The Poncelet Grid and Billiards in Ellipses

Mark Levi and Serge Tabachnikov

1. INTRODUCTION: THE PONCELET CLOSURE THEOREM. The Poncelet closure theorem (or Poncelet porism) is a classical result of projective geometry. Given nested ellipses γ and Γ, with γ inside Γ, one plays the following game: starting at a point x on Γ, draw a tangent line to γ until it intersects Γ at point y, repeat the construction, starting with y, and so on. One obtains a polygonal curve inscribed in Γ and circumscribed about γ. Suppose that this process is periodic: the nth point coincides with the initial one. Now start at a different point, say x_1. The Poncelet closure theorem states that the polygonal line again closes up after n steps (see Figure 1). We call these closed inscribed-circumscribed curves *Poncelet polygons*. Although the Poncelet theorem is almost two hundred years old,[1] it continues to attract interest (see [**2**], [**3**], [**4**], [**5**], [**6**], [**9**], [**12**], [**14**], [**15**], [**19**], [**23**], [**24**] for a sample of references).

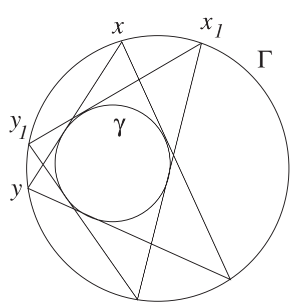

Figure 1. Poncelet polygons: even for (nonconcentric) circles, this is a non-trivial theorem.

It is hard to believe that one can still add anything new to such a well-studied subject! However, recently Rich Schwartz [**21**] discovered the following property of Poncelet polygons. Extending the sides of a Poncelet n-gon, one obtains a set of points called the *Poncelet grid* (see Figure 2, which is borrowed from [**21**]). The points of the Poncelet grid can be viewed as lying on a family of nested closed curves and also on a family of disjoint curves having radial directions.

More precisely, let $\ell_1, \dots, \ell_n$ be the lines containing the sides of the Poncelet polygon, enumerated in such a way that their tangency points to γ are in cyclic order. The Poncelet grid consists of the $n(n+1)/2$ points $\ell_i \cap \ell_j$. The indices are understood cyclically and, by convention, $\ell_j \cap \ell_j$ is the tangency point of ℓ_j with γ. Define sets P_k and Q_k as follows:

$$P_k = \cup_{i-j=k} \ell_i \cap \ell_j, \quad Q_k = \cup_{i+j=k} \ell_i \cap \ell_j. \tag{1}$$

The cases of odd and even n differ somewhat and, as in [**21**], we assume that n is odd. There are $(n+1)/2$ sets P_k, each containing n points, and n sets Q_k, each containing $(n+1)/2$ points.

[1] Poncelet discovered this result in 1813–14, when he was a prisoner of war in the Russian city of Saratov. He published his theorem in 1822, upon his return to France.

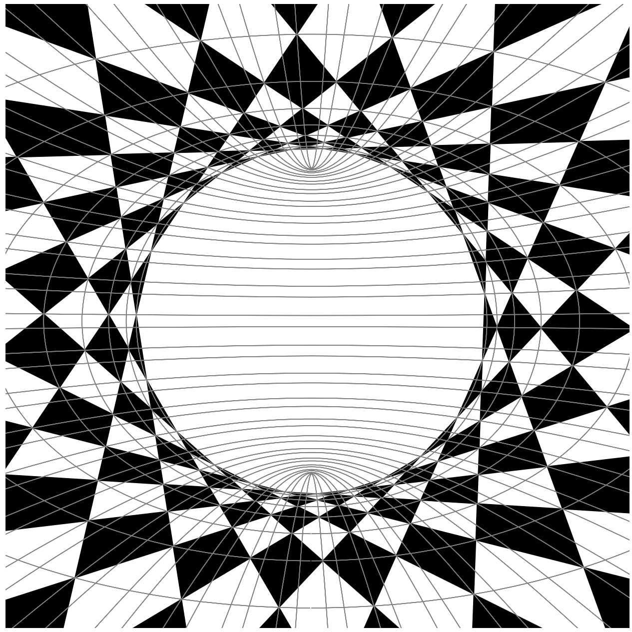

Figure 2. Poncelet grid.

The Schwartz theorem reads as follows:

Theorem 1. *The sets P_k lie on nested ellipses,*[2] *and the sets Q_k lie on disjoint hyperbolas. The complexified versions of these ellipses and hyperbolas have four common complex tangent lines. Furthermore, all the sets P_k are projectively equivalent to each other, and all the sets Q_k are projectively equivalent to each other.*

The proof in [**21**] consists in a study of properties of the underlying elliptic curve. In this article we give a different, more elementary proof, which deduces this theorem from properties of billiards in ellipses.

2. MATHEMATICAL BILLIARDS. In this section we survey (mostly with proofs) necessary facts about billiards. Detailed discussions can be found in [**11**], [**13**], [**24**] or [**25**]. The billiard system describes the motion of a free point inside a plane domain: the point moves with a constant speed along a straight line until it hits the boundary, at which point it reflects according to the familiar law of geometrical optics "the angle of incidence equals the angle of reflection."

We assume that the billiard table is a convex domain with a smooth boundary curve Γ. The billiard ball map acts on oriented lines that intersect the billiard table, sending the incoming billiard trajectory to the outgoing one. Let x, y, and z be points on Γ such that the line segment xy reflects to the line segment yz. The equal angles condition has a variational meaning.

