A \mathbf{u})^T - 2(\mathbf{x}^T A \mathbf{u}) + (\mathbf{x}^T A \mathbf{x})(1 - \mathbf{u}^T A \mathbf{u}) + \mathbf{u}^T A \mathbf{u} &= 0 \\ \mathbf{x}^T (A \mathbf{u} \mathbf{u}^T A) \mathbf{x} + \mathbf{x}^T [(1 - \mathbf{u}^T A \mathbf{u}) A] \mathbf{x} - 2(\mathbf{x}^T A \mathbf{u}) + \mathbf{u}^T A \mathbf{u} &= 0 \\ \mathbf{x}^T [A \mathbf{u} \mathbf{u}^T A + (1 - \mathbf{u}^T A \mathbf{u}) A] \mathbf{x} + [-2\mathbf{u}^T A] \mathbf{x} + [\mathbf{u}^T A \mathbf{u}] &= 0. \end{aligned}$$

By comparing the leading quadratic forms we see that the cone matrix C satisfies

$$C = A \mathbf{u} \mathbf{u}^T A + (1 - \mathbf{u}^T A \mathbf{u}) A,$$

at least up to a scalar multiple. For this choice of C we easily verify that $-2\mathbf{u}^T C = -2\mathbf{u}^T A$ and $\mathbf{u}^T C \mathbf{u} = \mathbf{u}^T A \mathbf{u}$, so the linear and constant terms also agree.

In summary, we find that the tangent cone from the point \mathbf{u} to the quadric surface $\mathbf{x}^T A \mathbf{x} = 1$ has the equation

$$\boxed{(\mathbf{x} - \mathbf{u})^T [A \mathbf{u} \mathbf{u}^T A + (1 - \mathbf{u}^T A \mathbf{u}) A] (\mathbf{x} - \mathbf{u}) = 0.}$$

4 Confocal Quadric Surfaces

The specific surfaces that we need are called *confocal quadric surfaces*. This is the key idea that I learned from Hilbert and Cohn-Vossen.

Definition. Fix distinct real numbers $a, b, c \in \mathbb{R}$. Then for any real parameter $k \in \mathbb{R}$ not in the set $\{a, b, c\}$ we consider the quadric surface

$$\mathbf{x}^T A_k \mathbf{x} = \frac{x^2}{a-k} + \frac{y^2}{b-k} + \frac{z^2}{c-k} = 1,$$

where the matrix A_k is defined by

$$A_k := \begin{pmatrix} 1/(a-k) & & \\ & 1/(b-k) & \\ & & 1/(c-k) \end{pmatrix}.$$

The surfaces $\mathbf{x}^T A_k \mathbf{x} = 1$ for various k describe a *confocal family of quadrics*. ///

For example, let us assume that $c < b < a$. Then the surface $\mathbf{x}^T A_k \mathbf{x} = 1$ is

- an ellipsoid when $k < c$,
- a hyperboloid of one sheet when $c < k < b$,
- a hyperboloid of two sheets when $b < k < a$,
- imaginary when $a < k$.

As k approaches one of the critical values $\{a, b, c\}$, the surface degenerates to a space curve in one of the three principal planes of the family. For example, as $k \rightarrow c$ we must have $z \rightarrow 0$ and hence the surface $\mathbf{x}^T A_k \mathbf{x} = 1$ degenerates to the space curve defined by

$$\frac{x^2}{a-c} + \frac{y^2}{b-c} = 1 \quad \text{and} \quad z = 0.$$

Note that this is an ellipse in the x, y -plane, called the *focal ellipse* of the system. For values of k near c we obtain either a very thin ellipsoid on the inside of the ellipse ($k < c$) or a very thin hyperboloid of one sheet on the outside of the ellipse ($c < k$). Similarly, as $k \rightarrow b$ or $k \rightarrow a$ the surface $\mathbf{x}^T A_k \mathbf{x} = 1$ degenerates to the space curve

$$\frac{x^2}{a-b} + \frac{z^2}{c-b} = 1 \quad \text{and} \quad y = 0$$

or

$$\frac{y^2}{b-a} + \frac{z^2}{c-a} = 1 \quad \text{and} \quad x = 0,$$

respectively. The first of these is a hyperbola in the x, z -plane, called the *focal hyperbola* of the system. For k on either side of b we obtain very thin hyperboloid of one sheet ($k < b$) or a very thin hyperboloid of two sheets ($b < k$). As $k \rightarrow a$ from the left, the two sheets of the hyperboloid converge to the y, z -plane from opposite sides and then disappear. The focal curve in the y, z -plane is imaginary.

[Remark: The focal **curves** of a confocal family of quadric **surfaces** generalize the focal **points** of a confocal family of quadric **curves**. To see this, one should fix two real numbers $a \neq b$ and consider the family of quadric curves

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

For more on this topic see Chapter 1 of Hilbert and Cohn-Vossen.]

Here is the “fundamental theorem” of the subject. We will prove the theorem for surfaces but it should be clear how to generalize the theorem to confocal quadrics in any number of dimensions.

