A geometric proof of the Siebeck-Marden theorem

Beniamin Bogosel

June 15, 2015

Abstract

Siebeck-Marden theorem relates the roots of a third degree polynomial and the roots of its derivative in a geometrical way. A few geometric arguments imply that every inellipse for a triangle is uniquely related to a certain logarithmic potential via its focal points. This fact provides a new direct proof of a general form of the result of Siebeck and Marden.

Given three non-collinear points $a, b, c \in \mathbb{C}$, we can consider the cubic polynomial $P(z) = (z-a)(z-b)(z-c)$, whose derivative $P'(z)$ has two roots f_1, f_2. Gauss-Lucas theorem is a well known result which states that given a polynomial Q with roots $z_1, ..., z_n$, the roots of its derivative Q' are in the convex hull of $z_1, ..., z_n$. In the simple case where we have only three roots, there is a more precise result. The roots f_1, f_2 of the derivative polynomial are situated in the interior of the triangle Δabc and they have an interesting geometric property: f_1 and f_2 are the focal points of the unique ellipse which is tangent to the sides of the triangle Δabc at its midpoints. This ellipse is called the Steiner inellipse associated to the triangle Δabc. In the rest of this note, we use the term inellipse to denote an ellipse situated in a triangle, which is tangent to all three of its sides. This geometric connection between the roots of P and the roots of P' was first observed by Siebeck (1864) [12] and was reproved by Marden (1945) [8]. There has been a substantial interest in this result in the past decade: see [3],[5, pp 137-140] [7],[9],[10],[11]. Kalman [7] called this result Marden's theorem, but in order to give credit to Siebeck, who gave the initial proof, we call this result Siebeck-Marden Theorem in the rest of this note. Apart from its purely mathematical interest, Siebeck-Marden Theorem has a few applications in engineering. In [2] this result is used to locate the stagnation points of a sistem of three vortices and in [6] this result is used to find the location of a noxious facility location in the three-city case.

The proofs of the Siebeck-Marden theorem found in the references presented above are either algebraic or geometric in nature. The initial motivation for writing this note was to find a more direct proof, based on geometric arguments. The answer was found by answering the following natural question: *Can we find two different inellipses with the same center?* Indeed, let's note that $(a+b+c)/3 = (f_1+f_2)/2$, which means that the centers of ellipses having focal points f_1, f_2 coincide with the centroid of the triangle Δabc. The geometric aspects of the problem can be summarized in the following questions:

1. Is an inellipse uniquely determined by its center?

2. Which points in the interior of the triangle Δabc can be centers of an inellipse?

3. What are the necessary and sufficient conditions required such that two points f_1 and f_2 are the focal points of an inellipse?

4. Is there an explicit connection between the center of the inellipse and its tangency points?

We give precise answers to all these questions in the next section, dedicated to the geometric properties of inellipses. Once these properties are established, we are able to prove a more general version of the Siebeck-Marden Theorem. The proof of the original Siebeck-Marden result will follow immediately from the two main geometric properties of the critical points f_1, f_2:

- The midpoint of $f_1 f_2$ is the centroid of Δabc.
- The points f_1, f_2 are isogonal conjugates relative to triangle Δabc.

We recall that two points f_1, f_2 are isogonal conjugates relative to triangle Δabc if the pairs of lines $(af_1, af_2), (bf_1, bf_2), (cf_1, cf_2)$ are symmetric with respect to the bisectors of the angles a, b, c, respectively.

1 Geometric properties of inellipses

We start by answering the third question raised above: *Which pairs of points can be the foci of an inellipse?* In order to have an idea of what is the expected answer, we can look at the following general configuration: suppose we have an ellipse $\mathcal{E}$ with foci f_1, f_2 and an exterior point a. Consider the two tangents at_1, at_2 to $\mathcal{E}$ which go through a. Then the angles $\angle t_1 a f_1$ and $\angle t_2 a f_2$ are equal.

A simple proof of this fact goes as follows. Construct g_1, g_2 the reflections of f_1, f_2 with respect to lines at_1, at_2, respectively (see Figure 1 left). Then the triplets of points $(f_1, t_2, g_2), (f_2, t_1, g_1)$ are collinear. To see this, recall the result, often called, Heron's Problem, which says that the minimal path from a point a to a point b which touches a line ℓ not separating a and b must satisfy the reflection angle condition. Now, it is enough to note that $f_1 g_2 = f_1 t_2 + f_2 t_2 = f_1 t_1 + f_2 t_1 = g_1 f_2$. Thus, triangles $\Delta a f_1 g_2, \Delta a g_1 f_2$ are congruent, which implies that the angles $\angle t_1 a f_1$ and $\angle t_2 a f_2$ are equal.

