Möbius transformations and ellipses

Adam Coffman* Marc Frantz†

September 6, 2006

1 Introduction

If $T : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$ is a Möbius transformation of the extended complex plane, it is well-known that the image under T of a line or circle is another line or circle. It seems natural to consider the image $T(\mathcal{E})$ of a non-circular ellipse $\mathcal{E} \subseteq \mathbb{C}$, although as shown in Figure 1, such a curve is not always an ellipse. For the sake of convenience, we will call a curve $\mathcal{C}$ a "möte," for "Möbius Transformation of an Ellipse," if $\mathcal{C} = T(\mathcal{E})$ for some non-circular ellipse $\mathcal{E}$ and Möbius transformation T. We will also call two curves $\mathcal{C}_1$ and $\mathcal{C}_2$ in $\mathbb{C} \cup \{\infty\}$ "Möbius equivalent" if there exists a Möbius transformation T such that $\mathcal{C}_2 = T(\mathcal{C}_1)$. Our main result is that two mötes, $T_1(\mathcal{E}_1)$ and $T_2(\mathcal{E}_2)$, are Möbius equivalent if and only if $\mathcal{E}_1$ and $\mathcal{E}_2$ are ellipses with the same eccentricity. In this sense, the eccentricity is an invariant of an ellipse not only under similarity transformations of the plane, but also under the larger group of Möbius transformations. In the last Section we briefly consider some other special plane curves in the extended complex plane.

2 Möbius transformations

We begin by summarizing some facts about Möbius transformations. The formula for a Möbius transformation $T : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$ is:

$$T(z) = \frac{az + b}{cz + d}, \tag{1}$$

or

$$T(z) = \frac{a\bar{z} + b}{c\bar{z} + d}, \tag{2}$$

with complex coefficients that satisfy $ad - bc \neq 0$, and the usual conventions for ∞ input and output, so that T is continuous, considering $\mathbb{C} \cup \{\infty\}$ as the Riemann sphere. Any such T is one-to-one and onto, and $\mathcal{M}$, the set of all Möbius

*Indiana University - Purdue University Fort Wayne
†Indiana University Bloomington

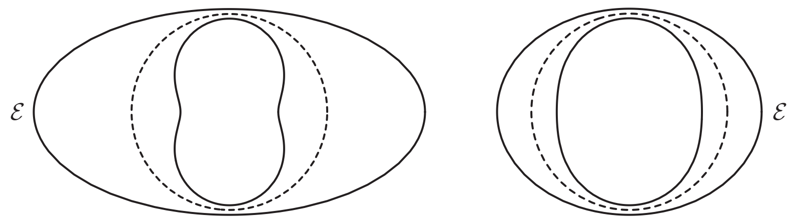

Figure 1: Ellipse $\mathcal{E}$ with a non-convex central inversion (left), and with a convex central inversion (right).

transformations, is a group under composition. For $T \in \mathcal{M}$, the following are equivalent: i) $T(\infty) = \infty$, ii) $c = 0$, and iii) T restricted to $\mathbb{C}$ is a similarity transformation of the plane. The similarities form a subgroup of $\mathcal{M}$.

Inversion in a circle is also a special case of a Möbius transformation. For a circle in $\mathbb{C}$ with center $q \in \mathbb{C}$ and radius $r > 0$, the formula for the inversion in that circle is $T(z) = \frac{r^2}{\bar{z}-\bar{q}} + q$.

Finally, we recall that Möbius transformations are conformal — they preserve angles where curves intersect (preserving or reversing the orientation of the angle in cases (1), (2), respectively), and they also preserve tangency.

From this point, we will use the term "ellipse" only for non-circular ellipses in the plane $\mathbb{C}$, and we will use the term "circle" to refer to both circles in $\mathbb{C}$ and extended straight lines, so a Möbius transformation takes a circle to a circle.

3 Central Inversions

As a special case of a möte, consider the image of an ellipse $\mathcal{E}$ under an inversion in a circle $\mathcal{C}$ so that $\mathcal{C}$ and $\mathcal{E}$ have the same center. Such a möte generated by a central inversion has appeared in different areas of geometry — it was described by Proclus [13] as a special case of the intersection of a torus and a plane, so it is called a "hippopede of Proclus."