Fundamental Theorem of Confocal Quadrics. Fix real numbers $a, b, c \in \mathbb{R}$ satisfying $c < b < a$ and consider the confocal family $\mathbf{x}^T A_k \mathbf{x} = 1$ of quadric surfaces. For each point $\mathbf{u} = (u, v, w) \in \mathbb{R}^3$ that is not on a principal plane (i.e., with $u, v, w \neq 0$) there exist exactly three surfaces of the family passing through \mathbf{u} . These surfaces correspond to parameters $k_1, k_2, k_3 \in \mathbb{R}$ satisfying

$$k_1 < c < k_2 < b < k_3 < a,$$

hence there is one surface of each topological type. The tangent planes to the three surfaces at \mathbf{u} are mutually perpendicular, and the *Cartesian coordinates* (u, v, w) are related to the *confocal coordinates* (k_1, k_2, k_3) as follows:

$$\begin{aligned} u^2 &= (a - k_1)(a - k_2)(a - k_3)/(b - a)(c - a) \\ v^2 &= (b - k_1)(b - k_2)(b - k_3)/(a - b)(c - b) \\ w^2 &= (c - k_1)(c - k_2)(c - k_3)/(a - c)(b - c). \end{aligned}$$

///

Before proving this we must derive the equation of the tangent plane at a given point on a quadric surface. So consider a general quadric surface $\mathbf{x}^T A \mathbf{x} = 1$ and let \mathbf{u} be any point on the surface, so that $\mathbf{u}^T A \mathbf{u} = 1$. Observe that

$$\boxed{\mathbf{x}^T A \mathbf{u} = 1}$$

is the equation of some plane passing through \mathbf{u} . I claim that this is the tangent plane.

To see why, we will show that any line $\mathbf{u} + t\mathbf{v}$ contained in the plane $\mathbf{x}^T \mathbf{A}\mathbf{u} = 1$ has at least double contact with the surface at \mathbf{u} . Indeed, since the line is contained in the plane we have for all t that

$$\begin{aligned}(\mathbf{u} + t\mathbf{v})^T \mathbf{A}\mathbf{u} &= 1 \\ \mathbf{u}^T \mathbf{A}\mathbf{u} + t\mathbf{v}^T \mathbf{A}\mathbf{u} &= 1 \\ 1 + t\mathbf{v}^T \mathbf{A}\mathbf{u} &= 1 \\ t\mathbf{v}^T \mathbf{A}\mathbf{u} &= 0,\end{aligned}$$

and it follows that $\mathbf{v}^T \mathbf{A}\mathbf{u} = 0$. Then the intersection of the line with the surface is determined by the following equation in t :

$$\begin{aligned}(\mathbf{u} + t\mathbf{v})^T \mathbf{A}(\mathbf{u} + t\mathbf{v}) &= 1 \\ \mathbf{u}^T \mathbf{A}\mathbf{u} + 2t\mathbf{v}^T \mathbf{A}\mathbf{u} + t^2 \mathbf{v}^T \mathbf{A}\mathbf{v} &= 1 \\ 1 + 0 + t^2 \mathbf{v}^T \mathbf{A}\mathbf{v} &= 1 \\ t^2 \mathbf{v}^T \mathbf{A}\mathbf{v} &= 0.\end{aligned}$$

If $\mathbf{v}^T \mathbf{A}\mathbf{v} = 0$ then we see that the line is completely contained in the surface, and if $\mathbf{v}^T \mathbf{A}\mathbf{v} \neq 0$ then we see that $t = 0$ is a double root as desired. In summary, we find that the plane $\mathbf{x}^T \mathbf{A}\mathbf{u} = 1$ is tangent to the surface $\mathbf{x}^T \mathbf{A}\mathbf{x} = 1$ at $\mathbf{x} = \mathbf{u}$.

Proof of the Fundamental Theorem. We have assumed that $c < b < a$. Let $\mathbf{u} = (u, v, w)$ be any point that is not on a principal plane (i.e., such that $u, v, w \neq 0$). We are looking for values of k such that $\mathbf{u}^T \mathbf{A}_k \mathbf{u} = 1$. In other words, we want

$$\begin{aligned}(u \ v \ w) \begin{pmatrix} 1/(a-k) & & \\ & 1/(b-k) & \\ & & 1/(c-k) \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} &= 1 \\ \frac{u^2}{a-k} + \frac{v^2}{b-k} + \frac{w^2}{c-k} &= 1 \\ (b-k)(c-k)u^2 + (a-k)(c-k)v^2 + (a-k)(b-k)w^2 &= (a-k)(b-k)(c-k).\end{aligned}$$

Therefore we will define the polynomial

$$\varphi_{\mathbf{u}}(k) := (b-k)(c-k)u^2 + (a-k)(c-k)v^2 + (a-k)(b-k)w^2 - (a-k)(b-k)(c-k).$$

We observe that $\varphi_{\mathbf{u}}(k)$ is a cubic polynomial in k with leading coefficient 1 and satisfying

$$\varphi_{\mathbf{u}}(c) = (a-c)(b-c)w^2 > 0,$$

$$\begin{aligned}\varphi_{\mathbf{u}}(b) &= (a-b)(c-b)v^2 < 0, \\ \varphi_{\mathbf{u}}(a) &= (b-a)(c-a)u^2 > 0.\end{aligned}$$

It follows that $\varphi_u(k)$ has three distinct real roots k_1, k_2, k_3 satisfying

$$k_1 < c < k_2 < b < k_3 < a.$$

By previous remarks we know that the surfaces corresponding to k_1, k_2, k_3 are an ellipsoid, a hyperboloid of one sheet and a hyperboloid of two sheets, respectively. Thus we have found one surface of each topological type passing through \mathbf{u} .