As a direct consequence, the foci of an inellipse for Δabc are isogonal conjugates relative to Δabc. The converse is also true and this results dates back to the work of Steiner [13] (see [1]).

Theorem 1. *(Steiner). Suppose that Δabc is a triangle.*

1. *If $\mathcal{E}$ is an inellipse for Δabc with foci f_1 and f_2, then f_1 and f_2 are isogonal conjugates relative to Δabc.*

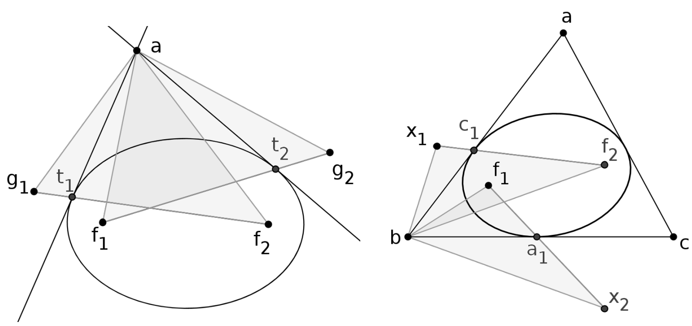

Figure 1: Left: basic property of the tangents to an ellipse. Right: construction of an inellipse starting from two isogonal conjugate points.

2. *If f_1 and f_2 are isogonal conjugates relative to Δabc, then there is a unique inellipse for Δabc with foci f_1 and f_2.*

Proof. The proof of 1. was discussed above, so it only remains to prove 2. Consider the points x_1, x_2 the reflections of f_1, f_2 with respect to the lines ab and bc (see Figure 1 right). The construction implies that $bf_1 = bx_1,\ bf_2 = bx_2$ and $\angle x_1 b f_2 = \angle f_1 b x_2$, which, in turn, implies that $x_1 f_2 = f_1 x_2$. We denote their common value with m. We denote $a_1 = f_1 x_2 \cap bc$ and $c_1 = x_1 f_2 \cap ab$. The construction of x_1, x_2 implies that $f_1 a_1 + f_2 a_1 = f_1 x_2 = f_2 x_1 = f_1 c_1 + f_2 c_1 = m$. Heron's Problem cited above implies that a_1 is the point which minimizes $x \mapsto f_1 x + f_2 x$ with $x \in bc$ and c_1 is the point which minimizes $x \mapsto f_1 x + f_2 x$ with $x \in ab$.

Thus, the ellipse characterized by $f_1 x + f_2 x = m$ is tangent to bc and ab in a_1 and, respectively c_1. A similar argument proves that this ellipse is, in fact, also tangent to ac. The unicity of this ellipse comes from the fact that m is defined as the minimum of $f_1 x + f_2 x$ where x is on one of the sides of Δabc, and this minimum is unique, and independent of the chosen side. $\square$

We are left to answer Questions 1,2 and 4. The first two questions were answered by Chakerian in [4] using an argument based on orthogonal projection. We provide a slightly different argument, which, in addition, gives us information about the relation between the barycentric coordinates of the center of the inellipse and its tangency points. In the proof of the following results we use the properties of real affine transformations of the plane.

Theorem 2. 1. *An inellipse for Δabc is uniquely determined by its center.*

2. *The locus of the set of centers of inellipses for Δabc is the interior of the medial*[1] *triangle for Δabc.*

[1]The medial triangle is the triangle formed by the midpoints of a triangle

3. *If the center of the inellipse $\mathcal{E}$ is $\alpha a + \beta b + \gamma c$, where $\alpha, \beta, \gamma > 0$ and $\alpha + \beta + \gamma = 1$, then the points of tangency of the inellipse divide the sides of Δabc in the ratios $(1-2\beta)/(1-2\gamma), (1-2\gamma)/(1-2\alpha), (1-2\alpha)/(1-2\beta)$.*

Proof. 1. We begin with the particular case where the ellipse $\mathcal{E}$ is the incircle, with center o, the incenter. Suppose $\mathcal{E}'$ is another inscribed ellipse, with center o, and denote by f_1, f_2 its focal points. We know that f_1, f_2 are isogonal conjugates relative to Δabc and the midpoint of $f_1 f_2$ is o, the center of the inellipse. Thus, if $f_1 \neq f_2$ then ao is at the same time a median and a bisector in triangle $af_1 f_2$. This implies that $ao \perp f_1 f_2$. A similar argument proves that $bo \perp f_1 f_2$ and $co \perp f_1 f_2$. Thus a, b, c all lie on a line perpendicular to $f_1 f_2$ in o, which contradicts the fact that Δabc is not degenerate. The assumption $f_1 \neq f_2$ leads to a contradiction, and therefore we must have $f_1 = f_2$, which means that $\mathcal{E}'$ is a circle and $\mathcal{E}' = \mathcal{E}$.