For sufficiently eccentric ellipses, it is clear from Figure 1 that the interior of a hippopede need not be convex, and so a hippopede is not necessarily an ellipse. In fact, no hippopede is an ellipse, but this is not so obvious for the less eccentric ellipse in Figure 1, which is close to a circle, and so is its oval-shaped image under inversion. Our first lemma proves this fact, considering without loss of generality an ellipse centered at the origin, and inversion in the unit circle.

Lemma 1 *If $\mathcal{E}$ is an ellipse centered at $0 \in \mathbb{C}$, and $T(z) = 1/\bar{z}$, then the image $T(\mathcal{E})$ is not an ellipse.*

PROOF: For any rotation around the origin, $R(z) = e^{i\theta} \cdot z$, an easy calculation checks $T = R^{-1} \circ T \circ R$. Since $T(\mathcal{E})$ is just a rotation of the möte $T(R(\mathcal{E}))$, $T(\mathcal{E})$

is an ellipse if and only if $T(R(\mathcal{E}))$ is an ellipse. So, we can assume $\mathcal{E}$ has been rotated into "standard position," with its major diameter on the real axis, and given in terms of the real coordinates (x,y) (the components of the complex coordinate $z = x + iy$) by:

$$\mathcal{E} = \left\{(x,y) : \frac{x^2}{A^2} + \frac{y^2}{B^2} = 1\right\}, \tag{3}$$

with $0 < B < A$.

In the group of similarity transformations, the möte $T(\mathcal{E})$ has the same symmetries as the ellipse: reflections in the major and minor axes, and a 180° rotational symmetry. This is easily seen in Figure 1, and also follows from a short calculation: if S is any of the similarity transformations $z \mapsto \bar{z}$, $z \mapsto -\bar{z}$, or $z \mapsto -z$, then $S \circ T = T \circ S$, so $S(T(\mathcal{E})) = T(S(\mathcal{E})) = T(\mathcal{E})$. If $T(\mathcal{E})$ is an ellipse, then its axes, which are its lines of reflection symmetry, must be the coordinate axes, and $T(\mathcal{E})$ contains the points $T(\pm A + 0i) = \pm\frac{1}{A} + 0i$ and $T(0 \pm Bi) = 0 \pm \frac{1}{B}i$, so it must have an equation of the form $A^2x^2 + B^2y^2 = 1$. However, it is easy to check that the point $(A+Bi)/\sqrt{2}$ is on $\mathcal{E}$, but $T((A+Bi)/\sqrt{2}) = \sqrt{2}(A+Bi)/(A^2+B^2)$ does not satisfy the equation $A^2x^2 + B^2y^2 = 1$ unless $A = B$. ∎

4 An intersection property

The following lemma will be useful in the proof of the main result. It states some familiar (or at least plausible) properties of conics, but it will also lead to a similar property of mötes as an immediate corollary.

Lemma 2 *A circle $\mathcal{C}$ and a conic $\mathcal{L}$ that does not contain $\mathcal{C}$ can have at most four points of intersection, and at most two points of tangency.*

PROOF: Recall that "circle" includes straight lines, but we take for granted that lines have the claimed properties. The circle $\mathcal{C}$, with finite radius $r > 0$ and center (U,V), admits a parametric equation of the form

$$\mathbf{p} : \mathbb{R} \to \mathbb{R}^2 : t \mapsto \left(\frac{2rt}{1+t^2} + U, \frac{r(1-t^2)}{1+t^2} + V\right). \tag{4}$$

The image $\mathbf{p}(\mathbb{R})$ covers the whole circle except one point, $(U, V - r)$. If this point happens to be an element of $\mathcal{L}$, there is some rotation transformation R which fixes the circle, so that $R((U, V - r)) \notin \mathcal{L}$. $R \circ \mathbf{p}$ is still given by quadratic rational functions of t, so if necessary we can replace $\mathbf{p}$ by $R \circ \mathbf{p}$ to get a parametrization of $\mathcal{C}$ that contains all the points of intersection. If $q(x,y) = 0$ is a non-zero quadratic implicit equation for $\mathcal{L}$, the composition $(q \circ \mathbf{p})(t)$ is zero exactly at the points where $\mathbf{p}(t)$ meets $\mathcal{L}$. Expanding $q \circ \mathbf{p}$ gives a rational function $\frac{N(t)}{D(t)}$ whose denominator is never zero, and whose numerator has degree at most 4. Since $\lim\limits_{t\to\infty} \mathbf{p}(t)$ exists and is not in $\mathcal{L}$,