By the remarks after the statement of the theorem, the tangent planes to these three surfaces at \mathbf{u} have the equations

$$\mathbf{x}^T A_{k_1} \mathbf{u} = 1, \quad \mathbf{x}^T A_{k_2} \mathbf{u} = 1 \quad \text{and} \quad \mathbf{x}^T A_{k_3} \mathbf{u} = 1,$$

respectively. To show that these planes are mutually perpendicular it is enough to show that the normal vectors $A_{k_1} \mathbf{u}$, $A_{k_2} \mathbf{u}$ and $A_{k_3} \mathbf{u}$ are mutually perpendicular, and for this we will use a clever trick. Observe that for any real numbers $k \neq \ell$ with $k, \ell \notin \{a, b, c\}$ we have the following ‘‘partial fractions’’ identity:

$$\boxed{A_k A_\ell = \frac{A_k - A_\ell}{k - \ell}.}$$

Now if $k \neq \ell$ are two elements of the set $\{k_1, k_2, k_3\}$ then by definition we have $\mathbf{u}^T A_k \mathbf{u} = 1$ and $\mathbf{u}^T A_\ell \mathbf{u} = 1$, and it follows that the vectors $A_k \mathbf{u}$ and $A_\ell \mathbf{u}$ are perpendicular:

$$(A_k \mathbf{u})^T (A_\ell \mathbf{u}) = \mathbf{u}^T (A_k A_\ell) \mathbf{u} = \frac{\mathbf{u}^T (A_k - A_\ell) \mathbf{u}}{k - \ell} = \frac{\mathbf{u}^T A_k \mathbf{u} - \mathbf{u}^T A_\ell \mathbf{u}}{k - \ell} = \frac{1 - 1}{k - \ell} = 0.$$

It only remains to solve for the Cartesian coordinates u, v, w in terms of the confocal coordinates k_1, k_2, k_3 . Since the cubic polynomial $\varphi_{\mathbf{u}}(k)$ has leading coefficient 1 and distinct roots k_1, k_2, k_3 we must have

$$\varphi_{\mathbf{u}}(k) = (k - k_1)(k - k_2)(k - k_3).$$

Then by substituting $k = a, b, c$ into $\varphi_{\mathbf{u}}(k)$ we obtain

$$\begin{aligned}(b-a)(c-a)u^2 &= \varphi_{\mathbf{u}}(a) = (a-k_1)(a-k_2)(a-k_3), \\ (a-b)(c-b)v^2 &= \varphi_{\mathbf{u}}(b) = (b-k_1)(b-k_2)(b-k_3), \\ (a-c)(b-c)w^2 &= \varphi_{\mathbf{u}}(c) = (c-k_1)(c-k_2)(c-k_3),\end{aligned}$$

as desired. □

5 Symmetries of a Tangent Cone

The final ingredient that we need is the result stated in the footnote of Hilbert and Cohn-Vossen that I alluded to in the Introduction. This fact was apparently well known in the early twentieth century⁸ but I have not been able to find a proof in the literature. The selection of topics in this note was chosen to make the proof as slick as possible. Here is the footnote.

Footnote 4 from Hilbert and Cohn-Vossen (pg. 24). *The following is another property of the confocal system, which, incidentally, includes the property just mentioned⁹ as a limiting case: The planes of symmetry of the tangent cone from any point \mathbf{u} in space to any surface of the system which does not enclose \mathbf{u} are the tangent planes at \mathbf{u} to the three surfaces of the system that pass through \mathbf{u} .*

To express this in our language we fix distinct real numbers $c < b < a$ and consider the confocal system of quadric surfaces $\mathbf{x}^T A_k \mathbf{x} = 1$ as in Section 4. For a generic point in space $\mathbf{u} \in \mathbb{R}^3$ (i.e., not on the principal planes) recall from the Fundamental Theorem that there exist three confocal surfaces through \mathbf{u} corresponding to some “confocal parameters” $k_1, k_2, k_3 \in \mathbb{R}$ satisfying

$$k_1 < c < k_2 < b < k_3 < a.$$

Furthermore, recall that tangent planes to the three confocal surfaces at \mathbf{u} are mutually perpendicular and are given by the equations

$$\mathbf{x}^T A_{k_1} \mathbf{u} = 1, \quad \mathbf{x}^T A_{k_2} \mathbf{u} = 1 \quad \text{and} \quad \mathbf{x}^T A_{k_3} \mathbf{u} = 1.$$

Here is the generic case of Hilbert and Cohn-Vossen’s statement.

Theorem. Let $\mathbf{u} \in \mathbb{R}^3$ be a generic point (i.e., with nonzero coordinates). Fix any generic confocal parameter $\ell \in \mathbb{R} \setminus \{a, b, c, k_1, k_2, k_3\}$ and consider the tangent cone from the point \mathbf{u} to the confocal surface $\mathbf{x}^T A_\ell \mathbf{x} = 1$. We know from Section 2 that this cone has equation $(\mathbf{x} - \mathbf{u})^T K_{\mathbf{u}, \ell} (\mathbf{x} - \mathbf{u}) = 0$ where the cone matrix is defined by

$$K_{\mathbf{u}, \ell} := A_\ell \mathbf{u} \mathbf{u}^T A_\ell + (1 - \mathbf{u} A_\ell \mathbf{u}^T) A_\ell.$$

I claim that this matrix has eigenvectors $A_{k_1} \mathbf{u}$, $A_{k_2} \mathbf{u}$ and $A_{k_3} \mathbf{u}$, with corresponding eigenvalues

$$\begin{aligned} \lambda_1 &= (\ell - k_2)(\ell - k_3) / (a - \ell)(b - \ell)(c - \ell), \\ \lambda_2 &= (\ell - k_1)(\ell - k_3) / (a - \ell)(b - \ell)(c - \ell), \\ \lambda_3 &= (\ell - k_1)(\ell - k_2) / (a - \ell)(b - \ell)(c - \ell). \end{aligned}$$

⁸Hilbert and Cohn-Vossen refer to the work of Otto Staude but they don’t provide a reference. It seems that Staude’s work is summarized in his textbook (1896).