Consider now the general case. Suppose that the inellipses $\mathcal{E}, \mathcal{E}'$ for Δabc have the same center. Consider an affine mapping h which maps $\mathcal{E}$ to a circle. Since h maps ellipses to ellipses and preserves midpoints, the image of our configuration by h is a triangle where $h(\mathcal{E})$ is the incircle and $h(\mathcal{E}')$ is an inscribed ellipse with the same center. This case was treated in the previous paragraph and we must have $h(\mathcal{E}) = h(\mathcal{E}')$. Thus $\mathcal{E} = \mathcal{E}'$.

2. To find the locus of the centers of inellipses for Δabc, it is enough to see which barycentric coordinates are admissible for the incircle of a general triangle. We recall that barycentric coordinates of a point p are proportional with the areas of the triangles $\Delta pbc, \Delta pca, \Delta pab$, and their sum is chosen to be 1. Thus, barycentric coordinates are preserved under affine transformations. The barycentric coordinates of the center of an inellipse with respect to a, b, c are the same as the barycentric coordinates of the incenter of the triangle $h(a), h(b), h(c)$. As before, h is the affine transformation which transforms the ellipse into a circle. Conversely, if the barycentric coordinates of the intcenter o with respect to $\Delta a'b'c'$ are x, y, z, then we consider the affine transformation which maps the triangle $\Delta a'b'c'$ onto the triangle Δabc. The circle is transformed into an inellipse, with center having barycentric coordinates x, y, z.

The barycentric coordinates of the incenter have the form

$$x = \frac{u}{u+v+w},\ y = \frac{v}{u+v+w},\ z = \frac{w}{u+v+w},$$

where u, v, w are the lengths of the sides of $\Delta a'b'c'$. Thus, we can see that $x + y + z = 1$ and $x < y + z, y < z + x, z < x + y$. One simple consequence of these relations is the fact that $x, y, z < 1/2$. Furthermore, since

$$x = \frac{\text{Area}(ob'c')}{\text{Area}(a'b'c')}, y = \frac{\text{Area}(oc'a')}{\text{Area}(a'b'c')}, z = \frac{\text{Area}(oa'b')}{\text{Area}(a'b'c')},$$

we can see that the previous relations for x, y, z are satisfied if and only if o is in the interior of the medial triangle for $\Delta a'b'c'$. Thus, the locus of the center of an inscribed ellipse is the interior of the medial triangle.

3. If the center of the inellipse $\mathcal{E}$ is $\alpha a + \beta b + \gamma c$ with $\alpha + \beta + \gamma = 1$, then consider an affine map h which transforms $\mathcal{E}$ into a circle. Let $\Delta a'b'c'$ be the image of Δabc by h. It is known that α, β, γ are proportional with the sidelengths of the triangle $\Delta a'b'c'$. Thus, the tangency points of $h(\mathcal{E})$ with respect to $\Delta a'b'c'$ divide its sides into ratios

$$\frac{\alpha+\gamma-\beta}{\alpha+\beta-\gamma}, \frac{\alpha+\beta-\gamma}{\beta+\gamma-\alpha}, \frac{\beta+\gamma-\alpha}{\alpha+\gamma-\beta}.$$

The affine map h does not modify the ratios of collinear segments, thus, $\mathcal{E}$ divides the sides of Δabc into the same ratios. $\square$

2 Inellipses and critical points of logarithmic potentials

The properties of inellipses described above allow us to state and prove a result which is a bit more general than Siebeck-Marden Theorem. In fact, every inellipse relates to the critical points of a logarithmic potential of the form

$$L(z) = \alpha \log(z-a) + \beta \log(z-b) + \gamma \log(z-c).$$

The following result gives a precise description of this connection.

Theorem 3. *Given Δabc and $\alpha, \beta, \gamma > 0$ with $\alpha + \beta + \gamma = 1$, the function $L(z) = \alpha \log(z-a) + \beta \log(z-b) + \gamma \log(z-c)$ has two critical points f_1 and f_2. These critical points are the foci of an inellipse which divides the sides of Δabc into ratios $\beta/\gamma, \gamma/\alpha, \alpha/\beta$.*

Conversely, given an inellipse $\mathcal{E}$ for Δabc, there exists a function of the form $L(z)$ like above whose critical points f_1, f_2 are the foci of $\mathcal{E}$.