Lemma 1. *The angles made by lines xy and yz with Γ (i.e., with the line tangent to Γ at y) are equal if and only if y is a critical point of the function $f(y) = |xy| + |yz|$ (where $|xy|$ signifies the Euclidean distance between x and y).*

[2] This part of the theorem was known to Darboux (see [**7**]).

Proof. Assume first that y is a free point, not confined to Γ. The gradient of the function $|xy|$ is the unit vector from x to y, and the gradient of $|yz|$ is the unit vector from z to y. By the Lagrange multipliers principle, y in Γ is a critical point of the function $f(y) = |xy| + |yz|$ if and only if the sum of the two gradients is orthogonal to Γ, which is equivalent to the fact that xy and yz make equal angles with Γ. ■

The space of oriented lines in the plane has a remarkable area form. An oriented line is characterized by its direction φ in $[0, 2\pi]$ and its signed distance p from an origin (see Figure 3). The coordinates (φ, p) identify the space of oriented lines with a cylinder on which the area form is $\omega = dp \wedge d\varphi$. This area form ω is invariant under isometries of the plane (and, up to a factor, is characterized by this invariance). It is widely used in integral geometry, for example, in the Crofton formula [**20**].

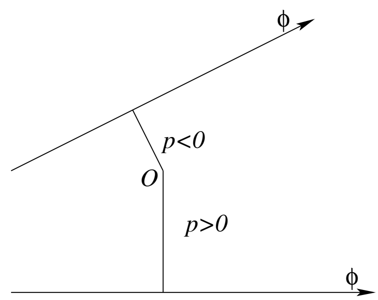

Figure 3. Coordinates in the space of oriented lines in the plane.

One can introduce a different coordinate system on the subset of oriented lines in the plane consisting of the lines that intersect the billiard table (i.e., the space on which the billiard ball map acts). Consider the curve Γ as described by the arclength parameter t, say starting at some specified point of Γ and tracing Γ out in the counterclockwise orientation. An oriented line L intersects the curve Γ twice; let $\Gamma(t)$ be the first of these intersection points, and let α in $[0, \pi]$ be the angle between L and the tangent vector $\Gamma'(t)$. Then (t, α) provide coordinates for the space of oriented lines that intersect the billiard table. It can be shown that, in these new coordinates, $\omega = \sin\alpha \, d\alpha \wedge dt$. We omit this rather straightforward computation (see, for example, [**24**] or [**25**]).

A fundamental property of the billiard ball map is its area-preserving character:

Theorem 2. *The area form ω is invariant under the billiard ball map.*

Proof. Let $\Gamma(t_1)$ be the second intersection point of an oriented line L with the curve Γ, and let α_1 be the angle between the line and the vector $\Gamma'(t_1)$. Then the billiard ball map sends (t, α) to (t_1, α_1).

Denote by $f(t, t_1)$ the distance between points $\Gamma(t)$ and $\Gamma(t_1)$. The partial derivative $\partial f/\partial t_1$ is the projection of the gradient of the distance $|\Gamma(t)\Gamma(t_1)|$ on the curve at point $\Gamma(t_1)$. This gradient is the unit vector in the direction from $\Gamma(t)$ to $\Gamma(t_1)$, and it makes angle α_1 with the curve, whence $\partial f/\partial t_1 = \cos\alpha_1$. Likewise, $\partial f/\partial t = -\cos\alpha$. Therefore

$$df = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial t_1} dt_1 = -\cos\alpha \, dt + \cos\alpha_1 \, dt_1,$$

so

$$0 = d^2 f = \sin\alpha \; d\alpha \wedge dt - \sin\alpha_1 \; d\alpha_1 \wedge dt_1.$$

This means that ω is an invariant area form.[3] ■

Remark 1. The billiard ball map is an example of a *discrete Lagrangian system* (see, for example, [**22**] or [**26**]). A discrete Lagrangian system on a manifold M is determined by a smooth function $f : M \times M \to \mathbf{R}$, a *Lagrangian*, satisfying certain convexity conditions. The Lagrangian determines a map $T : M \times M \to M \times M$ given by a variational principle: $T(x_0, x_1) = (x_1, x_2)$ if

$$f_2(x_0, x_1) + f_1(x_1, x_2) = 0, \tag{2}$$

where the subscript 1 or 2 indicates differentiation with respect to the first or the second variable, respectively.[4] In the case of billiards, M is a (topological) circle, the boundary of the billiard table, and f is the chord length (see Lemma 1). The map T has an invariant differential two-form. To obtain this form, take the exterior derivative of both sides of equation (2),

$$f_{12}(x_0, x_1)\; dx_0 + (f_{22}(x_0, x_1) + f_{11}(x_1, x_2))\; dx_1 + f_{12}(x_1, x_2)\; dx_2 = 0,$$

and then take the wedge product with dx_1 to see that

$$f_{12}(x_0, x_1)\; dx_0 \wedge dx_1 = f_{12}(x_1, x_2)\; dx_1 \wedge dx_2.$$

Thus

$$\omega = f_{12}(x_0, x_1)\; dx_0 \wedge dx_1$$

is a T-invariant two-form (in this argument, we use vector notation: $f_x dx$ means $f_{x^1} dx^1 + f_{x^2} dx^2 + \cdots$ in local coordinates, $dx_0 \wedge dx_1$ means $dx_0^1 \wedge dx_1^1 + dx_0^2 \wedge dx_1^2 + \cdots$, etc.). In the billiard case, we recover the area form introduced earlier.