⁹Their “property just mentioned” was a geometric description of our Main Theorem.

Since the parameters k_1, k_2, k_3 are distinct we observe that the (nonzero) eigenvalues $\lambda_1, \lambda_2, \lambda_3$ are also distinct. If the eigenvalues are not all of the same sign (i.e., if $k_1 < \ell < k_3$) then we conclude that $(\mathbf{x} - \mathbf{u})^T K_{\mathbf{u}, \ell} (\mathbf{x} - \mathbf{u}) = 0$ is a real non-circular cone with planes of symmetry equal to the tangent planes $\mathbf{x}^T A_{k_1} \mathbf{u} = 1$, $\mathbf{x}^T A_{k_2} \mathbf{u} = 1$ and $\mathbf{x}^T A_{k_3} \mathbf{u} = 1$ through \mathbf{u} . [Remark: It is interesting that that the planes of symmetry depend only on the point \mathbf{u} and not on the parameter ℓ .] ///

Proof. Consider any $i \in \{1, 2, 3\}$. To prove that $A_{k_i} \mathbf{u}$ is an eigenvector of $K_{\mathbf{u}, \ell}$ we will use the partial fractions identity

$$A_\ell A_{k_i} = \frac{A_\ell - A_{k_i}}{\ell - k_i},$$

which holds because $\ell \neq k_i$ and $\ell, k_i \notin \{a, b, c\}$. Since \mathbf{u} is on the surface $\mathbf{x}^T A_{k_i} \mathbf{x} = 1$ (i.e., $\mathbf{u}^T A_{k_i} \mathbf{u} = 1$) we have

$$\begin{aligned} K_{\mathbf{u}, \ell} A_{k_i} \mathbf{u} &= [A_\ell \mathbf{u} \mathbf{u}^T A_\ell + (1 - \mathbf{u} A_\ell \mathbf{u}^T) A_\ell] A_{k_i} \mathbf{u} \\ &= A_\ell \mathbf{u} \mathbf{u}^T A_\ell A_{k_i} \mathbf{u} + (1 - \mathbf{u} A_\ell \mathbf{u}^T) A_\ell A_{k_i} \mathbf{u} \\ &= A_\ell \mathbf{u} \mathbf{u}^T \left(\frac{A_\ell - A_{k_i}}{\ell - k_i} \right) \mathbf{u} + (1 - \mathbf{u} A_\ell \mathbf{u}^T) \left(\frac{A_\ell - A_{k_i}}{\ell - k_i} \right) \mathbf{u} \\ &= A_\ell \mathbf{u} \left(\frac{\mathbf{u}^T A_\ell \mathbf{u} - \mathbf{u}^T A_{k_i} \mathbf{u}}{\ell - k_i} \right) + (1 - \mathbf{u} A_\ell \mathbf{u}^T) \left(\frac{A_\ell \mathbf{u} - A_{k_i} \mathbf{u}}{\ell - k_i} \right) \\ &= A_\ell \mathbf{u} \left(\frac{\mathbf{u}^T A_\ell \mathbf{u} - 1}{\ell - k_i} \right) + (1 - \mathbf{u} A_\ell \mathbf{u}^T) \left(\frac{A_\ell \mathbf{u} - A_{k_i} \mathbf{u}}{\ell - k_i} \right) \\ &= \cancel{A_\ell \mathbf{u} \left(\frac{\mathbf{u}^T A_\ell \mathbf{u} - 1}{\ell - k_i} \right)} + \left(\frac{1 - \mathbf{u}^T A_\ell \mathbf{u}}{\ell - k_i} \right) A_\ell \mathbf{u} - \left(\frac{1 - \mathbf{u}^T A_\ell \mathbf{u}}{\ell - k_i} \right) A_{k_i} \mathbf{u} \\ &= \left(\frac{\mathbf{u}^T A_\ell \mathbf{u} - 1}{\ell - k_i} \right) A_{k_i} \mathbf{u}. \end{aligned}$$

It follows that $A_{k_i} \mathbf{u}$ is an eigenvector of $K_{\mathbf{u}, \ell}$ with eigenvalue $\lambda_i = (\mathbf{u}^T A_\ell \mathbf{u} - 1) / (\ell - k_i)$. To compute the eigenvalue explicitly we recall from the proof of the Fundamental Theorem that the cubic polynomial $\varphi_{\mathbf{u}}(k) := (a - k)(b - k)(c - k)(\mathbf{u}^T A_k \mathbf{u} - 1)$ has distinct roots k_1, k_2, k_3 and hence

$$\begin{aligned} (a - k)(b - k)(c - k)(\mathbf{u}^T A_k \mathbf{u} - 1) &= (k - k_1)(k - k_2)(k - k_3) \\ \mathbf{u}^T A_k \mathbf{u} - 1 &= (k - k_1)(k - k_2)(k - k_3) / (a - k)(b - k)(c - k). \end{aligned}$$

By substituting $k = \ell$ we obtain the eigenvalue

$$\lambda_i = \frac{\mathbf{u}^T A_\ell \mathbf{u} - 1}{\ell - k_i} = \frac{(\ell - k_1)(\ell - k_2)(\ell - k_3)}{(a - \ell)(b - \ell)(c - \ell)(\ell - k_i)},$$

which agrees with the claimed formulas for $\lambda_1, \lambda_2, \lambda_3$ when $i = 1, 2, 3$. □

To complete the proof of the Main Theorem it only remains to examine what happens when the point $\mathbf{u} \in \mathbb{R}^3$ approaches a principal plane and when the confocal parameter $\ell \in \mathbb{R}$ approaches one of the critical values $\{a, b, c\}$.

