The Beauty of Conics

Damjan Kobal
Department of Mathematics
FMF, University of Ljubljana
Slovenia
Damjan.Kobal@fmf.uni-lj.si

Invited Special Lecture
5^{rd} International Conference On Mathematics Education
AIMER
19 - 22 December, 2019
Indore, India

Abstract

The conics are likely one of the most inspiring chapters of geometry and elementary mathematics. Likely, they are also one of the most unjustifiably neglected chapters in the teaching of mathematics. The story of conics starts with the simple and intuitive notion of a 'distance' between the simplest geometric objects (points, lines, circles). Modest and careful observations (Leonardo da Vinci's Saper Vedere) then unveils conics in many different settings. Throughout 'this conics story' we will use only very elementary and intuitive geometry to learn a lot about conics. We shall encounter simple and inspiring teaching ideas, rich historical mathematical insights, persuasive intuitive proving concepts, motivational technology and mathematical applicability. Archimedes (more than 2200 years ago) was able to calculate the area of a parabola section – which would be quite a hard task even for many of today's users of Leibnitz and Newton's calculus – and Archimedes did it with mind-boggling simplicity. And after we comprehend Archimedes reasoning, we even appreciate and understand deeper what seems to be a very simple formula for an area of a triangle.

Key words: distance, parabola, ellipse, hyperbola

1 Introduction

Throughout this 'conics story' we will use very intuitive approach and simple 'hands on' introductory tasks to introduce and explain classical definitions of the conics and connect those definitions with even less known and more universal properties, which define and interconnect the conics. The reader is advised to follow the dynamic visualizations of the GeoGebra book 'The Beauty of Conics' at https://www.geogebra.org/m/phtegwnp.

2 Parabola by paper folding

Let us start with a simple task.

Hands on activity 1. *Take a regular sheet of paper and choose any point F somewhere close to one of the longer edges (Figure 1).*

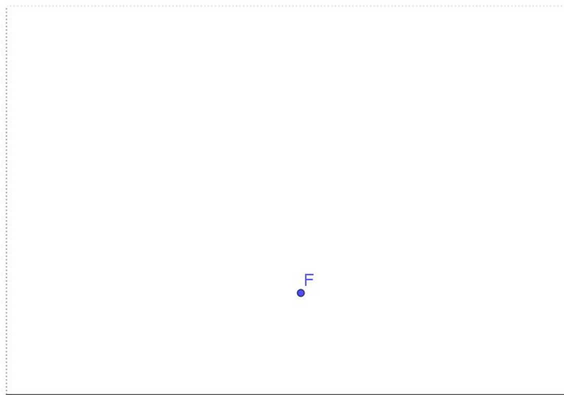

Figure 1: Sheet of paper with point F

Fold the paper several different ways ... so that in any of the folds the bottom edge meets point F.

After folding the paper several times the folds form an obvious 'parabola shape' (Figure 2 - see https://www.geogebra.org/m/phtegwnp#material/uDaaz8kc).

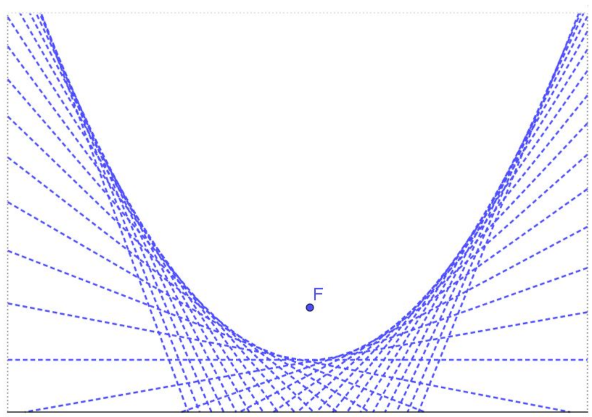

Figure 2: Sheet of the paper folded many times so that the bottom edge meets point F

Is the curve indicated by these folds really a parabola?

Definition 2. *Parabola $\mathcal{P}(p, F)$ given by line p and point $F \notin p$ is the set of all points T, which are equidistant from line p and point F. Line p and point F are respectively called directrix and focus of parabola $\mathcal{P}(p, F)$.*

Corollary 3. [Geometric (Archimedean) construction of a parabola]

- *We start with a given directrix p and point $F \notin p$.*
- *For any point $A \in p$, we draw a line through A perpendicular to p.*

- *For the same point A we draw a perpendicular bisector of AF.*
- *The two lines intersect at point T. Since T was obtained by perpendicular bisector of AF, the triangle ΔAFT is isosceles. Thus $AT = TF$. Since $AT \perp p$, by definition point T must be on the parabola.*
- *As point A travels along directrix p, point T travels along the parabola $\mathcal{P}(p, F)$.*