Another fact about billiards that we require for our discussion concerns caustics. A *caustic* is a curve C inside a billiard table with the following property: if a segment of a billiard trajectory is tangent to C, then so is each reflected segment (see Figure 4). We assume that caustics are smooth and convex.

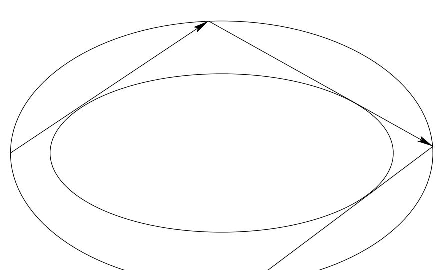

Figure 4. A caustic.

[3] We use in this argument simple properties of differential forms, in particular, the definition of the exterior differential d and its property $d^2 = 0$.

[4] The index i in $(x_0, x_1, \ldots, x_i, \ldots)$ is the discrete analog of time t for a continuous Lagrangian system with the position variable $x(t)$.

Consider a caustic γ of a billiard table with boundary Γ. Suppose that we "erase" the table and only the caustic remains. Can we recover Γ from γ? The answer is given by the following *string construction*: wrap a closed nonelastic string around γ, pull it tight at a point, and move this point around γ to obtain a curve Γ (see Figure 5).

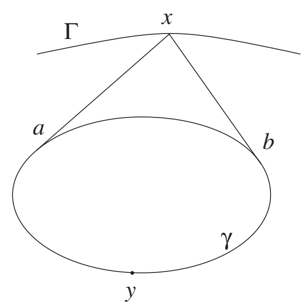

Figure 5. String construction.

Theorem 3. *If Γ is a curve generated by a string construction from a given convex curve γ, then the region bounded by Γ has γ as a caustic.*

Proof. Choose a reference point y on γ. For a point x in the plane, let $f(x)$ and $g(x)$ be the distances from x to y by going along a string from y to x on the left and on the right, respectively. Then Γ is a level curve of the function $f + g$. We want to prove that the angles made with Γ by the segments ax and bx depicted in Figure 5 are equal.

We claim that the gradient of f at x is the unit vector in the direction of ax. Indeed, ax is the direction of the fastest increase of f, and the directional derivative $D_{ax} f = 1$. Likewise, the gradient of g at x is the unit vector in the direction of bx. It follows that $\nabla(f + g)$ bisects the angle axb. Therefore ax and bx make equal angles with Γ. ∎

Note that the string construction provides a one-parameter family of billiard tables: the parameter is the length of the string. Note also that, by the same reasoning, the level curves of the function $f - g$ are orthogonal to Γ.

3. BILLIARDS IN ELLIPSES.

The optical properties of conics were already known to the ancient Greeks. In this section we examine their implications for billiards in ellipses.

First of all, recall the geometric definition of an ellipse: it is the locus E of points the sum of whose distances from two given points F_1 and F_2, called the foci of E, is constant. An ellipse can be constructed using a string whose ends are fixed at the foci (Figure 6). A hyperbola is defined similarly, but the sum of distances is replaced with the absolute value of their difference. Taking the segment F_1F_2 as γ in Theorem 3, it

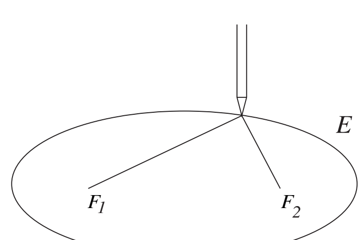

Figure 6. Construction of an ellipse.

follows that a ray passing through one focus reflects to a ray passing through the other focus.

The construction of an ellipse with given foci has a parameter, the length of the string. A family of conics all of which share the same foci is called *confocal*. An equation that describes a confocal family that includes both ellipses and hyperbolas is

$$\frac{x_1^2}{a_1^2+\lambda}+\frac{x_2^2}{a_2^2+\lambda}=1, \tag{3}$$

where λ is a parameter.

Fix foci F_1 and F_2. For a generic point P in the plane there exist a unique ellipse and a unique hyperbola with foci F_1 and F_2 that passes through P. The ellipse and the hyperbola are orthogonal to each other: this follows from the fact that the sum of two unit vectors is perpendicular to their difference (recall the proofs of Lemma 1 and Theorem 3).

The next theorem says that the billiard ball map in an ellipse is completely integrable, that is, it possesses an invariant quantity.[5]

Theorem 4. *A billiard trajectory inside an ellipse E remains forever tangent to a fixed confocal conic. More precisely, if some segment of a billiard trajectory does not intersect the focal segment F_1F_2 of E, then no segment of this trajectory intersects F_1F_2 and all segments are tangent to the same ellipse E' with foci F_1 and F_2; if some segment of a trajectory intersects F_1F_2, then all segments of this trajectory intersect F_1F_2 and all are tangent to the same hyperbola H with foci F_1 and F_2.*

Proof. We learned the following elementary geometric proof from [**10**] (the Russian original appeared about thirty five years ago). Let A_0A_1 and A_1A_2 be consecutive segments of a billiard trajectory. Assume that A_0A_1 does not intersect the segment F_1F_2; the other case is dealt with similarly. It follows from the optical property of an ellipse that the segments F_1A_1 and F_2A_1 make equal angles with the ellipse. Thus $\angle A_0A_1F_1 = \angle A_2A_1F_2$ (Figure 7).