6 Proof of the Main Theorem

In the previous section we computed the symmetries of the tangent cone from a generic point $\mathbf{u} \in \mathbb{R}^3$ to a generic surface $\mathbf{x}^T A_\ell \mathbf{x} = 1$ in the confocal family corresponding to the fixed parameters $c < b < a$. If $k_1 < c < k_2 < b < k_3 < a$ are the confocal coordinates of the point \mathbf{u} , we found that the tangent cone is real when $k_1 < \ell < k_3$ and that it is never circular.

In this section we will complete the proof of the Main Theorem by allowing the point \mathbf{u} to approach the principal planes and by allowing the parameter ℓ to approach one of the critical values $\{a, b, c\}$, i.e., by allowing the quadric surface $\mathbf{x}^T A_\ell \mathbf{x} = 1$ to degenerate to one of the *focal curves* of the confocal system:

$$\begin{aligned} x^2/(a-c) + y^2/(b-c) &= 1 & \text{and } z &= 0, \\ x^2/(a-b) + z^2/(c-b) &= 1 & \text{and } y &= 0, \\ y^2/(b-a) + z^2/(c-a) &= 1 & \text{and } x &= 0. \end{aligned}$$

Since $c < b < a$ we see that the first of these curves is an ellipse in the x, y -plane, the second is a hyperbola in the x, z -plane, and the third is an imaginary curve in the y, z -plane.

Letting the surface $\mathbf{x}^T A_\ell \mathbf{x} = 1$ approach a focal curve.

Let $\mathbf{u} \in \mathbb{R}^3$ be a fixed point not on a principal plane, with confocal coordinates $k_1 < c < k_2 < b < k_3 < a$. Let $(\mathbf{x} - \mathbf{u})^T K_{\mathbf{u}, \ell} (\mathbf{x} - \mathbf{u}) = 0$ be the tangent cone from the point \mathbf{u} to the surface $\mathbf{x}^T A_\ell \mathbf{x} = 1$. From the previous section we know that the matrix $K_{\mathbf{u}, \ell}$ has eigenvectors $A_{k_1} \mathbf{u}$, $A_{k_2} \mathbf{u}$ and $A_{k_3} \mathbf{u}$, with corresponding eigenvalues

$$\begin{aligned} \lambda_1 &= (\ell - k_2)(\ell - k_3) / (a - \ell)(b - \ell)(c - \ell), \\ \lambda_2 &= (\ell - k_1)(\ell - k_3) / (a - \ell)(b - \ell)(c - \ell), \\ \lambda_3 &= (\ell - k_1)(\ell - k_2) / (a - \ell)(b - \ell)(c - \ell). \end{aligned}$$

Note that the **eigenvectors** are independent of the parameter ℓ . The **eigenvalues** become undefined as ℓ approaches one of the critical values $\{a, b, c\}$, however this is easy to fix.

To see what happens as $\ell \rightarrow c$ (i.e., as the surface $\mathbf{x}^T A_\ell \mathbf{x} = 1$ degenerates to the focal ellipse) we observe that the matrix $(c - \ell)K_{\mathbf{u}, \ell}$ has the same eigenvectors as $K_{\mathbf{u}, \ell}$ but with eigenvalues

$$\begin{aligned} \lambda_1 &= (\ell - k_2)(\ell - k_3) / (a - \ell)(b - \ell), \\ \lambda_2 &= (\ell - k_1)(\ell - k_3) / (a - \ell)(b - \ell), \\ \lambda_3 &= (\ell - k_1)(\ell - k_2) / (a - \ell)(b - \ell). \end{aligned}$$

Since these eigenvalues are well-defined when $\ell \rightarrow c$, we conclude that the matrix $K_{\mathbf{u},c} := \lim_{\ell \rightarrow c} (c - \ell)K_{\mathbf{u},\ell}$ exists¹⁰ and is uniquely determined by having eigenvectors $A_{k_1}\mathbf{u}, A_{k_2}\mathbf{u}, A_{k_3}\mathbf{u}$ with corresponding eigenvalues

$$\begin{aligned}\lambda_1 &= (c - k_2)(c - k_3) / (a - c)(b - c), \\ \lambda_2 &= (c - k_1)(c - k_3) / (a - c)(b - c), \\ \lambda_3 &= (c - k_1)(c - k_2) / (a - c)(b - c).\end{aligned}$$

Now the equation $(\mathbf{x} - \mathbf{u})^T K_{\mathbf{u},c} (\mathbf{x} - \mathbf{u}) = 0$ defines the cone from the point \mathbf{u} to the focal ellipse. Since $k_1 < c < k_2 < k_3$ we find that the eigenvalues of $K_{\mathbf{u},c}$ satisfy

$$\lambda_1 > 0 > \lambda_2 > \lambda_3.$$

It follows that the cone from a generic point \mathbf{u} to the focal ellipse is real and non-circular. Below we will see what happens when \mathbf{u} is non-generic.