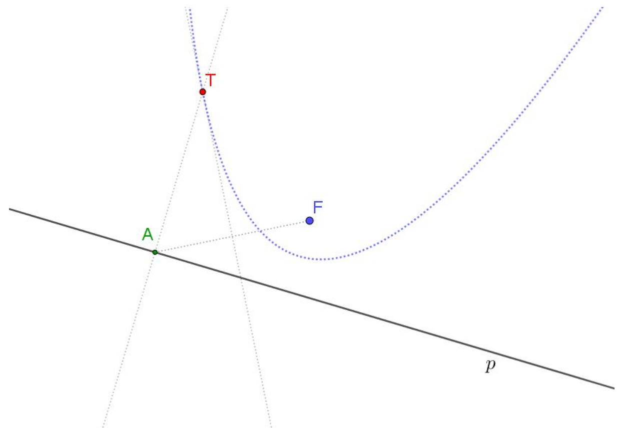

Figure 3: Geometric construction of a parabola

Theorem 4. *Definition of a parabola given by Definition 2 coincides, or better generalizes the definition of a parabola within the classical coordinate system given by a polynomial of the second degree $p(x) = ax^2 + bx + c$. Considering well known coordinate translations of the form $f(x) = a(x-p)^2 + q$ it is easy to prove, that (analytically given) parabola $y = ax^2$ is the same as parabola defined geometrically by directrix $y = -\frac{1}{4a}$ and focus $F(0, \frac{1}{4a})$. Or switching the axes, parabola given analytically by equation $x = ay^2$ is the parabola given geometrically by directrix $x = -\frac{1}{4a}$ and focus $F(\frac{1}{4a}, 0)$.*

Proof. Let us show, that parabola $y = ax^2$ has directrix $y = -\frac{1}{4a}$ and focus $F(0, \frac{1}{4a})$. Any point on the parabola is of the form (x, ax^2) and its distance from point $F(0, \frac{1}{4a})$ is equal to $\sqrt{x^2 + (ax^2 - \frac{1}{4a})^2}$, while its distance from the line $y = -\frac{1}{4a}$ is $ax^2 + \frac{1}{4a}$.

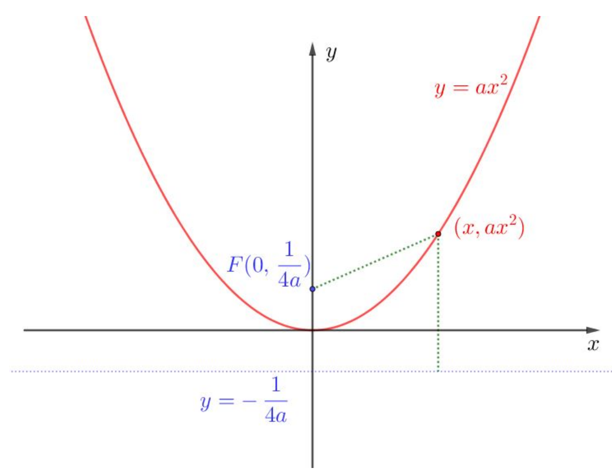

Figure 4: Focus and directrix of an analytically given parabola

Equality of the two expressions is confirmed by simple calculation:

$$\sqrt{x^2 + (ax^2 - \frac{1}{4a})^2} = ax^2 + \frac{1}{4a}$$

$$x^2 + (ax^2 - \frac{1}{4a})^2 = \left(ax^2 + \frac{1}{4a}\right)^2$$

$$x^2 + a^2x^4 - \frac{x^2}{2} + \frac{1}{16a^2} = a^2x^4 + \frac{x^2}{2} + \frac{1}{16a^2}.$$

□

Corollary 5. *Folding of the paper from 'hand on activity 1' represents the parabola given by Definition 2 with the edge of paper as directrix and point F as focus.*

Proof. Folds from 'hand on activity 1' are exactly perpendicular bisectors of AF from corollary 3. The 'beauty of parabola image' given by folding the paper becomes even more suggestive when we realize, that this folds represent 'tangents' of our parabola (intuitive concept of an envelope), as explained in the following corollary. □

Corollary 6. *In the construction of a parabola given by corollary 3, the perpendicular bisector of AF is tangent to the parabola $\mathcal{P}(p, F)$.*

Proof. We have already seen that point T on the perpendicular bisector r of AF is on the parabola $\mathcal{P}(p, F)$. Therefore, r is tangent to $\mathcal{P}(p, F)$ or it must intersect the parabola in another point U. As seen on figure 5, for any such point U we would have $UF = UA \neq UV$, which proves that for $U \neq T$ point U can not be equidistant from point F and directrix p. Thus, perpendicular bisector of AF is tangent to parabola $\mathcal{P}(p, F)$.

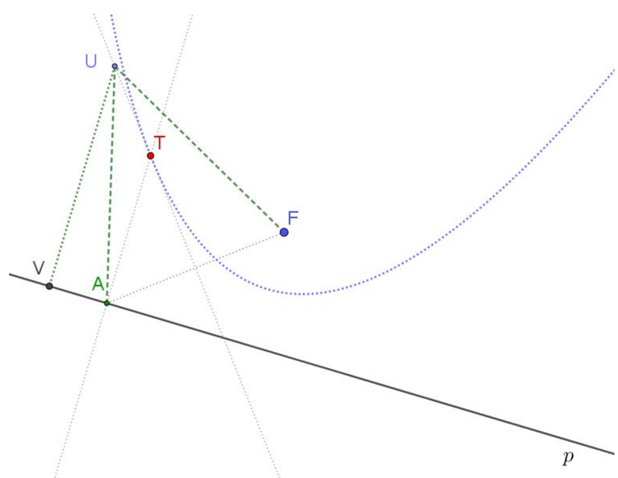

Figure 5: Perpendicular bisector is tangent

□

Remark: Analytical approach to parabola, which is given by the function of the form $f(x) = ax^2 + bx + c$ offers many advantages in analyzing 'parabola', but at the same time important geometric properties of a parabola, which were known already to the ancient Greeks, are much more accessible and intuitively easy to explain, if we use geometric approach. In the sequel we present and analyze some of these properties, which are undoubtedly much more elegantly explainable within the geometric approach.

