The Poncelet grid and the billiard in an ellipse

Mark Levi and Serge Tabachnikov
Department of Mathematics, Penn State University
University Park, PA 16802, USA

1 The closure theorem and the Poncelet grid

The Poncelet closure theorem (or Poncelet porism) is a classical result of projective geometry. Given two nested ellipses, γ and Γ, one plays the following game: choose 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 line, inscribed into Γ and circumscribed about γ. Suppose that this process is periodic: the n-th 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 will call these closed inscribed-circumscribed lines Poncelet polygons.

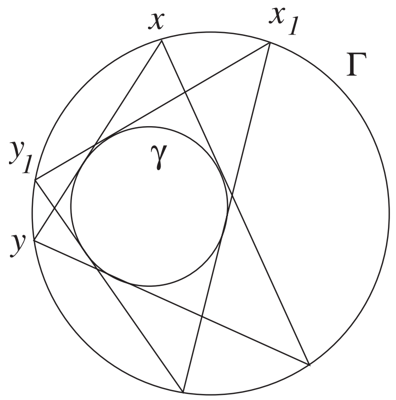

Figure 1: Poncelet polygons

Although the Poncelet theorem is almost 200 years old[1], it continues to attract interest: see [2, 3, 4, 5, 7, 10, 12, 13, 20, 21] for a sample of references.

It is hard to believe that one can still add anything new to such a well-studied subject! However, recently R. Schwartz [18] 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 borrowed from [18]. 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.

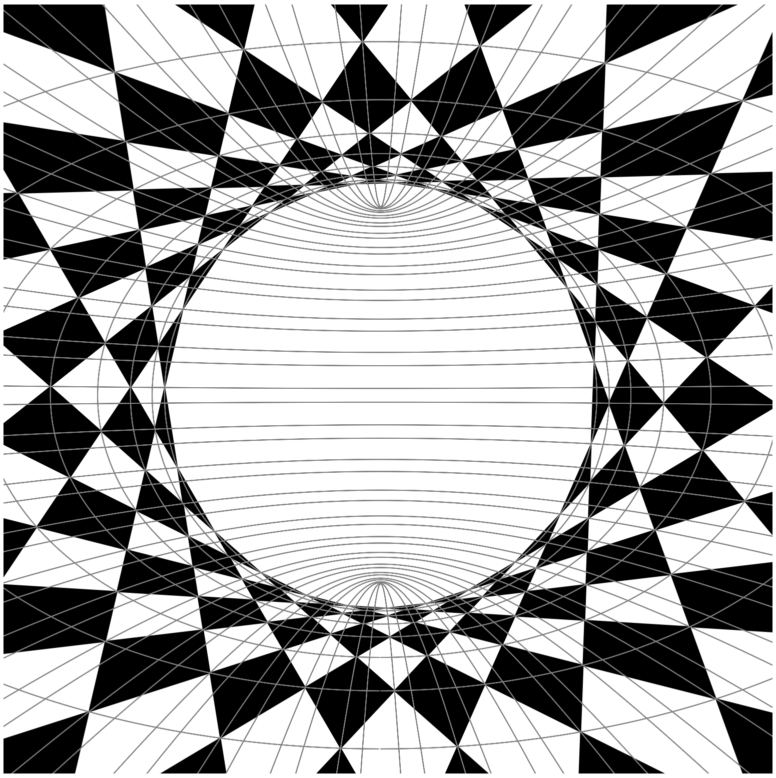

Figure 2: Poncelet grid

More precisely, let $\ell_1, \ldots, \ell_n$ be the lines containing the sides of the polygon, enumerated in such a way that their tangency points to γ are in cyclic order. The Poncelet grid consists of $n(n+1)/2$ points $\ell_i \cap \ell_j$. The indices

[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 returning to France.

are understood cyclically and, by convention, $\ell_j \cap \ell_j$ is the tangency point of ℓ_j with γ. Define the sets:

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

The cases of odd and even n differ somewhat and, as in [18], 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.

The Schwartz theorem states:

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

The proof in [18] consists in a study of properties of the underlying elliptic curve; we will give a different, more elementary, proof and deduce this theorem from properties of billiards in ellipses.

2 Mathematical billiards: general facts

In this section we survey (mostly, with proofs) necessary facts about billiards, see [9, 11, 21, 22] for detailed discussion.