$$\lim_{t\to\infty} q(\mathbf{p}(t)) = q(\lim_{t\to\infty} \mathbf{p}(t)) \neq 0,$$

so N is not identically zero, and there are at most four points of intersection. A point of intersection $\mathbf{p}(t_0)$ is a point of tangency exactly when $\frac{d\mathbf{p}}{dt}(t_0)$ is orthogonal to the gradient, $(\nabla q)(\mathbf{p}(t_0))$. By the chain rule,

$$\left(\frac{d}{dt}(q \circ \mathbf{p})\right)(t_0) = (\nabla q)(\mathbf{p}(t_0)) \cdot \left(\frac{d\mathbf{p}}{dt}(t_0)\right),$$

so the curve is tangent to $\mathcal{L}$ when t_0 is a root of both $(q \circ \mathbf{p})(t)$ and $\frac{d}{dt}(q \circ \mathbf{p})(t) = \frac{N'(t)}{D(t)} - \frac{N(t)D'(t)}{(D(t))^2}$, which implies t_0 is a root of $N(t)$ and $N'(t)$, and therefore a double root of the quartic $N(t)$; there can be at most two such double roots. ∎

Corollary 3 *Any möte $T(\mathcal{E})$ and any circle $\mathcal{C}$ can meet in at most four points, and can have at most two points of tangency.*

Proof: Applying T^{-1} gives $\mathcal{E}$, a non-circular ellipse, and a circle $T^{-1}(\mathcal{C})$, which by Lemma 2 can intersect in at most four points. Since Möbius transformations preserve tangency, the second claim also follows from the Lemma. ∎

In particular, a möte can meet a straight line in at most four points, and without going into the details, any möte is contained in the zero set of some polynomial in the (x, y) coordinates with degree at most 4 — obviously ellipses have degree 2, but the points on a möte that is a hippopede of Proclus satisfy an irreducible, degree 4 implicit equation.

5 The main results

The main result mentioned in the Introduction will be a consequence of the following Theorem.

Theorem 4 *If two ellipses $\mathcal{E}$, $\mathcal{E}'$ are Möbius equivalent, with $T(\mathcal{E}) = \mathcal{E}'$, then T is a similarity.*

Proof: If $T(\mathcal{E}) = \mathcal{E}'$, and R is a reflection symmetry of $\mathcal{E}'$, then $R \circ T$ also takes $\mathcal{E}$ onto $\mathcal{E}'$. So, without loss of generality we can assume $T \in \mathcal{M}$ is of the form (2), and we want to show that $c = 0$.

An ellipse has four points lying on its axes of reflection symmetry. We call these points "vertices," and any similarity transformation of the plane takes the vertices of an ellipse to the vertices of the image ellipse.

Since the similarities form a subgroup of $\mathcal{M}$, we can also assume that $\mathcal{E}$ is in standard position, Equation (3). Its vertices are the points $(\pm A, 0)$ and $(0, \pm B)$, or in complex coordinates, $\pm A$ and $\pm iB$.

Let $\mathcal{C}$ be a circle tangent to an ellipse $\mathcal{L}$ at a point z_0. We say that $\mathcal{C}$ is *bitangent* at z_0 if there is some point $z_1 \neq z_0$ so that $\mathcal{C}$ is tangent to $\mathcal{L}$ at z_1. From Figure 2, it is easy to see that at any point z_0 that is not a vertex, there are at least two bitangent circles: one interior, centered on the major axis, and one exterior, centered on the minor axis.

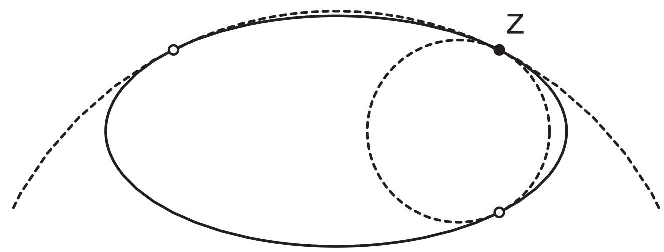

Figure 2: Bitangent circles (dashed) with a common point of tangency.

At a vertex of an ellipse, there is only one bitangent circle: if the other point of tangency were any point besides the opposite vertex, then by the reflection symmetry of the ellipse and circle, there would be a third point of tangency, contradicting Lemma 2.