Definition 7. *The line which is perpendicular to the directrix* p *and passes through the focus* F *is called the axe of the parabola* $\mathcal{P}(p, F)$.

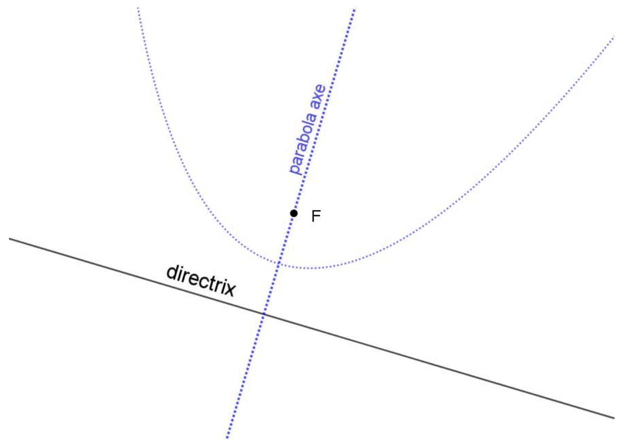

Figure 6: Directrix and axis of a parabola

Corollary 8. *Any ray parallel to the axis of a parabola bounces of a parabola in the direction of the focus as indicated on Figure 7.*

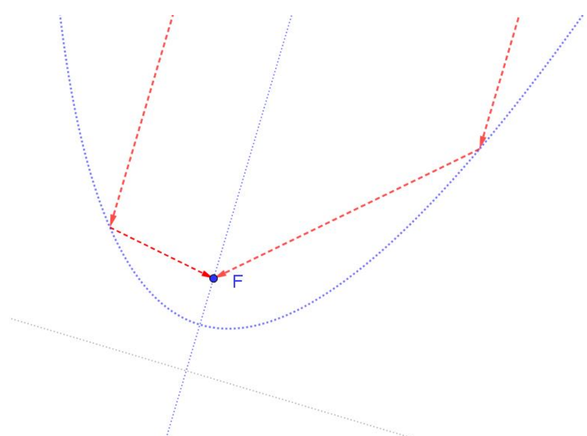

Figure 7: Rays parallel to axis bounce of the parabola towards the focus

Proof. Considering the fact that perpendicular bisector of AF from Corollary 3 (denoted by b on figure 8) is tangent to the parabola (Corollary 6), we only need to determine the equality of angles $\gamma = \beta$ on Figure 8. But $\alpha = \beta$ since triangle ΔAFT is isosceles and $\alpha = \gamma$ since α and γ are opposite angles. Thus $\gamma = \beta$.

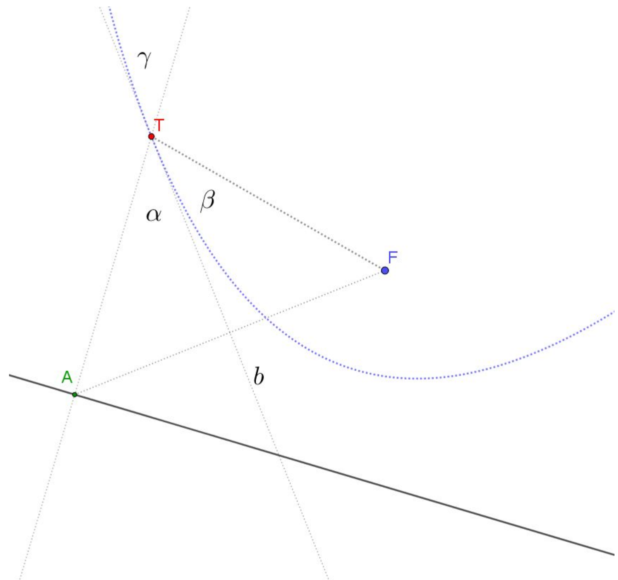

Figure 8: $\alpha = \beta = \gamma$

□

Corollary 9. *This geometric property of a parabola has been used in many technological devices like satellite antennas, radars and car lights.*

- *In satellite antennas and radars, which are of a parabolic shape, the sensor is put in the focus, where all the bounced off signals are collected (Figure 9).*

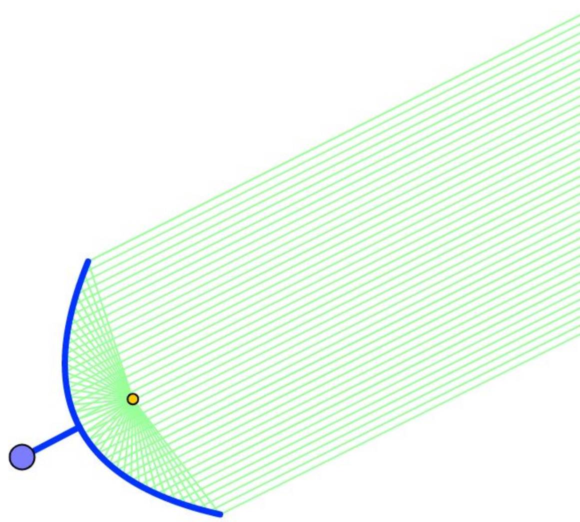

Figure 9: Signals are collected in the focus

- *For the long (flash) car light the source of light is put in the focus (Figure 10).*