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, where 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, 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 2.1 *The angles made by lines xy and yz with Γ are equal if and only if y is a critical point of the function $f(y) = |xy| + |yz|$ (where $|xy|$ is the Euclidean distance between x and y).*

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, $y \in \Gamma$ is a critical point of the function $|xy| + |yz|$ if and only if the sum of the two gradients is orthogonal to Γ, and this is equivalent to the fact that xy and yz make equal angles with Γ. $\Box$

The space of oriented lines in the plane has a remarkable area form. An oriented line is characterized by its direction φ and its signed distance p from an origin (the sign is determined by the right-hand rule). The coordinates (φ, p) identify the space of oriented lines with the cylinder; the area form is $\omega = dp \wedge d\varphi$. This area form is invariant under isometries of the plane (and, up to a factor, is characterized by this invariance); the form ω is widely used in integral geometry, for example, in the Crofton formula [17].

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 in the arc length parameter t. Let $\Gamma(t)$ be the first intersection point of an oriented line with the curve Γ and $\alpha \in [0, \pi]$ the angle between the line and the direction of the curve at point $\Gamma(t)$. Then (t, α) are coordinates on the space of oriented lines that intersect the billiard table.

One can prove that, in these new coordinates, $\omega = \sin\alpha \; d\alpha \wedge dt$. We omit this rather straightforward computation (see, e.g., [20, 21]).

A fundamental property of the billiard ball map is that it is area preserving.

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 with the curve Γ and α_1 the angle between the line and the curve at point $\Gamma(t_1)$. 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 from $\Gamma(t)$ to $\Gamma(t_1)$ and it makes angle α_1 with the curve; hence $\partial f/\partial t_1 = \cos\alpha_1$. Likewise, $\partial f/\partial t = -\cos\alpha$. Therefore

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

and hence
$$0 = d^2 f = \sin\alpha\ d\alpha \wedge dt - \sin\alpha_1\ d\alpha_1 \wedge dt_1.$$
This means that ω is an invariant area form. $\Box$

Remark 2.2 The billiard ball map is an example of a discrete Lagrangian system, see, e.g., [19, 23]. 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.[2] In the case of billiards, M is a circle, the boundary of the billiard table, and f is the chord length, cf. Lemma 2.1. The map T has an invariant differential 2-form. To obtain this form, take the exterior derivative 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 wedge multiply by dx_1:

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

Thus
$$\omega = f_{12}(x_0, x_1)\ dx_0 \wedge dx_1$$
is a T-invariant 2-form (in this argument, we use vector notation: $f_x dx$ means $f_{x^1} dx^1 + f_{x^2} dx^2 + \dots$, $dx_0 \wedge dx_1$ means $dx_0^1 \wedge dx_1^1 + dx_0^2 \wedge dx_1^2 + \dots$, in local coordinates, etc). In the billiard case, we recover the above introduced area form.

Another necessary fact about billiards concerns caustics. A caustic is a curve inside a billiard table such that if a segment of a billiard trajectory is tangent to this curve, then so is each reflected segment. We assume that caustics are smooth and convex.

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

Let Γ be the boundary of a billiard table and γ a caustic. Suppose that one erases the table, and only the caustic remains. Can one recover Γ from γ? The answer is given by the following *string construction*: wrap a closed non-stretchable string around γ, pull it tight at a point and move this point around γ to obtain a curve Γ.

Theorem 3 *Let Γ be a curve generated by a string construction from a given convex curve γ. Then the billiard inside Γ has γ as its caustic.*

Proof. Choose a reference point $y \in \gamma$. For a point $x \in \Gamma$, let $f(x)$ and $g(x)$ be the distances from x to y by going along the 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 by the segments ax and bx with Γ are equal; see figure 3.

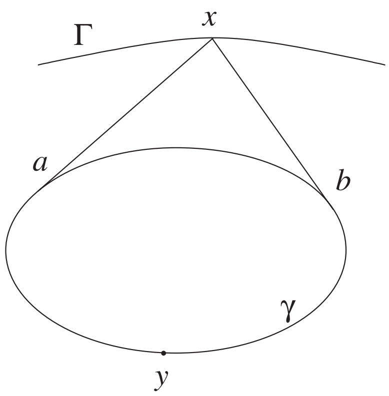

Figure 3: String construction

We claim that the gradient of f at x is the unit vector in the direction 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 bx. It follows that $\nabla(f + g)$ bisects the angle axb. Therefore ax and bx make equal angles with Γ. $\square$

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 curve of the function $f - g$ is orthogonal to Γ.

