GEOMETRIC CENTER OF MASS FOR POINTS ON CONIC SECTIONS: PROPERTIES, GENERALIZATIONS, APPLICATIONS, AND MYSTERIES

Boris Hanin (Century HS and Stanford University)

1. The Who, When, What, Where, and Why

In the summer of 2002, my family and I traveled to St. Petersburg, Russia to visit my grandparents. While there, my father told me about an intriguing theorem in geometry that he himself learned in 1972 from Dr. Zalman A. Skopets (1917-1984) who was married to my grandmother's best friend. Dr. Skopets was a prominent geometer and headed the Department of Geometry at the Yaroslavl' State Pedagogical University for many years. In the summers, my father would visit Dr. Skopets who would give him challenging problems in geometry. One such problem appealed to him so much that he remembered it even on that summer day in St. Petersburg some 30 years later. Here is that remarkable problem:

Let $A_1, A_2,\ldots, A_n$ be n arbitrary points on a circle and M be their geometric center of mass. Also, let $B_1, B_2,\ldots, B_n$ be the respective second points of intersection of the lines $A_1M, A_2M,\ldots, A_nM$ with the circle. Prove that

$$\frac{A_1M}{MB_1}+\frac{A_2M}{MB_2}+\ldots+\frac{A_nM}{MB_n}=n. \qquad (1)$$

Just as a clarification to the reader, AB stands for the length of the line segment connecting points A and B, and the geometric center of mass of several points is the point M such that equal masses positioned at these points and placed on a light sheet would balance if a fulcrum were placed at M. What is most amazing about equation (1) is that the positions of the points A_1, $A_2, \ldots, A_n$ are absolutely arbitrary, and some of these points may even coincide!

My father remembers that in discussing the problem, Dr. Skopets dropped an enigmatic phrase to the effect that this theorem goes back to "an old German text in geometry." Who created this mathematical gem? Knowing Dr. Skopets' mathematical taste and affections as well as the nature of the problem, my father suggested that the most likely candidate would be the great German geometer Jacob Steiner (1796–1863) who discovered many fundamental properties of the center of mass and moments of inertia. This opinion was later seconded by a differential geometer, Dr. Robert Fisher, Jr., who was my mentor at Idaho State University and from whom I have taken several math courses. Browsing through the two-volume Steiner's *Gesammelte Werke* [1] (German for collected works), however, did not reveal anything even remotely resembling the result in question. Thus, to this day, I have not been able to find the author and origin of the theorem.

For n=2 points on a circle, equality (1) is obvious because the geometric center of mass of two points lies halfway between them. So what first came to my mind is to check the theorem for other particular choices of points on a circle. The simplest among them is the case of three points on a circle (n=3). To my surprise, the calculations that arose were completely unmanageable, even for three points forming an isosceles triangle. However, with a little encouragement from my father to combine plane geometry with vector algebra, I eventually ended up proving equality (1). This proof was the main driving force behind the entire project; because of this, it is given in the last section.

Over the next two years, I worked on generalizing and extending this theorem to the other conic sections. The beauty of working on math is that all I needed was pencil and paper. I didn't have to keep lab hours or wait for something to grow or synthesize. Because this theorem deals exclusively with classical geometry, the only math I used was plane and three-dimensional geometry, vector algebra and the method of coordinates that I learned from my middle and high school geometry studies. So what made this research possible is really more curiosity, ingenuity, and inspiration rather than a vast knowledge of mathematics. However, as a result of this project, I learned many new things about the center of mass, discovered for myself the method of transformations, sharpened my skills in vector algebra and analytic geometry, and perhaps most importantly, honed my ability to construct mathematical proofs. For me, the most difficult part of the project was formulating exactly in what direction I wanted to expand the theorem.

2. The Main Work

One of the most natural questions that occurred to me when looking at formula (1) was as follows: Given points A_1, A_2,…, A_n, are there points inside the circle other than M for which equality (1) remains true? The answer is YES: if the geometric center of mass is replaced by the center of the circle, O, for example, then equation (1) obviously holds because every ratio would equal one. My intuition told me that it would be strange to have only two discrete points, the center of the circle and the center of mass, that satisfy this equation. In fact, when I looked closely at my original proof of the theorem (given in Section 4) I was able to find the set of *all* points inside the circle with property (1). That set turned out to be a circle with diameter OM, which of course includes O and M!