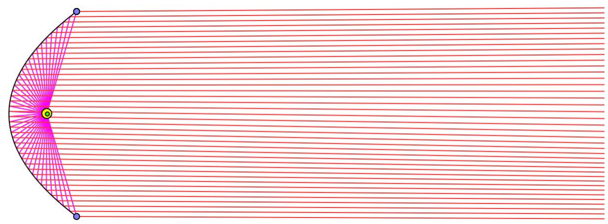

Figure 10: The source of long light is in the focus

- *For the short (dimmed) car light the source of light is put millimeters ahead of the focus and directed to the upper part of the parabola. (Figure 11).*

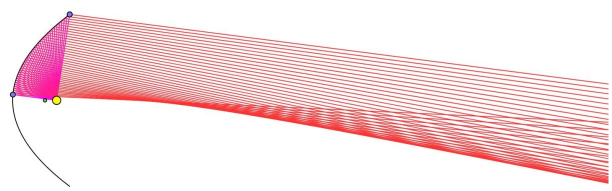

Figure 11: The source of short light is not in the focus

Positioning of the source of long and short car lights is clearly visible on the image of the real car bulb and on the scheme on Figure 12.

Figure 12: The source of short light is out of the focus

3 Ellipse by paper folding

Hands on activity 10. *Cut a circle (with center O) out of a sheet of paper and choose any point F somewhere close to the edge of a cut out circle (Figure 13).*

Figure 13: Circle shaped piece of paper with point F

Fold the paper several different ways ... so that in any of the folds the circle edge meets point F. Think about the analogy with 'Hands on activity 1'.

After folding the paper several times the folds form an obvious 'elliptic shape' (Figure 14 - see: `https://www.geogebra.org/m/phtegwnp#material/Z4bBvMX7`).

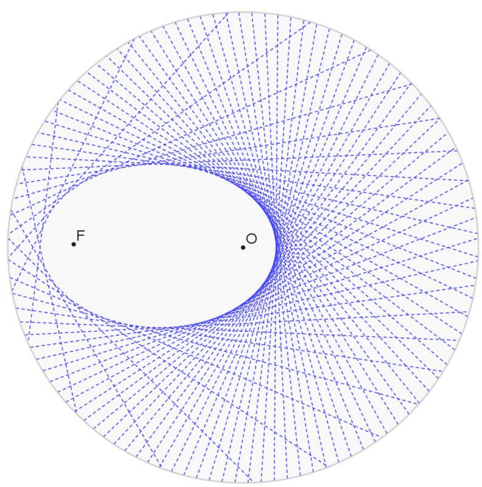

Figure 14: Circle folded many times so that the edge meets point F

Is the curve indicated by these folds really an ellipse?

Definition 11. [Gardener definition] *An ellipse defined by two points (called foci) F, G and distance d is the set of all points T such that*

$$FT + TG = d.$$

An ellipse can intuitively be described as the outer-most points that can be drawn by pencil bound by the rope of length d attached at both ends as indicated on the figure 15.

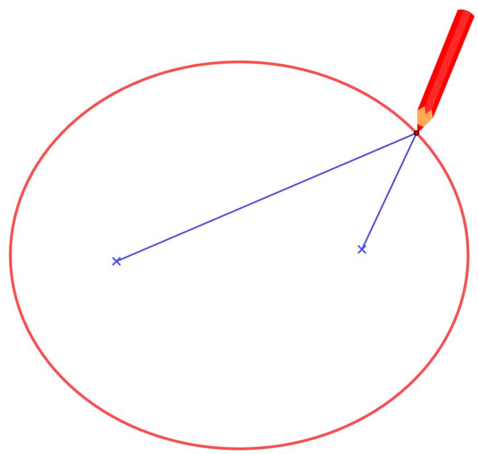

Figure 15: An ellipse

Theorem 12. *Definition of an ellipse given by Definition 11 coincides, or better generalizes the definition of an ellipse within the classical coordinate system given by the second degree equation* $\frac{(x-p)^2}{a^2} + \frac{(y-q)^2}{b^2} = 1$. *It is easy to prove, that (analytically given) ellipse* $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ *is the same as the ellipse defined geometrically by foci* $F(-\sqrt{a^2-b^2}, 0)$ *and* $G(\sqrt{a^2-b^2}, 0)$ *and distance* $d = 2a$ *in the case* $a > b$ *and by foci* $F(0, -\sqrt{b^2-a^2})$ *and* $G(0, \sqrt{b^2-a^2})$ *and distance* $d = 2b$ *in the case* $a < b$.