3 The billiard in an ellipse: integrability and its consequences

Optical properties of conics were already known to the Ancient Greeks. In this section we review billiards in ellipses and describe some consequences of their complete integrability.

First of all, recall the geometric definition of an ellipse: it is the locus of points whose sum of distances to two given points, F_1 and F_2, is fixed; these two points are called the foci. An ellipse can be constructed using a string whose ends are fixed at the foci, see figure 4. A hyperbola is defined similarly with the sum of distances replaced by the absolute value of their difference. Taking the segment F_1F_2 as γ in Theorem 3, it follows that a ray passing through one focus reflects to a ray passing through the other focus.

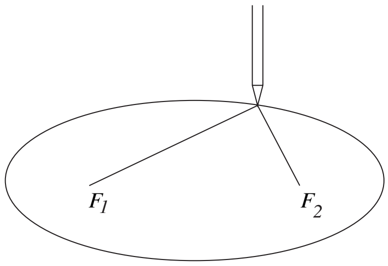

Figure 4: Gardener's construction of an ellipse

The construction of an ellipse with given foci has a parameter, the length of the string. The family of conics with fixed foci is called confocal. The equation of a confocal family, including 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 F_1 and F_2. Given a generic point in the plane, there exist a unique ellipse and a unique hyperbola with foci F_1, F_2 passing through the point. 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 its difference; cf. proofs of Lemma 2.1 and Theorem 3.

The next theorem says that the billiard ball map in an ellipse is integrable, that is, possesses an invariant quantity.[3]

Theorem 4 *A billiard trajectory inside an ellipse forever remains tangent to a fixed confocal conic. More precisely, if a segment of a billiard trajectory does not intersect the segment F_1F_2, then all the segments of this trajectory do not intersect F_1F_2 and are all tangent to the same ellipse with foci F_1 and F_2; and if a segment of a trajectory intersects F_1F_2, then all the segments of this trajectory intersect F_1F_2 and are all tangent to the same hyperbola with foci F_1 and F_2.*

Proof. We learned the following elementary geometrical proof from [8] (the Russian original appeared about 35 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; so the segments A_0A_1 and A_1A_2, and hence the angles $A_0A_1F_1$ and $A_2A_1F_2$ are equal, see figure 5.

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 the angles F_2BA_1 and F_1BA_0 are equal, 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. Indeed, $F_1'A_1 = F_1A_1$ 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 above, $\angle A_0A_1F_1 = \angle A_0A_1F_1'$ 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. □

Thus the billiard inside an ellipse has a 1-parameter family of caustics consisting of confocal ellipses. Theorems 3 and 4 imply the following Graves theorem: *wrapping a closed non-stretchable string around an ellipse produces a confocal ellipse*, see [3, 16].

[3]The billiard ball map inside a multi-dimensional ellipsoid is completely integrable as well, see, e.g., [14].