Similarly we can define the matrices $K_{\mathbf{u},b} := \lim_{\ell \rightarrow b} (b - \ell)K_{\mathbf{u},\ell}$ and $K_{\mathbf{u},a} := \lim_{\ell \rightarrow a} (a - \ell)K_{\mathbf{u},\ell}$, which both have the same eigenvectors $A_{k_1}\mathbf{u}, A_{k_2}\mathbf{u}$ and $A_{k_3}\mathbf{u}$. The eigenvalues of $K_{\mathbf{u},b}$ are

$$\begin{aligned}\lambda_1 &= (b - k_2)(b - k_3) / (a - b)(c - b), \\ \lambda_2 &= (b - k_1)(b - k_3) / (a - b)(c - b), \\ \lambda_3 &= (b - k_1)(b - k_2) / (a - b)(c - b),\end{aligned}$$

which satisfy

$$\lambda_3 < 0 < \lambda_1 < \lambda_2.$$

It follows that the cone $(\mathbf{x} - \mathbf{u})^T K_{\mathbf{u},b} (\mathbf{x} - \mathbf{u}) = 0$ from a generic point \mathbf{u} to the focal hyperbola is real and non-circular.

Finally we observe that the eigenvalues of $K_{\mathbf{u},a}$ are

$$\begin{aligned}\lambda_1 &= (a - k_2)(a - k_3) / (b - a)(c - a), \\ \lambda_2 &= (a - k_1)(a - k_3) / (b - a)(c - a), \\ \lambda_3 &= (a - k_1)(a - k_2) / (b - a)(c - a),\end{aligned}$$

which satisfy

$$0 < \lambda_1 < \lambda_2 < \lambda_3.$$

Thus the cone $(\mathbf{x} - \mathbf{u})^T K_{\mathbf{u},a} (\mathbf{x} - \mathbf{u}) = 0$ from a generic point \mathbf{u} to the imaginary focal curve is imaginary, as expected.

Letting the point \mathbf{u} approach a principal plane.

¹⁰Unfortunately it seems that this matrix does not have a nice closed formula.

Fix real numbers $c < b < a$ as before and recall from the Fundamental Theorem that each generic point $\mathbf{u} = (u, v, w) \in \mathbb{R}^3$ is contained in three (mutually perpendicular) confocal quadric surfaces corresponding to some parameters

$$k_1 < c < k_2 < b < k_3 < a.$$

Conversely, any real numbers k_1, k_2, k_3 satisfying these inequalities correspond to three confocal surfaces that intersect (perpendicularly) at the eight points $\mathbf{u} = (u, v, w)$ defined by

$$\begin{aligned} u^2 &= (a - k_1)(a - k_2)(a - k_3) / (b - a)(c - a) \\ v^2 &= (b - k_1)(b - k_2)(b - k_3) / (a - b)(c - b) \\ w^2 &= (c - k_1)(c - k_2)(c - k_3) / (a - c)(b - c). \end{aligned}$$

Thus we observe that the point \mathbf{u} approaches the three principal planes precisely when the parameters k_1, k_2, k_3 approach the critical values c, b, a from the left:

$$\begin{aligned} u \rightarrow 0 &\Leftrightarrow k_1 \rightarrow c \text{ from the left,} \\ v \rightarrow 0 &\Leftrightarrow k_2 \rightarrow b \text{ from the left,} \\ w \rightarrow 0 &\Leftrightarrow k_3 \rightarrow a \text{ from the left.} \end{aligned}$$

As long as the values k_1, k_2, k_3 remain distinct we find that the matrix $K_{\mathbf{u}, \ell}$ of the cone from \mathbf{u} to any surface $\mathbf{x}^T A_\ell \mathbf{x} = 1$ (including the degenerate cases when $\ell \in \{a, b, c\}$) has distinct eigenvalues $\lambda_1, \lambda_2, \lambda_3$, and thus it remains non-circular.

Under what conditions do we get a circular cone? In other words, under what conditions does the cone matrix $K_{\mathbf{u}, \ell}$ have a repeated eigenvalue? As the confocal parameters k_1, k_2, k_3 move around, we observe from the explicit formulas for the eigenvalues $\lambda_1, \lambda_2, \lambda_3$ that

$$\lim(\lambda_i - \lambda_j) = 0 \quad \Leftrightarrow \quad \lim(k_i - k_j) = 0.$$

In other words, the cone from \mathbf{u} to a confocal surface or focal curve becomes circular precisely when two of the confocal parameters k_1, k_2, k_3 approach each other. Since the parameters of a generic point satisfy

$$k_1 < c < k_2 < b < k_3 < a,$$

we see that it is **impossible** for k_1 and k_3 to approach each other. Thus we have two cases:

Case 1: The point \mathbf{u} approaches the focal hyperbola. As $k_2 \rightarrow b \leftarrow k_3$ we find in the limit that the coordinates of the point $\mathbf{u} = (u, v, w)$ satisfy

$$\begin{aligned} u^2 &= (a - k_1)(a - b)(a - b) / (b - a)(c - a) = (a - b)(a - k_1) / (a - c), \\ v^2 &= (b - k_1)(b - b)(b - b) / (a - b)(c - b) = 0, \\ w^2 &= (c - k_1)(c - b)(c - b) / (a - c)(b - c) = -(c - b)(c - k_1) / (a - c). \end{aligned}$$