Proof. Let us prove the statement for the case $a > b$. For any point $T(x, y)$, its distances to points $F(-\sqrt{a^2-b^2}, 0)$ and $G(\sqrt{a^2-b^2}, 0)$ are (as in figure 16) given respectively by $\sqrt{(x+\sqrt{a^2-b^2})^2+y^2}$ and $\sqrt{(x-\sqrt{a^2-b^2})^2+y^2}$.

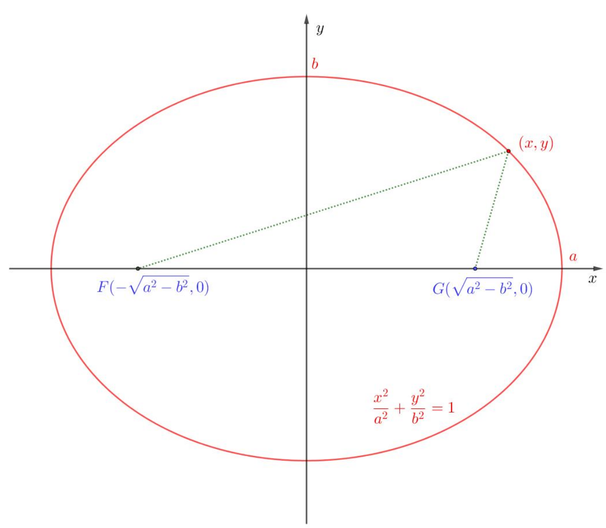

Figure 16: Foci of an analytically given ellipse

Simplifying the equation

$$\sqrt{(x+\sqrt{a^2-b^2})^2+y^2} + \sqrt{(x-\sqrt{a^2-b^2})^2+y^2} = 2a$$

yields exactly the equation of our analytically defined ellipse

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

which finishes the proof. ☐

Corollary 13. *Folding of the 'circle shaped paper' from 'hand on activity 10' represents an ellipse given by Definition 11 with the two foci being the chosen point F and the center 0 of the 'paper circle' and with distance d being the radius of 'paper circle'.*

Proof. Folds from 'hand on activity 10' are exactly perpendicular bisectors of AF, where A is any point on the edge of the circle, as seen on figure 17. Notice the analogy with 'hand on activity 1' and the construction of a parabola in Corollary 3.

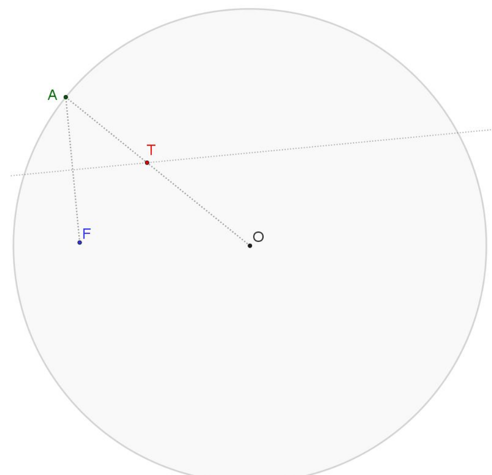

Figure 17: Folding circle's edge to meet point F

Triangle ΔATF is isosceles and $AT = TF$. Since OA is the radius of our circle and $OA = OT + TA = OT + TF$, by definition 11 point T travels along the ellipse defined by distance d being the radius of our circle and foci F and O. Furthermore, we see that point T is equidistant from the circle and point F. Analogous arguments as in Corollary 6 prove, that 'our folds' (that is perpendicular bisectors of AF) represent 'tangents' of our ellipse (intuitive concept of an envelope). ☐

Corollary 14. **[Alternative definition of an ellipse]** *An ellipse can be defined as the set of points that are equidistant from a circle and a point (within the circle).*

Proof. Follows from discussion in the proof of Corollary 13. ☐

Corollary 15. *Any ray starting in one of the two foci bounces of the ellipse in the direction of the second focus as indicated on Figure 18*

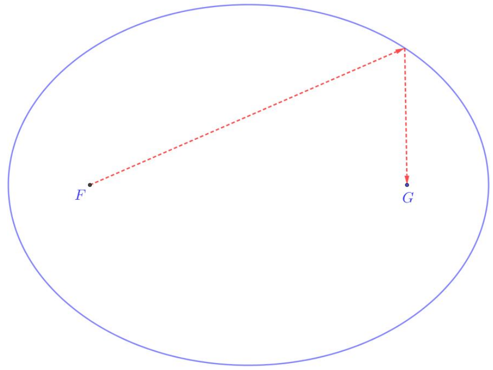

Figure 18: Ray starting in one focus bounces of to the second of the two foci

Proof. The statement and also the proof are quite analogous to those in Corollary 8. Considering the fact that perpendicular bisector of AF (denoted by b on figure 19) is tangent to the ellipse (discussion in the proof of Corollary 13), we only need to determine the equality $\gamma = \beta$ on Figure 19. But $\alpha = \beta$ since triangle ΔAFT is isosceles and $\alpha = \gamma$ since α and γ are opposite angles. Thus $\gamma = \beta$.