If there are so many (infinitely many, unless M=O) points inside the circle for which the theorem holds, could there possibly be any points outside the circle for which it holds as well? Again, the answer is YES. An intuitively appealing reason for this is that if the distance between point M and the circle is very large compared it its radius, then each ratio in (1) would be very close to 1 so that the sum of these ratios would be very close to n. Note that points B_1, B_2,…, B_n can be constructed even if point M lies outside of the circle, and hence equality (1) makes perfect sense. A key to proving the existence of points M outside the circle with property (1) is a very interesting transformation of the plane called *inversion with respect to a circle* that fixes the circle and interchanges its interior and exterior [2, p.452]. Using the inversion and applying the same methods as in the proof of the theorem for the points on the circle with diameter OM, I was able to find that there is a whole line of points outside the circle for which equality (1) holds. Furthermore, this line is the image of the circle with diameter OM under inversion with respect to our original circle.

As I went over my work for the circle, I noticed a lot of the circles I drew looked more like ellipses. Although this was because my artistic abilities are rather poor, it did give me an idea: maybe this theorem is true for points on an ellipse as well. After all, ellipses are just stretched or compressed circles. I quickly realized that such squashing and stretching is a linear transformation of the plane that does not affect the ratios of lengths of collinear segments and maps the geometric center of mass of the original points into the geometric center of mass of their images. So, all the properties of the equation (1) that I proved for the circle must indeed be true for the ellipse.

This discovery, together with all the work I did for the circle, really made me think of the theorem in a new way. I felt more and more that the original theorem was just a special case of something much bigger. Of course, this could easily have proved to be wishful thinking, but for this particular problem it turned out to be right on target. In fact, the next big step in my project was realizing that all of the properties of the center of mass that I proved for circles and ellipses must be true for parabolas as well. Again, I did some sample calculations to check equality (1) in several specific cases, and they matched the theorem's predictions exactly.

The main idea is to consider parabolas as ellipses with one focus at infinity. This approach comes from the fact that circles, ellipses, and parabolas (and hyperbolas as well) are all conic sections. That is, they can be viewed as intersections of the surface of a cone with a plane that passes through, and can be rotated around, a given point on the axis of the cone different from its vertex (see Figure 1).