This implies that

$$u^2 / (a - b) + w^2 / (c - b) = 1 \quad \text{and} \quad v = 0,$$

which tells us that \mathbf{u} is on the focal hyperbola of the system. At the same time, the eigenvalues of the cone from \mathbf{u} to the surface $\mathbf{x}^T A_\ell \mathbf{x} = 1$ approach the values

$$\begin{aligned}\lambda_1 &= (\ell - b)(\ell - b) / (a - \ell)(b - \ell)(c - \ell) = (b - \ell) / (a - \ell)(c - \ell), \\ \lambda_2 &= (\ell - k_1)(\ell - b) / (a - \ell)(b - \ell)(c - \ell) = (k_1 - \ell) / (a - \ell)(c - \ell), \\ \lambda_3 &= (\ell - k_1)(\ell - b) / (a - \ell)(b - \ell)(c - \ell) = (k_1 - \ell) / (a - \ell)(c - \ell).\end{aligned}$$

Since $k_1 < c < b$ implies $k_1 \neq b$ we observe that $\lambda_1 \neq \lambda_2 = \lambda_3$. The cone is real precisely when $k_1 < \ell < b$ and in this case the Theorem at the end of Section 2 says that the cone is **circular** with angle of aperture θ satisfying

$$\cos^2 \theta = \frac{\lambda_3}{\lambda_3 - \lambda_1} = \frac{k_1 - \ell}{k_1 - b}.$$

We can view the point \mathbf{u} locally as a function of the parameter k_1 , which satisfies $k_1 < \min\{c, \ell\}$. As $-\infty \leftarrow k_1$ the point \mathbf{u} on the focal hyperbola goes to infinity and we have $\cos^2 \theta \rightarrow 1$, or $\theta \rightarrow 0$. That is, from infinitely far away the surface $\mathbf{x}^T A_\ell \mathbf{x} = 1$ looks like a point. If $\ell < c$ (i.e., if the surface $\mathbf{x}^T A_\ell \mathbf{x} = 1$ is an ellipsoid) then as $k_1 \rightarrow \ell$ the point \mathbf{u} approaches an ‘‘umbilic point’’ on the surface of the ellipsoid and the (circular) tangent cone flattens out into the tangent plane at the umbilic point. In the limiting case $\ell \rightarrow c$, the point \mathbf{u} approaches one of the foci $(x, y, z) = (\pm\sqrt{a - b}, 0, 0)$ of the focal ellipse. If $c < \ell$ (i.e., if the surface $\mathbf{x}^T A_\ell \mathbf{x} = 1$ is a hyperboloid of one sheet) then as $k_1 \rightarrow c$ we have $\cos^2 \theta \rightarrow (c - \ell) / (c - b)$ and the angle of aperture reaches the **maximum** value $\theta = \arccos \sqrt{(c - \ell) / (c - b)}$.

In all cases, the axis of symmetry of the cone is given by the eigenvector

$$A_{k_1} \mathbf{u} = \left(\frac{u}{a - k_1}, \frac{v}{b - k_1}, \frac{w}{c - k_1} \right) = \left(\frac{u}{a - k_1}, 0, \frac{w}{c - k_1} \right),$$

which I claim is **tangent** to the focal hyperbola at the point \mathbf{u} . Indeed, let us view the point $\mathbf{u} = (u, v, w)$ locally as a function of the parameter k_1 . By differentiating the formula for u^2 with respect to k_1 we obtain

$$2uu' = (u^2)' = [(a - b)(a - k_1) / (a - c)]' = -(a - b) / (a - c) = -u^2 / (a - k_1),$$

and hence $u' = (-1/2) \cdot u / (a - k_1)$. Similarly we see that $v' = (-1/2) \cdot v / (b - k_1)$ and $w' = (-1/2) \cdot w / (c - k_1)$, thus the tangent vector to the focal hyperbola at \mathbf{u} is given by

$$(u', v', w') = \left(\frac{-u}{2(a - k_1)}, \frac{-v}{2(b - k_1)}, \frac{-w}{2(c - k_1)} \right) = -\frac{1}{2} A_{k_1} \mathbf{u}$$

as desired. ///

Case 2: The point \mathbf{u} approaches the focal ellipse. As $k_1 \rightarrow c \leftarrow k_2$ we find in the limit that the coordinates of the point $\mathbf{u} = (u, v, w)$ satisfy

$$\begin{aligned}u^2 &= (a - c)(a - c)(a - k_3) / (b - a)(c - a) = (a - c)(a - k_3) / (a - b), \\ v^2 &= (b - c)(b - c)(b - k_3) / (a - b)(c - b) = -(b - c)(b - k_3) / (a - b), \\ w^2 &= (c - c)(c - c)(c - k_3) / (a - c)(b - c) = 0.\end{aligned}$$

This implies that

$$u^2/(a-c) + v^2/(b-c) = 1 \quad \text{and} \quad w = 0,$$

which tells us that \mathbf{u} is on the focal ellipse of the system. At the same time, the eigenvalues of the cone from \mathbf{u} to the surface $\mathbf{x}^T A_\ell \mathbf{x} = 1$ approach the values

$$\begin{aligned} \lambda_1 &= (\ell - c)(\ell - k_3)/(a - \ell)(b - \ell)(c - \ell) = (k_3 - \ell)/(a - \ell)(b - \ell), \\ \lambda_2 &= (\ell - c)(\ell - k_3)/(a - \ell)(b - \ell)(c - \ell) = (k_3 - \ell)/(a - \ell)(b - \ell), \\ \lambda_3 &= (\ell - c)(\ell - c)/(a - \ell)(b - \ell)(c - \ell) = (c - \ell)/(a - \ell)(b - \ell). \end{aligned}$$

Since $c < b < k_3$ implies $c \neq k_3$ we observe that $\lambda_1 = \lambda_2 \neq \lambda_3$. The cone is real precisely when $c < \ell < k_3$ and in this case the Theorem at the end of Section 2 says that the cone is **circular** with angle of aperture θ satisfying

$$\cos^2 \theta = \frac{\lambda_1}{\lambda_1 - \lambda_3} = \frac{k_3 - \ell}{k_3 - c}.$$