Let z_0 be a vertex of $\mathcal{E}$, but, suppose toward a contradiction that $T(z_0)$ is not a vertex of $\mathcal{E}'$. Then, at $T(z_0)$, $\mathcal{E}'$ has two bitangent circles, and there are distinct points $z_1, z_2 \in \mathcal{E}$ so that these circles are tangent to $\mathcal{E}'$ at $T(z_0)$ and $T(z_1)$, and $T(z_0)$ and $T(z_2)$, respectively. Since T^{-1} preserves circles and the tangency of curves, these two bitangent circles of $\mathcal{E}'$ at $T(z_0)$ are transformed to two bitangent circles of $\mathcal{E}$ at z_0, contradicting the previous observation that there is only one bitangent circle at the vertex z_0 of $\mathcal{E}$. So, $T(A)$, $T(-A)$, $T(iB)$, and $T(-iB)$ must be the vertices of $\mathcal{E}'$.

The images of the axes of $\mathcal{E}$ must therefore be circles orthogonal to the ellipse $\mathcal{E}'$ at its vertices. Since $\mathcal{E}'$ is non-circular, these circles must be straight lines, the axes of $\mathcal{E}'$, so the pairs $T(A)$, $T(-A)$, and $T(iB)$, $T(-iB)$ are the pairs of opposing vertices of $\mathcal{E}'$.

Now suppose again by way of contradiction that $c \neq 0$. By the previous paragraph, the point $-\overline{d/c}$ which T maps to ∞ must lie on both the major and minor axes of $\mathcal{E}$, so $d = -\overline{d/c} = 0$. Hence T has the form $T(z) = a + b/\bar{z}$, with $b \neq 0$. Since $z \mapsto a + bz$ is a similarity, the image of $\mathcal{E}$ under the map $z \mapsto 1/\bar{z}$ must be an ellipse, however, this contradicts Lemma 1. ∎

The first Corollary is just the special case of the Theorem where $\mathcal{E} = \mathcal{E}'$. It appears as a statement without proof in [16].

Corollary 5 *The only Möbius transformations that are symmetries of an ellipse are similarities.* ∎

The next Corollary is the result on the invariance of the eccentricity.

Corollary 6 *If two mötes $T_1(\mathcal{E}_1)$ and $T_2(\mathcal{E}_2)$ are Möbius equivalent, then $\mathcal{E}_1$ and $\mathcal{E}_2$ have the same eccentricity.*

Proof: The hypothesis is that there exists $T \in \mathcal{M}$ such that $T(T_1(\mathcal{E}_1)) = T_2(\mathcal{E}_2)$, and this implies $\mathcal{E}_2 = T_2^{-1}(T(T_1(\mathcal{E}_1)))$. By the Theorem, the composition

$T_2^{-1} \circ T \circ T_1$ is a similarity transformation, so $\mathcal{E}_2$ and $\mathcal{E}_1$ are similar and have the same eccentricity. ∎

The notion that eccentricity is an invariant of mötes seems to have been known to the authors of [4] and [9], but not explicitly stated in that way. Their interest was in the description of conics by quadratic rational parametric maps, analogous to formula (4), and how these rational maps transform under Möbius transformations.

Corollary 7 *Any möte is either an ellipse or the image of an ellipse under inversion in a circle, and these cases are mutually exclusive.*

Proof: First, the mutually exclusive part follows from the Theorem: for ellipses $\mathcal{E}$, $\mathcal{E}'$, there can be no inversion R such that $R(\mathcal{E}) = \mathcal{E}'$, since R is a Möbius transformation that is not a similarity.

If the möte $T(\mathcal{E})$ is not an ellipse, then T is not a similarity, so T is of the form (1) or (2) with $c \neq 0$, and $T(\infty) = a/c$. We can compose T with an inversion that sends a/c back to ∞: let $R(z) = \frac{1}{\bar{z} - a/c} + \frac{a}{c}$. Then a short calculation will check that $R \circ T$ is a similarity transformation, so $R(T(\mathcal{E}))$ is an ellipse $\mathcal{E}'$, and it follows from $R = R^{-1}$ that $T(\mathcal{E}) = R(\mathcal{E}')$. ∎

The Famous Curves web site [11] at St. Andrews has an interactive demonstration where the user can manipulate a circle of inversion to see various images of the ellipse in the plane. Corollary 7 says that every möte can be generated this way, and the reader is encouraged to experiment with this to get an idea of what shapes a möte may have.