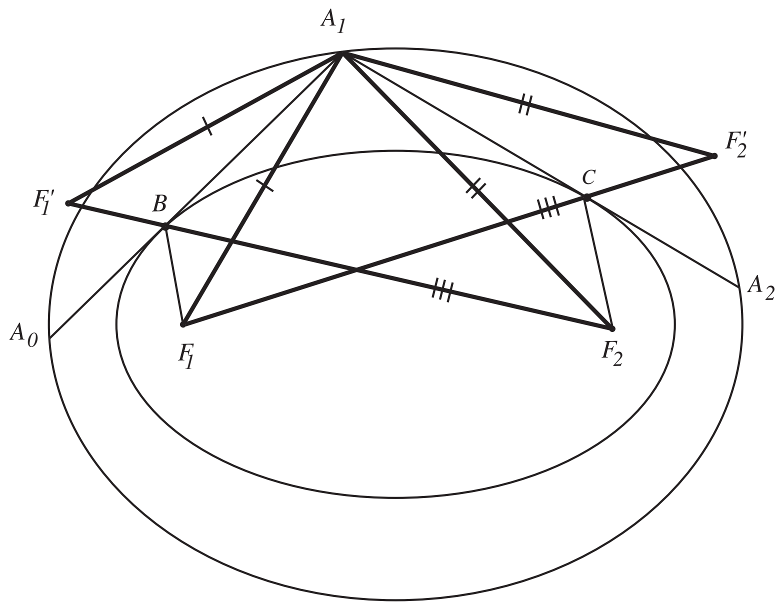

Figure 5: Integrability of the billiard in an ellipse

The space of oriented lines intersecting an ellipse is a cylinder. This cylinder is foliated by invariant curves of the billiard ball map, see figure 6 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. The two singular points of this curve represent the major axis with two opposite orientations, a 2-periodic billiard trajectory. Another 2-periodic trajectory is the minor axis represented by 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, and the invariant curves inside the ∞-shaped curve to 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, see figure 6 on the right.

The integrability of the billiard ball map makes it possible to choose a cyclic coordinate on each invariant curve, say, x mod 1, such that the map is given by a shift $x \mapsto x + c$; the value of the constant c depends on the invariant curve. This construction plays the central role in our paper, and its multi-dimensional analog is in the heart of the Arnold-Liouville theorem in the theory of completely integrable systems, see [1].

Choose a function f on the cylinder whose level curves are the invariant

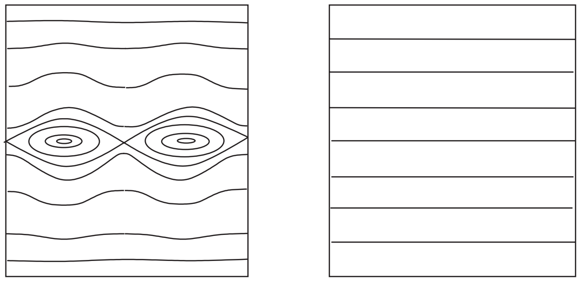

Figure 6: Phase space of the billiard ball map in an ellipse and in a circle

curves of the billiard ball map. Let γ be a curve $f = C$. Consider the curve γ_ε given by $f = C + \varepsilon$. For an interval $I \subset \gamma$, 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, one multiplies the length of every segment by the same factor. Choose 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 $x \mapsto x + const$ (we do not discuss explicit formulas for the parameter x, given by elliptic integrals; in what follows, we obtain numerous geometric consequences from the very fact that such a parameter exists).

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 $x \mapsto x + c$.

Let us summarize. Consider an ellipse Γ and a confocal ellipse γ, a caustic for the billiard in Γ. The billiard ball map can be considered as a self-map of γ (it sends point a to b in figure 3). 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 5 *The billiard ball maps associated with* Γ *and* Γ' *commute, see figure 7.*

Proof. The shifts $x \mapsto x + c$ and $x \mapsto x + c'$ commute. $\square$

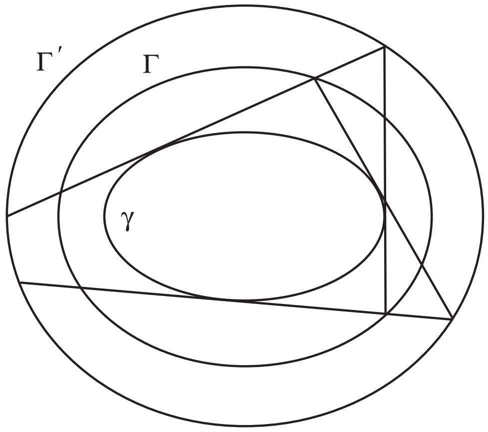

Figure 7: Commuting billiard ball maps

In particular, let γ degenerate to the segment through the foci. Then the rays in figure 7 pass through the foci, and Corollary 5 implies the following "most elementary theorem of Euclidean geometry"[4] discovered by M. Urquhart [6]: $AB + BF = AD + DF$ if and only if $AC + CF = AE + EF$; see figure 8, left. The reader is challenged to find an elementary proof of this theorem.

Another consequence is a Poncelet-style closure theorem.

Corollary 6 *Assume that a billiard trajectory in an ellipse* Γ*, tangent to a confocal ellipse* γ*, is* n*-periodic. Then every billiard trajectory in* Γ*, tangent to* γ*, is* n*-periodic.*

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

Finally, being only a particular case of the Poncelet porism, Corollary 6 implies its general version. This is because a generic pair of nested ellipses is

[4] The name coined by Pedoe [15].