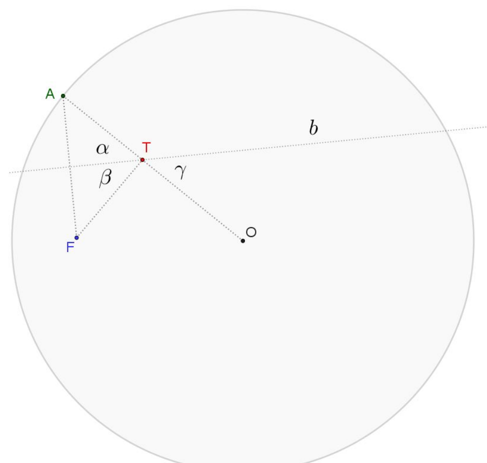

Figure 19: $\alpha = \beta = \gamma$

□

4 Hyperbola by transparent paper folding

Hands on activity 16. *Draw a circle (with center O) on a transparent paper and choose any point F outside the circle (Figure 20).*

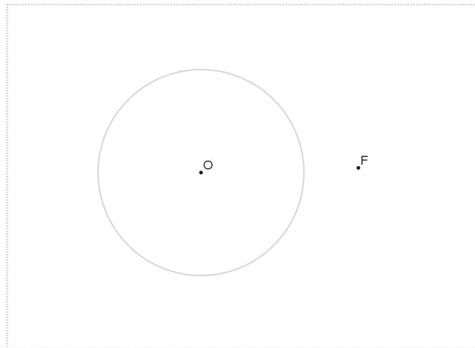

Figure 20: Circle and an outside point F on transparent paper

Fold the paper several different ways ... so that in any of the folds the circle edge meets point F. Think about the analogy with 'Hands on activity 1 and 10'.

After folding the paper several times the folds form an obvious 'hyperbolic shape' (Figure 21). Compare the dynamic visualization at https://www.geogebra.org/m/phtegwnp#material/Z4bBvMX7 and observe how an ellipse changes to hyperbola just by pulling the focus F outside the given circle.

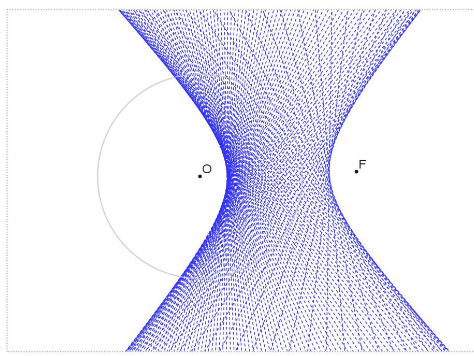

Figure 21: Paper folded many times so that the edge of the circle meets point F

Is the curve indicated by these folds really a hyperbola?

Definition 17. *A hyperbola defined by two points (called foci) F, G and distance d is the set of all points T such that*

$$|FT - TG| = d.$$

Theorem 18. *Definition of a hyperbola given by Definition 17 coincides, or better generalizes the definition of a hyperbola within the classical coordinate system given by the second degree equation $\frac{(x-p)^2}{a^2} - \frac{(y-q)^2}{b^2} = 1$. It is easy to prove, that (analytically given) hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ is the same as the hyperbola defined geometrically by foci $F(-\sqrt{a^2+b^2}, 0)$ and $G(\sqrt{a^2+b^2}, 0)$ and distance $d = 2a$.*

Proof. For any point $T(x, y)$, its distances to points $F(-\sqrt{a^2+b^2}, 0)$ and $G(\sqrt{a^2+b^2}, 0)$ are (as in figure 22) given respectively by $\sqrt{(x+\sqrt{a^2+b^2})^2+y^2}$ and $\sqrt{(x-\sqrt{a^2+b^2})^2+y^2}$.

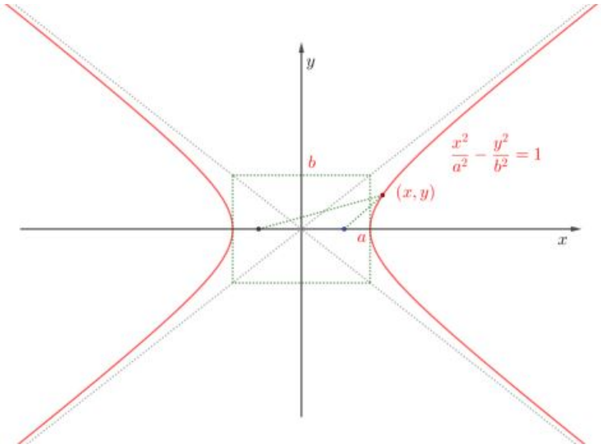

Figure 22: Foci of an analytically given hyperbola

Simplifying the equation

$$\left| \sqrt{(x+\sqrt{a^2+b^2})^2+y^2} - \sqrt{(x-\sqrt{a^2+b^2})^2+y^2} \right| = 2a$$

yields exactly the equation of our analytically defined hyperbola

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

which finishes the proof. □

Corollary 19. *Folding of the paper from 'hand on activity 16' represents the hyperbola given by Definition 17 with the two foci being the chosen point F and the center 0 of the circle and with distance d being the radius of the circle.*

Proof. Folds from 'hand on activity 16' are perpendicular bisectors of AF, where A is any point on the edge of the circle, as seen on figure 23. Notice the analogy with 'hand on activity 1 and 10'.