6 The other conics

In addition to ellipses, one might also consider hyperbolas and parabolas, and in the Riemann sphere, these conic curves become closed only when they also include the point at infinity. The topic of images of conics under inversions in circles is frequently addressed in books ([8], [17], [18]) and web sites ([5], [6], [11]) on plane curves, and the same curves also appear as "pedal curves" of conics ([1]). The term "hippopede of Proclus" applies to the central inversion of either an ellipse or a hyperbola — but in the hyperbolic case, the curve is self-intersecting, as in Figure 3.

The self-intersection will appear in any Möbius transformation of a hyperbola (a *möth*?), although for similarities, the self-intersection remains at the point ∞. This makes an analogue of Theorem 4 for hyperbolas easier to prove than the ellipse case. The interior angle in either of the loops is a conformal invariant and depends only on the eccentricity of the hyperbola, so the analogue of Corollary 6 is also easy: the eccentricity is a Möbius invariant for hyperbolas. So, we see the key to the proof of Theorem 4 was to show that an ellipse has distinguished points (the vertices) with respect to Möbius transformation geometry, even though an ellipse in $\mathbb{C}$ has no points with distinguishing topological or conformal properties.

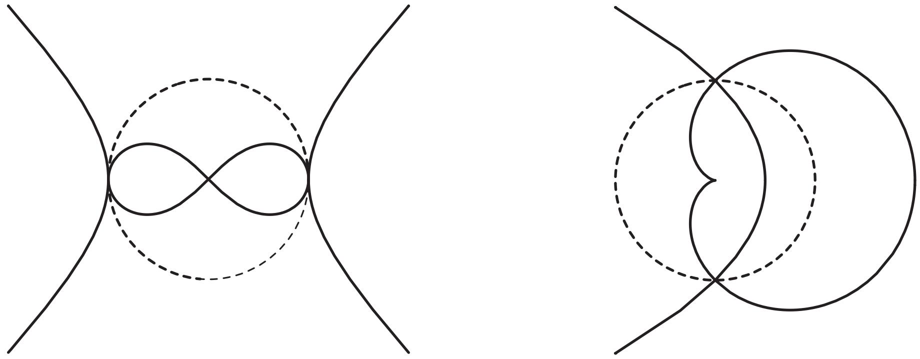

Figure 3: Two conics and their inverses with respect to a circle. Hyperbola and lemniscate (left); parabola and cardioid (right).

A Möbius transformation of a parabola is seen to have a cusp singularity, as in Figure 3, which distinguishes the image from the other Möbius transformations of conics. References [7], [10], and [12] use the term "nodal biquadratic" to refer to a curve that is Möbius equivalent to an ellipse or hyperbola, and "cuspidal biquadratic" to refer to a curve that is Möbius equivalent to a parabola. We can generalize the main result, Corollary 6, as follows:

Proposition 8 *Let $\mathcal{L}_1$ be a non-degenerate conic (circle, ellipse, parabola, or hyperbola), $\mathcal{L}_2$ another non-degenerate conic, and let T_1 and T_2 be Möbius transformations. If $T_1(\mathcal{L}_1)$ and $T_2(\mathcal{L}_2)$ are Möbius equivalent, then $\mathcal{L}_1$ and $\mathcal{L}_2$ have the same eccentricity.* ∎

In addition to its original description as an intersection of a plane with a torus, the elliptic hippopede also occurs as an intersection of a plane with Fresnel's elasticity surface ([15]). The arclength of both types of hippopede curves was a topic of investigation by Booth, [2], [3], so they have also been called "lemniscates of Booth," of elliptic and hyperbolic types. They also appear in some interesting applications, for example, in mechanical linkages ([5], [18]) and fluid physics ([14]). Some of the references on special plane curves refer to the elliptic hippopede of Proclus as merely an "oval,"[1] but some other special cases of inverses of conics are more famously named curves, including the limaçons of Etienne Pascal, as shown in the following table.

[1]Quoting E. Fermi, "If I could remember the names of all these particles, I would have been a botanist."