Reflect F_1 in A_0A_1 to F_1', and F_2 in A_1A_2 to F_2', and set

$$B = F_1'F_2 \cap A_0A_1,\ C = F_2'F_1 \cap A_1A_2.$$

Consider the ellipse with foci F_1 and F_2 that is tangent to A_0A_1. Since $\angle F_2BA_1 = \angle F_1'BA_0 = \angle F_1BA_0$, this ellipse touches A_0A_1 at the point B. Likewise an ellipse with foci F_1 and F_2 touches A_1A_2 at the point C. One wants to show that these two ellipses coincide or, equivalently, that $F_1B + BF_2 = F_1C + CF_2$, which boils down to $F_1'F_2 = F_1F_2'$.

We claim that the triangles $F_1'A_1F_2$ and $F_1A_1F_2'$ are congruent. First, $F_1'A_1 = F_1A_1$ and $F_2A_1 = F_2'A_1$ by symmetry. In addition, $\angle F_1'A_1F_2 = \angle F_1A_1F_2'$. Indeed, $\angle A_0A_1F_1 = \angle A_2A_1F_2$, as we remarked earlier, $\angle A_0A_1F_1 = \angle A_0A_1F_1'$ and $\angle A_2A_1F_2 = \angle A_2A_1F_2'$ by symmetry, and the angle $F_1A_1F_2$ is a common part of the angles $F_1'A_1F_2$ and $F_1A_1F_2'$. Hence $F_1'F_2 = F_1F_2'$, and the result follows. ■

We infer that the billiard ball map inside an ellipse has a one-parameter family of caustics consisting of confocal ellipses. Theorems 3 and 4 also imply the following

[5]The billiard ball map inside a multidimensional ellipsoid is completely integrable as well (see, for example, [**16**]).

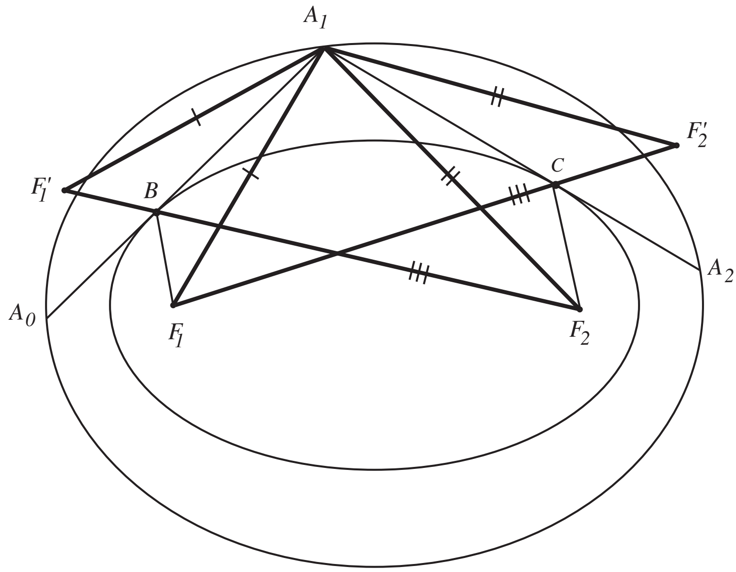

Figure 7. Integrability of the billiard in an ellipse.

theorem of Graves: *wrapping a closed nonelastic string around an ellipse produces a confocal ellipse* (see [**3**], [**18**]).

As noted earlier, the space of oriented lines intersecting an ellipse E is topologically a cylinder. This cylinder is foliated by invariant curves of the billiard ball map, as shown in Figure 8 on the left. Each curve represents the family of rays tangent to a fixed confocal conic. The ∞-shaped curve corresponds to the family of rays through the foci of E. The two singular points of this curve represent the major axis with two opposite orientations, a period-two billiard trajectory. Another period-two trajectory is the minor axis, which is represented by the two centers of the regions inside the ∞-shaped curve. The invariant curves outside the ∞-shaped curve correspond to the rays that are tangent to confocal ellipses, while the invariant curves inside the ∞-shaped curve represent the rays that are tangent to confocal hyperbolas. For comparison, we also give a (much simpler) phase portrait of the billiard ball map in a circle (Figure 8 on the right).

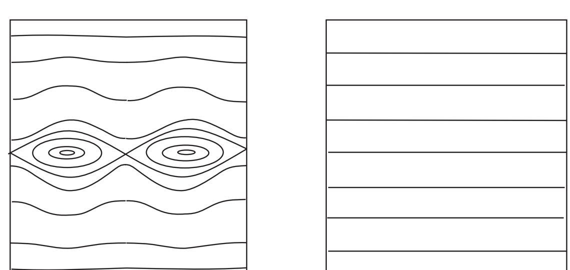

Figure 8. Phase space of the billiard ball map in an ellipse and in a circle.

The integrability of the billiard ball map makes it possible to choose a cyclic coordinate on each invariant curve (say, x modulo 1) such that the map is given by a shift $x \mapsto x + c$, in which the value of the constant c depends on the invariant curve. This construction plays the central role in our paper, and its multidimensional analog lies at the heart of the Arnol'd-Liouville theorem in the theory of completely integrable systems [**1**].