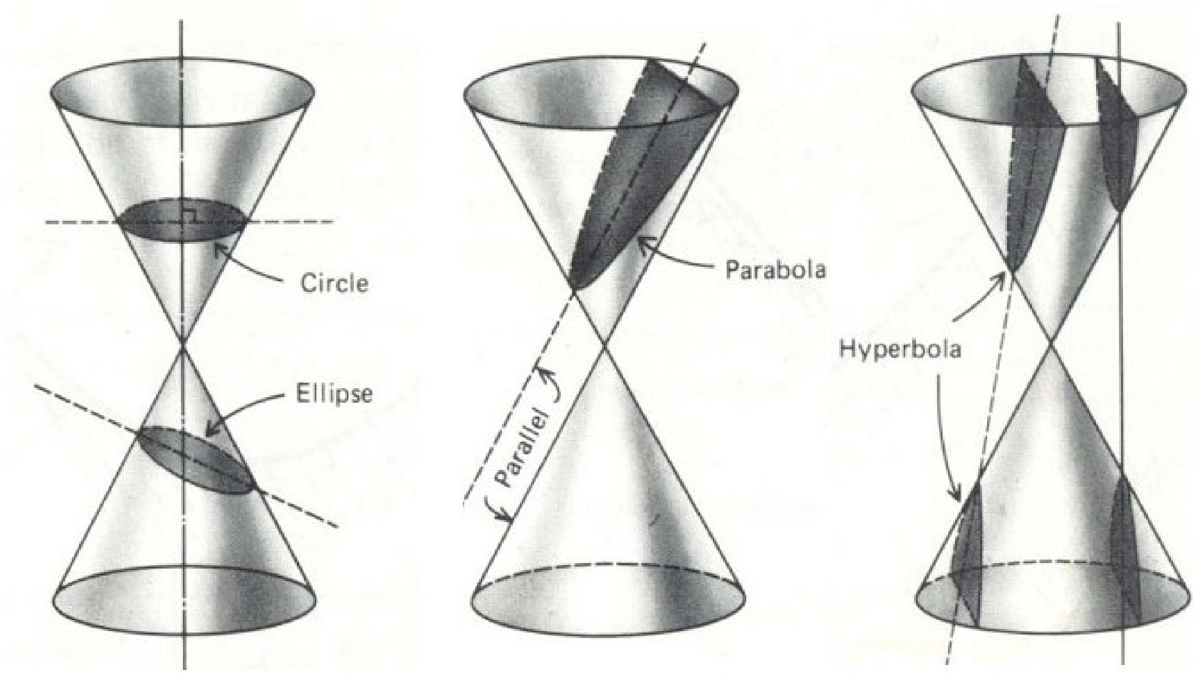

Figure 1

When the plane is perpendicular to the axis of the cone then the intersection of the cone and the plane is a circle. If we start to rotate the plane then the resulting conic section will be an ellipse until the plane will become parallel to the line that generates the cone by rotation around its axis. For this position of the plane, the conic section is a parabola. As we continue to rotate the plane, the produced conic sections will be hyperbolas. An important conclusion from this mental experiment is that a parabola can be viewed as a limiting case of ellipses. Now, to show property (1) for a parabola, imagine n given points on a parabola. Let us rotate the plane back by a small angle to get an ellipse. The rotated n points on the parabola can be projected onto the ellipse along the rotated plane from the center of rotation. For these new points, as we already know, equation (1) holds. Notice that, as the angle by which we rotate our plane to go from a parabola to an ellipse gets closer to zero, the points on the ellipse and their center of mass get closer to the original points on the parabola and their center of mass. Then, taking the limit as the angle of rotation of the plane tends towards zero in property (1) for the points on an ellipse yields this property for our n points on the parabola. Despite the simplicity of this argument, it took me several weeks to hammer out its details and make it into a rigorous mathematical proof.

The proof of the property (1) for the parabola has two significant drawbacks. First, it only established this property for the geometric center of mass. My sample calculations and my intuition, however, suggested that there are other points, similar to those for the circle and the ellipse, for which (1) is true. Second, I couldn't use this method to prove the theorem for the last of the conic sections, hyperbolas, because they cannot be represented as limiting cases of either ellipses or parabolas. I did not resolve these difficulties until after I proved property (1) for the hyperbola.

Hyperbolas turned out to be especially tricky and at the same time illuminating as to the true nature of the theorem. It turned out that if our n points on a hyperbola belong to the same branch, then equality (1) is satisfied. However, if both branches of the hyperbola contain some of these points, then (1) may not true anymore. For example, let us consider the hyperbola $y=1/x$ with $n=3$ and $A_1=(1/3,3)$, $A_2=(3,1/3)$, and $A_3=(-1,-1)$. The coordinates of the geometric center of mass are arithmetic means of the respective coordinates of the points. Thus, $M=(7/9,7/9)$. Using some standard analytic geometry we find that the second points of intersection of the lines A_1M, A_2M and A_3M with the hyperbola are $B_1=(3/5,5/3)$, $B_2=(5/3,3/5)$ and $B_3=(1,1)$. Then

$$A_1M/MB_1+A_2M/MB_2+A_3M/MB_3= 2.5+2.5+8=13$$

so that (1) fails! Note, however, that if we change the signs of the first two ratios, making them negative, then the sum of the ratios would be equal to 3, as required by (1). Also, observe that for these ratios the point M lies outside the segments A_1B_1 and A_2B_2. This proved to be a general rule: if a point P belongs to the line AB but does not lie between points A and B, then the ratio AP/PB should be viewed as negative. With this convention, I showed using analytic geometry that (1) holds true for arbitrary points on a hyperbola. Moreover, using the same method I found *all* points M in the plane with property (1) for hyperbolas. Finally, I adapted this method to obtain a parallel result for parabolas as well as confirm all my results for circles and ellipses that I initially found using plane geometry, vector algebra, and transformations.

The idea to use signed ratios of segments occurred to my mentor, Dr. Fisher. It resulted from dynamic constructions on The Geometer's Sketchpad [3] that he used to confirm my results for circles, ellipses and parabolas as well as to find out why their direct extension fails if applied to hyperbolas. These constructions made formula (1) come alive: if you build a conic section with some points on it and drag one of the points along the curve, then the program automatically and immediately recomputes the left-hand side of (1).

Finally, all the results for circles readily extend to spheres of any number of dimensions because, in the proof of the original theorem for circles, I used vector algebra and some geometric arguments, which are true in any number of dimensions (see Section 4). Furthermore, linear transformations of spheres produce similar results for the surfaces of ellipsoids of any number of dimensions.

This is as far as my theoretical work went. From here, my father and my mentor, Dr. Fisher, were so excited by my work, that they themselves began thinking about the theorem and ended up proving its far-reaching extension for points (and even arbitrary mass distributions) on general quadratic surfaces in any number of dimensions.

3. Application to Moments of Inertia

At this point I felt my project was missing some real-world applications. I knew there had to be some because the center of mass of an object is an extremely important physical characteristic. My father and I found an interesting application of the main theorem of my project in the case of spheres to moments of inertia that I studied in my multivariate calculus class. The moment of inertia of an object is an important characteristic of the distribution of its mass in space that plays the same role in the rotational motion of an object (with its mass concentrated at its center of mass) as its mass does in the object's linear motion.