We can view the point \mathbf{u} locally as a function of the parameter k_3 , which satisfies $\max\{b, \ell\} < k_3 < a$. If $b < \ell$ (i.e., if the surface $\mathbf{x}^T A_\ell \mathbf{x} = 1$ is a hyperboloid of two sheets) then as $\ell \leftarrow k_3$ the point \mathbf{u} approaches an ‘‘umbilic point’’ on the surface and the (circular) tangent cone flattens out into the tangent plane at the umbilic point. If $\ell < b$ (i.e., if the surface $\mathbf{x}^T A_\ell \mathbf{x} = 1$ is a hyperboloid of one sheet) then as $b \leftarrow k_3$ we have $\cos^2 \theta \rightarrow (b - \ell)/(b - c)$ and the angle of aperture reaches the **maximum** value $\theta = \arccos \sqrt{(b - \ell)/(b - c)}$. In the limiting case $\ell \rightarrow b$ the point \mathbf{u} approaches one of the foci $(x, y, z) = (\pm\sqrt{a - c}, 0, 0)$ of the focal hyperbola. For any value of ℓ , the angle of aperture reaches the **minimum** value $\theta = \arccos \sqrt{(a - \ell)/(a - c)}$ when $k_3 \rightarrow a$, i.e., when \mathbf{u} approaches one of the points $(x, y, z) = (0, \pm\sqrt{b - c}, 0)$. (Since the point \mathbf{u} is trapped on an ellipse it can’t get infinitely far away.)

Finally, the axis of symmetry of the cone is given by the eigenvector $A_{k_3} \mathbf{u}$. Using a similar argument to the previous case we see that this axis is tangent to the focal ellipse at \mathbf{u} . ///

These results hold for any value of ℓ as long as the corresponding cone is real. The Main Theorem is just a summary of these results for the cases when $\ell \in \{a, b, c\}$.

7 Conclusion

To conclude the note I will answer the three questions from the Introduction in plain language. Let us consider a central and non-degenerate quadric curve in the real x, y -plane:

$$\frac{x^2}{\alpha} + \frac{y^2}{\beta} = 1.$$

We assume that the parameters $\alpha, \beta \in \mathbb{R}$ satisfy $\alpha > \beta$ and are not both negative. Thus our curve is a non-circular ellipse (when $\alpha > \beta > 0$) or a non-rectangular hyperbola (when $\alpha > 0 > \beta$). If we define $(a, b, c) := (\alpha, \beta, 0)$ then our curve becomes

$$\frac{x^2}{a-c} + \frac{y^2}{b-c} = 1,$$

which we identify as either the focal ellipse or the focal hyperbola of a certain family of confocal quadric surfaces in x, y, z -space.

Question 1: From which points in space does our curve look like a circle?

Answer: It looks like a circle from points on the the other real focal curve defined by:

$$\frac{x^2}{a-b} + \frac{z^2}{c-b} = \frac{x^2}{\alpha-\beta} + \frac{z^2}{-\beta} = 1 \quad \text{and} \quad y = 0.$$

If our curve is an ellipse/hyperbola in the x, y -plane then the other focal curve is a hyperbola/ellipse in the x, z -plane, passing through the foci of the original curve.

Question 2: In which direction should we look to see the circle?

Answer: If we are sitting on the focal curve in the x, z -plane then we should look in the direction of the tangent line. The focal curve in the x, y -plane then looks like a circle centered on this line.

Question 3: How big is the circle?

Answer: The apparent size of the circle depends on the angle of aperture θ of the corresponding circular cone.

Suppose our curve is an ellipse ($\alpha > \beta > 0$) and that \mathbf{u} lies on the focal hyperbola. As \mathbf{u} approaches one of the foci $(x, y, z) = (\pm\sqrt{\alpha-\beta}, 0, 0)$ of the ellipse, the cone becomes flat and the ellipse looks like an infinitely big circle. As \mathbf{u} goes to infinity the ellipse looks like an infinitesimally small circle.

On the other hand, suppose that our curve is a hyperbola ($\alpha > 0 > \beta$) and that \mathbf{u} lies on the confocal ellipse. As \mathbf{u} approaches one of the foci $(x, y, z) = (\pm\sqrt{\alpha-\beta}, 0, 0)$ of the hyperbola, the cone becomes flat and the hyperbola looks like an infinitely big circle. As \mathbf{u} approaches one of the points $(x, y, z) = (0, 0, \pm\sqrt{-\beta})$ the angle of aperture reaches the **minimum** value $\theta = \arccos \sqrt{\alpha/(\alpha-\beta)}$.

References

- Elton, John H. (2009). *Indefinite quadratic forms and the invariance of the interval in Special Relativity*. <https://arxiv.org/abs/0904.3913>
- Hilbert, David and Cohn-Vossen, Stefan (1932). *Geometry and the Imagination*. English translation by P. Nemenyi published by AMS Chelsea.
- Sommerville, D.M.Y. (1934). *Analytical Geometry of Three Dimensions*. Cambridge University Press.
- Spain, Barry (1960). *Analytical Quadrics*. Pergamon Student Editions.
- Staude, Otto (1896). *Die Focaleigenschaften der Flächen zweiter Ordnung. Ein neues Kapitel zu den Lehrbüchern der analytischen Geometrie des Raumes*. Leipzig, Teubner.