Choose a function f on the cylinder whose level curves are the invariant curves of the billiard ball map. Let γ be the curve with equation $f = C$, and let the curve γ_ε be given by $f = C + \varepsilon$. For an interval I on γ consider the area $\omega(I, \varepsilon)$ between γ and γ_ε over I. Define the "length" of I as

$$\lim_{\varepsilon \to 0} \frac{\omega(I, \varepsilon)}{\varepsilon}.$$

Choosing a different function f multiplies the length of every segment on the curve γ by the same factor. Select a coordinate x so that the length element is dx; this coordinate is well defined up to an affine transformation. Normalizing x so that the total length is 1 determines x up to a shift $x \mapsto x + c$ (we do not discuss explicit formulas for the parameter x, which involves elliptic integrals; in what follows, we obtain numerous geometric consequences from the mere fact that such a parameter exists). Note that the value of the constant c depends on the invariant curve γ.

The billiard ball map preserves the area element ω and the invariant curves. Therefore it preserves the length element on the invariant curves, that is, the billiard map is given by the formula $x \mapsto x + c$.

We summarize. Consider an ellipse Γ and a confocal ellipse γ, a caustic for the billiard ball map in Γ. This map can be regarded as a self-map of γ (it sends point a to point b in Figure 5). We have introduced a parameter x on γ in which the billiard ball map is a shift $x \mapsto x + c$. The choice of the parameter x depends only on the area form in the space of oriented lines and the foliation of this space by the curves consisting of tangent lines to confocal ellipses. That is, *the parameter x depends only on γ, and not on* Γ. In contrast, the billiard ball map and therefore the constant c depend on the ellipse Γ as well.

Let Γ' be another confocal ellipse containing γ. Then Γ and Γ' share the family of caustics, in particular, the ellipse γ. Therefore the billiard ball map associated with Γ' is also a shift in the parameter x.

Corollary 1. *The billiard ball maps associated with* Γ *and* Γ' *commute.*

Proof. The shifts $x \mapsto x + c$ and $x \mapsto x + c'$ commute (see Figure 9). ∎

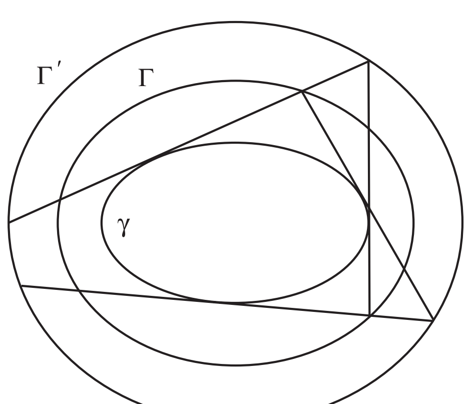

Figure 9. Commuting billiard ball maps.

In particular, let γ degenerate to the segment through the foci. Then the rays in Figure 9 pass through the foci, and Corollary 1 implies the following result, described by Pedoe [**17**] as the "most elementary theorem of Euclidean geometry", which was discovered by M. Urquhart [**8**]: $AB + BF = AD + DF$ if and only if $AC + CF = AE + EF$ (Figure 10, left). The reader is challenged to find an elementary proof of this theorem.

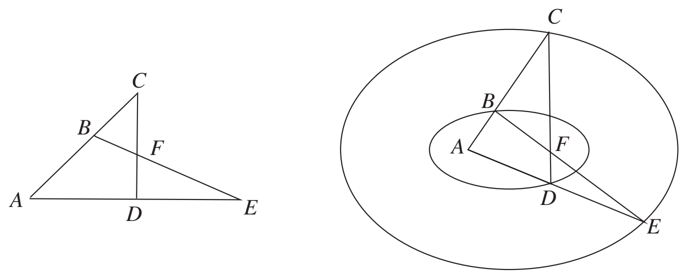

Figure 10. The most elementary theorem of Euclidean geometry.

A second consequence is the following Poncelet-style closure theorem:

Corollary 2. *If some billiard trajectory in an ellipse Γ is tangent to a confocal ellipse γ and has period n, then every billiard trajectory in Γ tangent to γ has period n.*

Proof. In the appropriate coordinate on γ, the billiard ball map is $x \mapsto x + c$. A point has period n if and only if nc is an integer. This condition does not depend on x, and the result follows. ∎

Finally, although Corollary 2 is only a special case of the Poncelet porism, it implies the general version. This is because a generic pair of nested ellipses is projectively equivalent to a pair of confocal ones[6] (this proof of the Poncelet porism is mentioned in [**26**]).

More precisely, consider the complexified situation. Two conics have four common tangent lines, and one has a one-parameter family of conics sharing these four tangents.

Lemma 2. *A confocal family of conics consists of all conics tangent to four fixed lines.*

Proof. Recall the notion of projective duality (see, for example, [**3**]). Let $P = (a : b : c)$ be a point in the projective plane described by its homogeneous coordinates. Assign to P the line L_P whose equation, in the same homogeneous coordinate system, is $ax + by + cz = 0$. We obtain a one-to-one correspondence $P \mapsto L_P$ between points and lines in the projective plane. This correspondence preserves the incidence relation: if a point P lies on a line L_Q, then the point Q lies on the line L_P. This correspondence extends to smooth curves: a curve is a one-parameter family of points, they correspond to a one-parameter family of lines, and these lines envelope a curve, which is called *projectively dual* to the original one. Projective duality is duality indeed: applied twice, it yields the original curve.

[6]This means that there exists a projective transformation of the plane that takes two nested ellipses to confocal ones.