The application is a generalization of Steiner's (the same Steiner who we suspected came up with formula (1)) formula for moments of inertia in the case of spherical objects. In its original form, Steiner's formula allows one to compute the moment of inertia of an object around any point in space given the object's mass and its moment of inertia around its center of mass. If the object in question is a sphere, with some arbitrary mass distribution on it, then it turned out that, in order to compute its moment of inertia around any point in space, it is sufficient to know the moment of inertia of the object around one of the points that satisfy property (1), not necessarily the center of mass. In reality, determining the moment of inertia of an object around its center of mass may be difficult because the center of mass is often inaccessible. The result stated above solves the inaccessibility problem because, as we know, there are infinitely many points outside the sphere (actually, a whole plane of them) for which property (1) holds. This could be very useful in astronomy, for example, because planets and stars are roughly spherical and can be viewed as spherical shells.

4. Proof of the Original Theorem for a Circle

In this final section, I just can't resist giving the proof of property (1) for the circle. It is really elementary, and I think quite elegant. Throughout most of the proof, I use an arbitrary point P in the plane, instead of M. This will hopefully make the fact that M is not the only point for which property (1) holds more believable. Please also note that, to save space, I use sigma notation throughout the proof. By definition,

$$\sum_{i=1}^{n}\frac{\mathrm{A_iM}}{\mathrm{MB_i}}=\frac{\mathrm{A_1M}}{\mathrm{MB_1}}+\frac{\mathrm{A_2M}}{\mathrm{MB_2}}+\ldots+\frac{\mathrm{A_nM}}{\mathrm{MB_n}}.$$

Theorem. *Let A_1, A_2,…, A_n be n points on a circle, M be their geometric center of mass, and B_i be the points of intersection of lines A_iM with the circle, respectively, i=1,2,…, n. Then*

$$\sum_{i=1}^{n}\frac{\mathrm{A_iM}}{\mathrm{MB_i}}=\mathrm{n}\ .$$

Proof. Let O be the center of the circle and r be its radius. Pick any point P inside the circle. We will investigate

$$\sum_{i=1}^{n}\frac{\mathrm{A_iP}}{\mathrm{PB_i}}. \tag{2}$$

Considering any chord AB and the diameter that both pass through the point P we find, by the intersecting chords theorem for the circle [4, p.116-118], that

$$\mathrm{AP\cdot PB=(r+OP)(r-OP)=r^2-OP^2}.$$

Then the sum (2) can be rewritten as follows:

$$\sum_{i=1}^{n}\frac{\mathrm{A_iP}}{\mathrm{PB_i}}=\sum_{i=1}^{n}\frac{\mathrm{A_iP^2}}{\mathrm{A_iP\cdot PB_i}}=\frac{1}{r^2-OP^2}\ \sum_{i=1}^{n}A_i\mathrm{P}^2\ . \tag{3}$$

Further, using vector algebra we have

$A_iP^2=\mathbf{A_iP}^2=(\mathbf{OP}-\mathbf{OA_i})^2=\mathbf{OP}^2-2\mathbf{OA_i}\bullet\mathbf{OP}+\mathbf{OA_i}^2=OP^2+r^2-2\mathbf{OA_i}\bullet\mathbf{OP}$, i=1,2,…,n.

Therefore,

$$\sum_{i=1}^{n} \mathrm{A_iP}^2 = n(r^2+OP^2) - 2\sum_{i=1}^{n}\mathbf{OP}\bullet\mathbf{OA}_i = n(r^2+OP^2) - 2\mathbf{OP}\bullet\sum_{i=1}^{n}\mathbf{OA}_i\,. \qquad (4)$$

By the definition of the geometric center of mass,

$$\sum_{i=1}^{n}\mathbf{OA_i} = \mathrm{n}\mathbf{OM}\,.$$

Thus, (4) becomes

$$\sum_{i=1}^{n} \mathrm{A_iP}^2 = n(r^2+OP^2) - 2n\mathbf{OP}\bullet\mathbf{OM}\,.$$

Setting here P=M we have

$$\sum_{i=1}^{n} \mathrm{A_iM}^2 = \mathrm{n}(r^2+OM^2) - 2\mathrm{n}OM^2 = \mathrm{n}(r^2 - O\mathrm{M}^2)\,.$$

Finally, from (3) with P = M we obtain the required formula.

The theorem is proved.

References

[1] J. Steiner, *Gesammelte Werke*, vol. 1-2, Chelsea Publishing Company, 2nd ed., 1971.
[2] R. Bix, *Topics in Geometry*, Academic Press, Boston, 1994.
[3] J. King, *Geometry Through the Circle with The Geometer's Sketchpad*, Key Curriculum Press, Berkeley, CA, 1996.
[4] M. Berger, *Geometry*, vol. 1-2, Springer-Verlag, New York, N.Y., 1987.