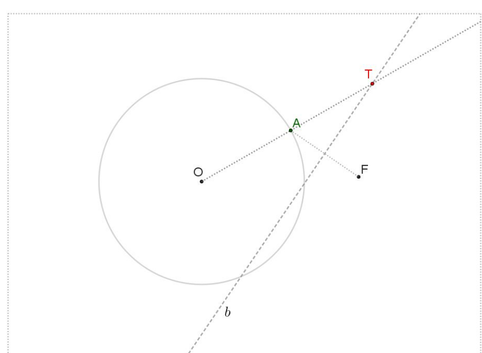

Figure 23: Folding cirle's edge to meet (outside) point F

Triangle ΔATF is isosceles and $AT = TF$. Since OA is the radius of our circle and $OA = OT - TA = OT - TF$, by definition 17 point T travels along the hyperbola defined by distance d (being the radius of our circle) and foci F and O. Furthermore, we see that point T is equidistant from the circle and point F. Analogous arguments as in Corollary 6 and in the proof of Corollary 13 prove, that 'our folds' (that is perpendicular bisectors of AF) represent 'tangents' of our hyperbola (intuitive concept of an envelope). □

Corollary 20. **[Alternative definition of a hyperbola]** *A hyperbola can be defined as the set of points that are equidistant from a circle and a point (outside the circle).*

Proof. Follows from discussion in the proof of Corollary 19. □

5 The unifying view on the Ellipse, Parabola and Hyperbola

There is yet another unifying view on the similarity between ellipse, parabola and hyperbola.

Let us recall that parabola is defined as the set of points T which are equidistant from given directrix and focus. Using notation $d(p,T)$ and $d(F,T)$ to denote distance between line p and point T and between points F and T respectively, we could also define parabola as the set of points for which

$$\frac{d(F,T)}{d(T,p)} = k = 1.$$

Theorem 21. **[The unifying definition of Ellipse, Parabola and Hyperbola]** *Given directrix p and focus point F, the set of all points T for which the ratio of distances*

$$\frac{d(F,T)}{d(T,p)} = k \quad \textit{is constant, is} \quad \begin{cases} \textit{an} \ \textbf{ellipse,} & \textit{if } k < 1, \\ \textit{a} \ \textbf{parabola,} & \textit{if } k = 1, \\ \textit{a} \ \textbf{hyperbola,} & \textit{if } k > 1. \end{cases}$$

Proof. First of all, if $k = 1$, then by Definition 2 our statement is correct. So we can assume that $k \neq 1$. The easiest way to proceed our proof is to use coordinate system. To further simplify our notation let our directrix be the line $x = 0$ and focus $F(1,0)$ as on Figure 24.

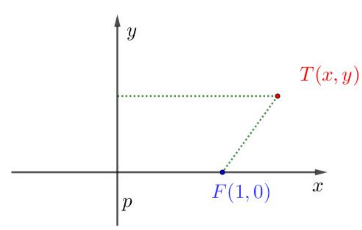

Figure 24: Comparing the distance of point T from point $(1,0)$ and line $x = 0$

Assuming $\frac{d(F,T)}{d(T,p)} = k$, we get simple equation

$$\frac{\sqrt{(x-1)^2 + y^2}}{x} = k$$

which easily simplifies into

$$\frac{(x - \frac{1}{1-k^2})^2}{(\frac{k}{1-k^2})^2} + \frac{y^2}{\frac{k^2}{1-k^2}} = 1.$$

From this equation it is obvious, that for $k < 1$ we get an ellipse (with center $S = (\frac{1}{1-k^2}, 0)$ and axes $a = \frac{k}{1-k^2}$ and $b = \frac{k}{\sqrt{1-k^2}}$), while for $k > 1$ we get a hyperbola (with center $S = (-\frac{1}{k^2-1}, 0)$ and axes $a = \frac{k}{k^2-1}$ and $b = \frac{k}{\sqrt{k^2-1}}$) as indicated on Figure 25

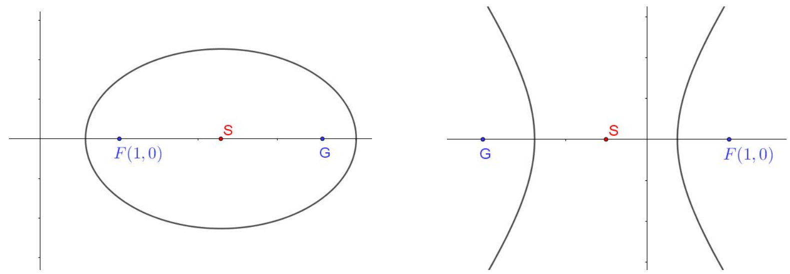

Figure 25: Ellipse for $k < 1$ and hyperbola for $k > 1$

From center S and focus F we could also easily calculate the other focus as $G(\frac{1+k^2}{1-k^2}, 0)$. □

There is yet another intuitively interesting and unifying perspective on the ellipse, parabola and hyperbola, which is essentially equivalent to the theorem 21. For that we need the theorem of Apollonius.