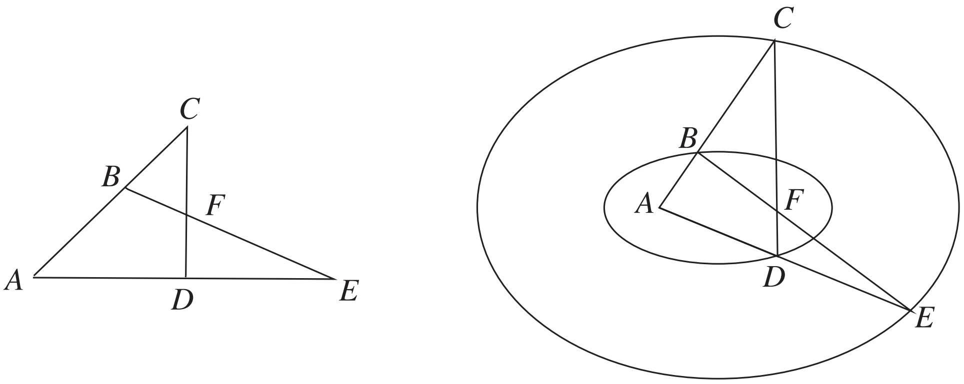

Figure 8: The most elementary theorem of Euclidean geometry

projectively equivalent to a pair of confocal ones (this proof of the Poncelet porism is mentioned in [23]).

More precisely, consider the complexified situation. Two conics have four common tangent lines, and one has a 1-parameter family of conics sharing these four tangents.

Lemma 3.1 *A confocal family of conics consists of the conics, tangent to four fixed lines.*

Proof. A curve, projectively dual to a conic, is a conic. The 1-parameter family of conics, dual to the confocal family (3), is given by the equation

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

This is an equation of a pencil, a 1-parameter family of conics that pass through four fixed points; these are the intersections of the two conics, $a_1^2x_1^2 + a_2^2x_2^2 = 1$ and $x_1^2 + x_2^2 = 1$. Projective duality interchanges points and tangent lines; applied again, it yields a 1-parameter family of conics sharing four tangent lines. □

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

4 Back to the Poncelet grid

Let γ and Γ be a pair of nested ellipses and P a Poncelet n-gon circumscribing γ and inscribed into Γ. Applying a projective transformation, we assume that γ and Γ are confocal.

Let x be the parameter on γ introduced in Section 3. Choosing the origin appropriately, the tangency points of the sides of P with γ have coordinates

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

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 a confocal ellipse to 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 next result, which is an (apparently new) addition to Theorem 3, the string construction.

Theorem 7 *Apply the string construction to an oval γ and let p, p' be two points on the curve Γ, see figure 9. Then, in the limit $p' \to p$, the lines pp' and qq' become orthogonal.*

Corollary 8 *The locus of intersections of tangent lines to an ellipse γ at $\gamma(c-x)$ and $\gamma(c+x)$ is a confocal hyperbola. In particular, the set Q_k lies on a confocal hyperbola.*

Proof of the corollary. Since points p and p' in figure 9 lie on a confocal ellipse, the "lengths" (measured via the parameter introduced in Section 3) of the arcs aa' and bb' are equal. It follows from Theorem 7 that the locus in question is a curve orthogonal to the family of confocal ellipses, that is, a confocal hyperbola. Therefore the set Q_k lies on a confocal hyperbola. $\Box$

Proof of Theorem 7. We will give two arguments, geometrical and analytical.

Let p and p' be infinitesimally close. By Theorem 3, the arc pp' bisects the angles qpq' and $qp'q'$. Let ε be the distance between p and p'. Dilate with

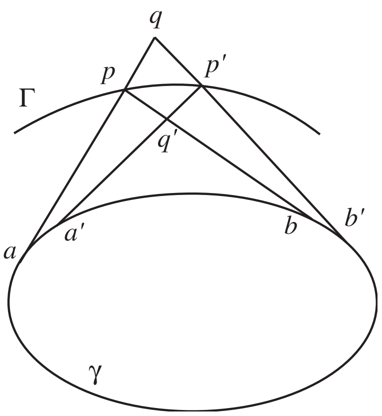

Figure 9: Addition to the string construction

factor $1/\varepsilon$. The angles do not change, the arc pp' becomes straight (to order ε), and, in the limit $\varepsilon \to 0$, we obtain a quadrilateral $pqp'q'$, symmetric with respect to the diagonal pp'. Hence $pp' \perp qq'$.