Curve	Eccentricity	Center of Inversion	Image
hyperbola	$e > 1$	center	hyperbolic lemniscate of Booth = self-intersecting hippopede
hyperbola	$e > 1$	focus	self-intersecting limaçon
hyperbola	$e = 2$	focus	trisectrix
hyperbola	$e = 2$	vertex	trisectrix of Maclaurin
hyperbola	$e = \sqrt{2}$	vertex	right strophoid
hyperbola	$e = \sqrt{2}$	center	lemniscate of Bernoulli
parabola	$e = 1$	focus	cusped limaçon = cardioid
parabola	$e = 1$	vertex	cissoid of Diocles
ellipse	$0 < e < 1$	center	elliptic lemniscate of Booth = simple hippopede
ellipse	$0 < e < 1$	focus	simple limaçon

References

[1] R. Alperin, *A grand tour of pedals of conics*, Forum Geometricorum **4** (2004), 143–151. MR 2130228.

[2] J. Booth, *Researches on the geometrical properties of elliptic integrals*, Philosophical Transactions of the Royal Society of London **142** (1852), 311–416, and **144** (1854), 53–69.

[3] J. Booth, *A Treatise on Some New Geometrical Methods*, Longmans, Green, Reader, and Dyer, London, Vol. I, 1873, and Vol. II, 1877.

[4] R. Deaux, *Introduction to the Geometry of Complex Numbers*, transl. from the revised French ed. by H. Eves, Ungar Pub. Co., New York, 1957. MR 0086745 (19,236f).

[5] R. Ferréol, *Encyclopédie des Formes Mathématiques Remarquables*, `www.mathcurve.com/courbes2d/courbes2d.shtml`

[6] S. Gray and S. Venit, *National Curve Bank*,

`curvebank.calstatela.edu/home/home.htm`

[7] E. Kasner, *The invariant theory of the inversion group: geometry upon a quadric surface*, Trans. AMS (4) **1** (1900), 430–498. MR 1500550.

[8] J. D. Lawrence, *A Catalog of Special Plane Curves*, Dover, 1972.

[9] F. Morley and F. V. Morley, *Inversive Geometry*, Chelsea, New York, 1954.

[10] F. Morley and B. Patterson, *On algebraic inversive invariants*, American J. of Math. (2) **52** (1930), 413–424. MR 1506765.

[11] J. O'Connor, E. Robertson, and B. Soares, Famous Curves Index, MacTutor History of Mathematics Archive,

www-history.mcs.st-and.ac.uk/history/Curves/Curves.html

[12] B. Patterson, *The inversive plane*, Amer. Math. Monthly (9) **48** (1941), 589–599. MR 0006034 (3,251c).

[13] Proclus, *A Commentary on the First Book of Euclid's Elements*, transl. by G. Morrow, Princeton, 1970. MR 0269463 (42 #4359).

[14] S. Richardson, *Some Hele-Shaw flows with time-dependent free boundaries*, J. Fluid Mech. **102** (1981), 263–278. MR 0612095 (82d:76023).

[15] D. Struik, *Lectures on Classical Differential Geometry*, reprint of 1961 2nd ed., Dover, New York, 1988. MR 0939369 (89b:53002).

[16] J. Wilker, *Möbius equivalence and Euclidean symmetry*, Amer. Math. Monthly (4) **91** (1984), 225–247. MR 0736877 (85b:51007).

[17] R. Yates, *Curves and Their Properties*, reprint of *A Handbook on Curves and Their Properties*, 1952, NCTM Classics in Mathematics Education, 1974. MR 0023075 (9,301b).

[18] C. Zwikker, *The Advanced Geometry of Plane Curves and Their Applications*, Dover, New York, 1963. MR 0166651 (29 #3924).

About the authors:

Adam Coffman is an Associate Professor of Mathematics at IPFW, and a recently inducted member of the state's newest ΠΜΕ chapter, Indiana Zeta. His web site www.ipfw.edu/math/Coffman/ shows computer graphics that illustrate his research in geometry, including some pictures related to the topic of this article.

Marc Frantz is a Research Associate in the Department of Mathematics at Indiana University. He previously taught mathematics at IUPUI and had a thirteen-year career as a painter and picture framer. He currently has an NSF grant that includes co-writing a math & art text (with Annalisa Crannell) that will be published by Princeton University Press.

Adam Coffman

Department of Mathematical Sciences, Indiana University - Purdue University Fort Wayne, Fort Wayne, IN 46805-1499. CoffmanA@ipfw.edu

Marc Frantz

Department of Mathematics, Indiana University, Bloomington, IN 47405. mfrantz@indiana.edu