A curve projectively dual to a conic is a conic. The one-parameter family of conics, dual to the confocal family (3), is given by the equation

$$(a_1^2 + \lambda)x_1^2 + (a_2^2 + \lambda)x_2^2 = 1$$

(the reader is encouraged to make an explicit computation). This is an equation of a pencil, a one-parameter family of conics that pass through four fixed points, namely, the intersection points of the conics, $a_1^2 x_1^2 + a_2^2 x_2^2 = 1$ and $x_1^2 + x_2^2 = 0$, that is, the points $(1 : \pm i : \pm\sqrt{a_1^2 - a_2^2})$. Projective duality interchanges points and tangent lines; applied again, it yields a one-parameter family of conics sharing four tangent lines. ∎

Since projective transformations act transitively on quadruples of lines in general position, a generic pair of conics is projectively equivalent to a pair of confocal ones.

4. THE PONCELET GRID. Let γ and Γ be a pair of nested ellipses, and let P be a Poncelet n-gon circumscribing γ and inscribed in Γ. Applying a projective transformation, we may assume that γ and Γ are confocal.

Let x be the parameter on γ introduced in section 3. When the origin is chosen appropriately, the tangency points of the sides of P with γ have coordinates

$$0,\ \frac{1}{n},\ \frac{2}{n},\ \dots,\ \frac{n-1}{n}.$$

The set P_k in (1) lies on the locus of intersections of the tangent lines to γ at points $\gamma(x)$ and $\gamma(x + k/n)$, where x varies from 0 to 1. This locus is a confocal ellipse for which the billiard trajectories tangent to γ close up after n reflections and k turns around γ (periodic trajectories with rotation number k/n). Thus P_k lies on an ellipse confocal with γ.

Likewise, the set Q_k in (1) lies on the locus of intersections of the tangent lines to γ at points $\gamma(x)$ and $\gamma(k/n - x)$. We want to show that this locus is a confocal hyperbola. To this end we need the following result, which is an (apparently new) addition to Theorem 3, the string construction (see Figure 11).

Theorem 5. *If the string construction is applied to an oval γ,*[7] *if p and p' are distinct points on the resulting curve Γ, and if q and q' are the intersection points of the pairs of tangent lines ap, $b'p'$ and $a'p'$, $b'p$, then as $p' \to p$ the lines pp' and qq' become orthogonal.*

Corollary 3. *The locus of intersections of tangent lines to an ellipse γ at $\gamma(c - x)$ and $\gamma(c + x)$ is a hyperbola confocal with γ. In particular, the set Q_k lies on a confocal hyperbola.*

Proof. Since the points p and p' in Figure 11 lie on the confocal ellipse Γ, the "lengths" (measured via the parameter introduced in section 3) of the arcs aa' and bb' are equal. It follows from Theorem 5 that the locus in question is a curve orthogonal to the family of confocal ellipses, thus it is a confocal hyperbola. Therefore the set Q_k lies on a confocal hyperbola. ∎

Proof of Theorem 5. We give two arguments, one geometric and the other analytic. Let p and p' be infinitesimally close. By Theorem 3, the arc pp' bisects the angles

[7] An oval is a smooth closed strictly convex curve.

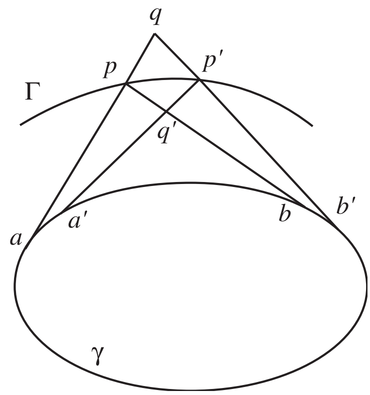

Figure 11. Addition to the string construction.

qpq' and $qp'q'$. If ε is the distance between p and p', we dilate the plane with factor $1/\varepsilon$. The angles do not change, the arc pp' becomes essentially straight, and as $\varepsilon \to 0$ we obtain a quadrilateral $pqp'q'$ that is symmetric with respect to the diagonal pp'. Hence pp becomes perpendicular to qq'.

Analytically, we compute how fast the points a and b of tangency determined by a point p outside γ move as we move p. Let the speeds of these points along γ be v_1 and v_2; let the tangent segments ap and bp have lengths l_1 and l_2; let the angular velocity of the lines ap and bp be ω_1 and ω_2; and let k_1 and k_2 be the curvatures of γ at a and b. Denote the velocity vector of point p by u.

Then $k_1 = \omega_1/v_1$ and $\omega_1 = u_1/l_1$, where u_1 is the component of u perpendicular to ap. An analogous relation holds for the variables k_2, ω_2, v_2, l_2, and u_2. It follows that

$$\frac{v_2}{v_1} = \frac{l_1 k_1}{l_2 k_2} \cdot \frac{u_2}{u_1}.$$

Consider two choices for u: one a vector tangent to Γ, the other a vector perpendicular to it. Because of the equal angles property (Theorem 3), in the first case we have $u_1 = u_2$ whereas in the second case $u_1 = -u_2$. Thus the ratios v_2/v_1 in the two cases have the same absolute value but opposite signs. This is equivalent to the orthogonality of pp' and qq'. ■

5. ELLIPTIC COORDINATES AND LINEAR EQUIVALENCE. It remains to show that the sets P_k are projectively (actually, linearly) equivalent for all values of k, and likewise for the sets Q_k. Given an ellipse γ, we again let x be the parameter described in section 3. Note that the map $x \mapsto x + 1/2$ is the central symmetry with respect to γ. In particular, the tangent lines at points $\gamma(x)$ and $\gamma(x + 1/2)$ are parallel.