Theorem 22. **[Apollonius theorem]** *For given points A and B, the set of all points T for which the ratio of distances*

$$\frac{d(A,T)}{d(T,B)} = k \quad \textit{is constant, is} \quad \begin{cases} a \quad \textbf{line} \textit{ (perpendicular bisector)}, & \textit{if } k = 1, \\ a \quad \textbf{circle}, & \textit{if } k \neq 1. \end{cases}$$

Proof. The proof is elementary, but we skip it here and refer the reader to the dynamic visualization of the proof in the chapter 'Parabola, Ellipse, Hyperbola Analogy' at https://www.geogebra.org/mphtegwnp#material/G5hyWkFB of the GeoGebra book 'The Beauty of Conics' at https://www.geogebra.org/m/phtegwnp'. □

Definition 23. *For the purpose of simplicity let for any given pair of points A and B and $k \in \mathbb{R}^+$ call* **Apollonius circle** *the associate Apollonius circle in the case of $k \neq 1$ or the perpendicular bisector of AB in the case of $k = 1$.*

Considering Theorem 22 the very elementary construction of a parabola from Corollary 3 can be nicely generalized into the construction of an ellipse, a parabola or a hyperbola.

Corollary 24. **[Construction of ellipse, parabola or hyperbola]**

- *We start with a given directrix p and point $F \notin p$.*
- *For any point $A \in p$, we draw a line r through A perpendicular to p.*
- *Assuming $k \in \mathbb{R}^+$ we draw Apollonius circle a associated to points F, A and number k.*
- *As A travels along directrix p, the locus of the intersection points between line r and Apollonius circle a, describe exactly points of an ellipse, a parabola or a hyperbola, depending on the value of k.*

For the visualization of the corollary statements see the last applet at https://www.geogebra.org/mphtegwnp#material/G5hyWkFB.

6 Dandelin Spheres and Conic Sections

It is a well known fact and one that can be easily proved, that parabola, ellipse and hyperbola as defined and described in definitions and theorem 2, 11, 17 and 21 can all be visualized as intersections of an infinite cone and an appropriate plane. The detailed explanation can not be done within this short presentation, but let us direct the reader to study the dynamic visualization at https://www.geogebra.org/m/phtegwnp#material/gSUZ53Km with just a couple of hints:

- Foci are the points where the so called Dandelin spheres touch the plane 'defining' (for example) an ellipse.
- For any sphere and any two tangents from a given point, the two 'tangent segments' are of the same length.
- Directrix (from) theorem 21 is the intersection line between the plane 'defining' (for example) an ellipse and the plane determined by circle where Dandelin sphere touches the cone.

7 Archimedes Parabola's Section

We would like to finish this short presentation by another beautiful view on a parabola, which is over two thousand years old and belongs to Archimedes. Just think about the simplicity and power of the simple formula for the area of a triangle, which is given by the base and the height and is fully independent on the 'shape' or position of the 'top vertex'. As presented on the figure 26, does not matter where we choose the vertex C_1 and C_2 as far as we keep the base and the height unchanged, the area of a triangle would remain $\frac{1}{2} \cdot b \cdot h$.

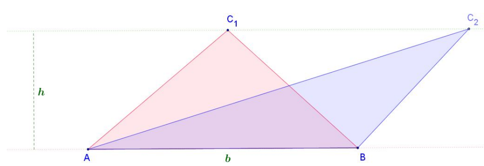

Figure 26: The area of a triangle is determined by base and height

It is very interesting that an analogous formula holds for parabola section. First it is not even trivial to see, that as with a triangle, does not matter where on the parallel line from the base line we put the point C_1 or C_2 as seen on figure 28, we can always have a parabola with the same base and given height - reaching that height at the chosen point. And the area of the parabola section is constant and given by the formula $\frac{2}{3} \cdot b \cdot h$.

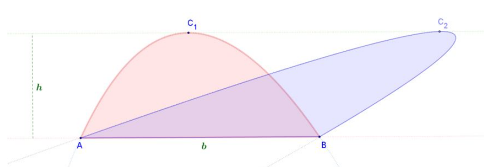

Figure 27: The area of a parabola section is determined by base and height

The proof of this ingenious Archimedes idea is a wonderful intuitive and elementary geometric reasoning, which would nevertheless require some more time and space than we have. We address the reader to the dynamic visualizations given at https://www.geogebra.org/m/phtegwnp#material/Uf8rr2fr, with a couple of hints:

- The starting point of the Archimedes proof is the geometric definition of a parabola (as in definition 2).
- *Archimedes triangle* is a triangle defined by any two points on a parabola and the third triangle vertex being the intersection of the two tangents to the parabola at the initial two points on a parabola.
- In any Archimedean triangle the line connecting the midpoint of the two vertex on a parabola and the third vertex of the triangle is parallel to the axes of the parabola.
- Archimedean triangle is split into four triangles: one inside parabola and one outside. The inside triangle has twice the area of the outside triangle. The remaining two triangles are also Archimedean triangles and the recursion step can be repeated.

Notably the same ratio of $\frac{2}{3}$ as it appears in this wonderful formula for the area of a parabola section is also the ratio between the volume of a cylinder circumscribed to a given sphere. And it is the formula that also belongs to Archimedes and has the most prominent role, as it is carved on Fields medal together with Archimedes portrait and name.

Figure 28: Archimedes portret, name and to a given sphere circumscribed cylinder carving on Fields Medal

References

[1] Kobal, D.; The Beauty of Conics, GeoGebra book at https://www.geogebra.org/m/phtegwnp, Accessed on December 22, 2019.