Proof. Denote f_1, f_2 the roots of

$$L'(z) = \frac{\alpha}{z-a} + \frac{\beta}{z-b} + \frac{\gamma}{z-c},$$

which means that f_1, f_2 are roots of

$$z^2 - (\alpha(b+c) + \beta(a+c) + \gamma(a+b))z + \alpha bc + \beta ca + \gamma ab = 0.$$

Without loss of generality, we can suppose that $a = 0$ and that the imaginary axis is the bisector of the angle $\angle bac$ (equivalently $bc < 0$). In this case we have $f_1 f_2 = \alpha bc < 0$, and thus the imaginary axis is the bisector of the angle $\angle f_1 a f_2$. Repeating the same argument for b and c we deduce that f_1, f_2 are isogonal conjugates relative to Δabc. Steiner's result (Theorem 1) implies that f_1, f_2 are the foci of an inellipse $\mathcal{E}$ for Δabc. The center of this inellipse has barycentric coordinates

$$o = \frac{\beta+\gamma}{2}a + \frac{\alpha+\gamma}{2}b + \frac{\alpha+\beta}{2}c,$$

which, according to Theorem 2, implies that $\mathcal{E}$ is the unique inellipse for Δabc, which divides the sides of Δabc in ratios $\beta/\gamma, \gamma/\alpha, \alpha/\beta$.

Conversely, given an inellipse $\mathcal{E}$ for Δabc, its tangency points must be of the form $\beta/\gamma, \gamma/\alpha, \alpha/\beta$ for some $\alpha, \beta, \gamma > 0$, $\alpha + \beta + \gamma = 1$. We choose $L(z) = \alpha \log(z-a) + \beta \log(z-b) + \gamma \log(z-c)$ and, according to the first part of the proof, the critical points f_1, f_2 of $L'(z)$ are the foci of an ellipse $\mathcal{E}'$ which divides the sides of Δabc into ratios $\beta/\gamma, \gamma/\alpha, \alpha/\beta$. This means that $\mathcal{E} = \mathcal{E}'$ and $L(z)$ is the associated logarithmic potential. $\square$

References

[1] R. E. Allardice. Note on the dual of a focal property of the inscribed ellipse. *Ann. of Math. (2)*, 2(1-4):148–150, 1900/01.

[2] H. Aref and M. Brøns. On stagnation points and streamline topology in vortex flows. *J. Fluid Mech.*, 370:1–27, 1998.

[3] E. Badertscher. A simple direct proof of Marden's theorem. *Amer. Math. Monthly*, 121(6):547–548, 2014.

[4] G. Chakerian. A distorted view of geometry. In R. Honsberger, editor, *Mathematical Plums*, volume 4 of *The Dolciani Mathematical Expositions*, pages ix+182. Mathematical Association of America, Washington, D.C., 1979.

[5] V. Dragović and M. Radnović. *Poncelet porisms and beyond*. Frontiers in Mathematics. Birkhäuser/Springer Basel AG, Basel, 2011. Integrable billiards, hyperelliptic Jacobians and pencils of quadrics.

[6] M. J. Kaiser. The dynamics and internal geometry of the three-city noxious location problem. *Math. Comput. Modelling*, 17(10):81–98, 1993.

[7] D. Kalman. An elementary proof of Marden's theorem. *Amer. Math. Monthly*, 115(4):330–338, 2008.

[8] M. Marden. A note on the zeros of the sections of a partial fraction. *Bull. Amer. Math. Soc.*, 51:935–940, 1945.

[9] D. Minda and S. Phelps. Triangles, ellipses, and cubic polynomials. *Amer. Math. Monthly*, 115(8):679–689, 2008.

[10] S. Northshield. Geometry of cubic polynomials. *Math. Mag.*, 86(2):136–143, 2013.

[11] J. L. Parish. On the derivative of a vertex polynomial. *Forum Geom.*, 6:285–288 (electronic), 2006.

[12] J. Siebeck. Ueber eine neue analytische behandlungweise der brennpunkte. *J. Reine Angew. Math. 64*, 1864.

[13] J. Steiner. Géométrie pure. Développement d'une série de théorèmes relatifs aux sections coniques. *Ann. Math. Pures Appl. [Ann. Gergonne]*, 19:37–64, 1828/29.