For a point P outside of γ we draw the tangent segments PA and PB to the ellipse and write $x - y$ and $x + y$ for the coordinates of the points A and B, respectively, where $0 \le y < 1/4$. Then (x, y) are coordinates for P. We proved in section 4 that the coordinate curves $y = c$ and $x = c$, where c is a constant, are ellipses and hyperbolas, respectively, that are confocal with γ.

As in section 4, the Poncelet grid is made by intersecting the tangent lines at points $\gamma(i/n)$ $(i = 0, 1, \ldots, n-1)$. The (x, y)-coordinates of the points of the grid are

$$\left(\frac{k}{2n} + \frac{j}{n}, \frac{k}{2n}\right) \left(k = 0, 1, \ldots, \frac{n-1}{2};\ j = 0, 1, \ldots, n-1\right).$$

Fixing the second coordinate yields an angular set P, while fixing the first one gives rise to a radial set Q.

An ellipse E with equation

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1$$

also determines *elliptic coordinates* in the plane. Through a point P there passes a unique hyperbola and a unique ellipse from the confocal family of conics (3). The elliptic coordinates of P are the respective values λ_1 and λ_2 of the parameter λ in (3). The hyperbolas and ellipses from the confocal family (3) are the coordinate curves of this coordinate system, $\lambda_1 = c$ and $\lambda_2 = c$, respectively. The Cartesian coordinates of P are expressed in terms of the elliptic ones as follows:

$$x_1^2 = \frac{(a_1^2 + \lambda_1)(a_1^2 + \lambda_2)}{a_1^2 - a_2^2}, \ x_2^2 = \frac{(a_2^2 + \lambda_1)(a_2^2 + \lambda_2)}{a_2^2 - a_1^2} \tag{4}$$

(the Cartesian coordinates are determined up to the symmetries of E: $(x_1, x_2) \mapsto (\pm x_1, \pm x_2)$).

Thus the coordinates (x, y) and the elliptic coordinates (λ_1, λ_2) have the same coordinate curves, families of ellipses and hyperbolas confocal with γ. It follows that λ_1 is a function of x and λ_2 a function of y.

Let Γ_λ and Γ_μ be two ellipses (or two hyperbolas) from a confocal family of conics (3). Consider the linear map $A_{\lambda,\mu}$ whose matrix relative to the standard basis for $\mathbf{R}^2$ is

$$\mathrm{Diag}\left(\sqrt{\frac{a_1^2 + \mu}{a_1^2 + \lambda}}, \sqrt{\frac{a_2^2 + \mu}{a_2^2 + \lambda}}\right).$$

This map transforms Γ_λ to Γ_μ. The following lemma is classical and goes back to J. Ivory.[8]

Lemma 3. *If Γ_λ and Γ_μ are two ellipses* (*respectively, two hyperbolas*) *from the family* (3) *and if P is a point of Γ_λ, then the points P and $Q = A_{\lambda,\mu}(P)$ lie on the same confocal hyperbola* (*respectively, ellipse*).

Proof. Unfortunately, we do not know a geometrical proof, so our argument will be computational. We consider the case when Γ_λ and Γ_μ are ellipses, the hyperbolic case being similar. Let (λ_1, λ_2) and (μ_1, μ_2) be the elliptic coordinates of points P and Q. Then $\lambda_2 = \lambda$ and $\mu_2 = \mu$. We want to prove that $\lambda_1 = \mu_1$.

If (x_1, x_2) and (X_1, X_2) are the Cartesian coordinates of P and Q then we have the formulas in (4) and the similar relations

$$X_1^2 = \frac{(a_1^2 + \mu_1)(a_1^2 + \mu_2)}{a_1^2 - a_2^2}, \ X_2^2 = \frac{(a_2^2 + \mu_1)(a_2^2 + \mu_2)}{a_2^2 - a_1^2}. \tag{5}$$

On the other hand, $Q = A_{\lambda,\mu}(P)$, whence

[8]Ivory was studying the gravitational potential of the infinitely thin shell between homothetic ellipsoids, the so-called homeoid.

$$X_1^2 = \frac{a_1^2 + \mu}{a_1^2 + \lambda} x_1^2 = \frac{(a_1^2 + \lambda_1)(a_1^2 + \mu_2)}{a_1^2 - a_2^2},$$

and likewise for X_2^2. Combined with (5), this yields $\lambda_1 = \mu_1$, as claimed. ∎

Now we can prove that the sets P_k and P_m are linearly equivalent: the equivalence is given by the maps $\pm A_{\lambda,\mu}$, where the sign depends on whether $k - m$ is even or odd. The argument for the sets Q_k is similar.

The respective (x, y)-coordinates of the sets P_k and P_m are

$$\left(\frac{k}{2n} + \frac{j}{n}, \frac{k}{2n}\right), \left(\frac{m}{2n} + \frac{j}{n}, \frac{m}{2n}\right) \ (j = 0, 1, \ldots, n - 1).$$

Now P_k and P_m lie on confocal ellipses Γ_λ and Γ_μ, respectively. According to Lemma 3, the map $A_{\lambda,\mu}$ preserves the first elliptic coordinate, hence the x-coordinate. Therefore the coordinates of the points of the set $A_{\lambda,\mu}(P_k)$ are

$$\left(\frac{k}{2n} + \frac{j}{n}, \frac{m}{2n}\right) \ (j = 0, 1, \ldots, n - 1).$$

If m has the same parity as k, this coincides with the set P_m, whereas if the parity of m is opposite to that of k, then this set is centrally symmetric to the set P_m.

ACKNOWLEDGMENTS. Many thanks to Rich Schwartz for fruitful discussions of his beautiful theorem and to the referees and Bruce Palka for their helpful suggestions. We are grateful to the Mathematics Institute at Oberwolfach for its hospitality; the second author also thanks Max-Planck-Institut in Bonn for its support. Both authors were partially supported by NSF grants, DMS-9704554 and DMS-0244720, respectively.

REFERENCES

1. V. Arnol'd, *Mathematical Methods of Classical Mechanics*, Springer-Verlag, New York, 1989.
2. W. Barth and T. Bauer, Poncelet theorems, *Expos. Math.* **14** (1996) 125–144.
3. M. Berger, *Geometry*, Springer-Verlag, Berlin, 1987.
4. H. Bos, C. Kers, F. Oort, and D. Raven, Poncelet's closure theorem, *Expos. Math.* **5** (1987) 289–364.
5. S.-J. Chang and R. Friedberg, Elliptical billiards and Poncelet's theorem, *J. Math. Phys.* **29** (1988) 1537–1550.
6. S.-J. Chang, K. J. Shi, Billiard systems on quadric surfaces and the Poncelet theorem, *J. Math. Phys.* **30** (1989) 798–804.
7. G. Darboux, *Principes de Géométrie Analytique*, Gauthier-Villars, Paris, 1917.
8. D. Elliott and M. L. Urquhart, *J. Aust. Math. Soc.* **8** (1968) 129–133.
9. P. Griffith and J. Harris, A Poncelet theorem in space, *Comm. Math. Helv.* **52** (1977) 145–160.
10. V. Gutenmacher and N. Vasilyev, *Lines and Curves*, Birkhäuser Verlag, Basel, 2004.
11. A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, Cambridge, 1995.
12. J. King, Three problems in search of a measure, this MONTHLY **101** (1994) 609–628.
13. V. Kozlov and D. Treshchev, *Billiards. A Genetic Introduction to the Dynamics of Systems with Impacts*, American Mathematical Society, Providence, RI, 1991.
14. B. Mirman, Numerical ranges and Poncelet curves, *Linear Algebra Appl.* **281** (1998) 59–85.
15. ———, Sufficient conditions for Poncelet polygons not to close, this MONTHLY **112** (2005) 351–356.
16. J. Moser, Geometry of quadrics and spectral theory, in *The Chern Symposium*, Springer-Verlag, New York, 1980, pp. 147–188.
17. D. Pedoe, The most "elementary" theorem of Euclidean geometry, *Math. Mag.* **49** (1976) 40–42.
18. K. Poorrezaei, Two proofs of Graves's theorem, this MONTHLY **110** (2003) 826–830.
19. E. Previato, Some integrable billiards. SPT 2002: Symmetry and perturbation theory, 181–195, World Sci. Publ., River Edge, NJ, 2002.

20. L. Santalo, *Integral Geometry and Geometric Probability*, Cambridge University Press, Cambridge, 2004.
21. R. Schwartz, The Poncelet grid, *Adv. Geometry* **7** (2007) 157-175.
22. Y. Suris, *The Problem of Integrable Discretization: Hamiltonian Approach*, Birkhäuser Verlag, Basel, 2003.
23. S. Tabachnikov, Poncelet's theorem and dual billiards, *L'Enseign. Math.* **39** (1993) 189–194.
24. ———, *Billiards*, Société Mathémathique de France, Panoramas et Synthèses, no. 1, Paris, 1995.
25. ———, *Geometry and Billiards*, American Mathematical Society, Providence, RI, 2005.
26. A. Veselov, Integrable mappings, *Russ. Math. Surv.* **46** (1991) 1–51.

MARK LEVI is a professor in the Department of Mathematics at the Pennsylvania State University. He received his Ph.D. from Courant Institute under the direction of Jürgen Moser. His research interests include differential equations and dynamical systems with applications to physical problems. One of his hobbies is creating and collecting physical devices that he uses to illustrate mathematical phenomena. This collection includes a 7-foot cycloidal trough used to demonstrate three remarkable properties of the cycloid discovered by Huygens over 300 years ago. His other hobby is inventing and collecting physical solutions to mathematical problems.
Department of Mathematics, The Pennsylvania State University, University Park, PA 16802
levi@math.psu.edu

SERGE TABACHNIKOV is a professor in the Department of Mathematics at the Pennsylvania State University. He received his Ph.D. from the Moscow State University in 1987; since 1990 he works in the USA. His research interests include differential topology, differential geometry, and dynamical systems, especially mathematical billiards. In the late 1980s he was the Head of Mathematics Department of "Kvant" (Quantum), a monthly Russian magazine devoted to physics and mathematics and somewhat similar to this Monthly. Since 2000, he is the Director of the Mathematics Advanced Study Semesters (MASS) program at Penn State.
Department of Mathematics, The Pennsylvania State University, University Park, PA 16802
tabachni@math.psu.edu