Analytically, let us see how fast the points a and b move as one moves the point p (not necessarily confined to Γ). 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 points a and b. Denote the velocity vector of point p by w.

Then $k_1 = \omega_1/v_1$ and $\omega_1 = w_1/l_1$ where w_1 is the component of w perpendicular to ap. Likewise, for the variables with index 2. It follows that

$$\frac{v_2}{v_1} = \frac{l_1 k_1}{l_2 k_2} \cdot \frac{w_2}{w_1}.$$

Consider two choices of w: tangent to Γ and perpendicular to it. Because of the equal angles property, Theorem 3, in the first case we have $w_1 = w_2$, and in the second case, $w_1 = -w_2$. Thus the ratio v_2/v_1 in both cases will have the same value and opposite signs. This is equivalent to orthogonality of xp' and qq'. $\square$

5 Elliptic coordinates and linear equivalence of the sets Ps and of the sets Qs

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, let x be the parameter on it described in Section 3. Note that the map $x \mapsto x + 1/2$ is central symmetry of the ellipse; in particular, the tangent lines at points $\gamma(x)$ and $\gamma(x + 1/2)$ are parallel.

For a point P outside of γ, draw tangent segments PA and PB to the ellipse, and let $x - y$ and $x + y$ be the coordinates of the points A and B, where $0 \le y < 1/4$. Then (x, y) are coordinates of the point P. We proved in Section 4 that the coordinate curves $y = const$ and $x = const$ are ellipses and hyperbolas, 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

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

Fixing the second coordinate yields an angular set P and fixing the first one yields a radial set Q.

An ellipse

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

also determines *elliptic coordinates* in the plane. Through a point P there passes a unique ellipse and a unique hyperbola from the confocal family of conics (3). The elliptic coordinates of P are the respective values of the parameter, λ_1 and λ_2 in (3). The hyperbolas and ellipses from the confocal family (3) are the coordinate curves of this coordinate system, $\lambda_1 = const$ and $\lambda_2 = const$, respectively. Cartesian coordinates of point 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 an ellipse: $(x_1, x_2) \mapsto (\pm x_1, \pm x_2)$).

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

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

This map takes Γ_λ to Γ_μ. The following lemma is classical and goes back to J. Ivory.[5]

Lemma 5.1 *If Γ_λ and Γ_μ are two ellipses (respectively, two hyperbolas) and P is a point of Γ_λ then the points P and $Q = A_{\lambda,\mu}(P)$ lie on the same confocal hyperbola (resp., ellipse).*

Proof. Unfortunately, we do not know a geometrical proof, so our argument will be computational.

We will consider the case when Γ_λ and Γ_μ are ellipses. 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 $\lambda_1 = \mu_1$.

Let (x_1, x_2) and (X_1, X_2) be the Cartesian coordinates of points P and Q. One has formulas (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, $Q = A_{\lambda,\mu}(P)$, hence

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

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 $\pm A_{\lambda,\mu}$, depending on whether $k - m$ is even or odd. The argument for the sets Q_k is similar.

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

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

$$\left(\frac{k}{2n} + \frac{j}{n}, \frac{k}{2n}\right) \quad \text{and} \quad \left(\frac{m}{2n} + \frac{j}{n}, \frac{m}{2n}\right); \; j = 0, 1, \dots, n-1.$$

The sets P_k and P_m lie on confocal ellipses Γ_λ and Γ_μ. According to Lemma 5.1, the map $A_{\lambda,\mu}$ preserves the first elliptic coordinate, and hence the x-coordinate. Therefore the coordinates of the points of the set $A_{\lambda,\mu}(P_k)$ are

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

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

Acknowledgments. Many thanks to R. Schwartz for fruitful discussions of his beautiful theorem and to the referees for their helpful suggestions. We are grateful to the Mathematics Institute at Oberwolfach for its hospitality and to NSF for partial support. The second author is also grateful to MPPIM, Bonn for its invariable hospitality.

References

[1] V. Arnold. *Mathematical methods of classical mechanics.* Springer-Verlag, 1989.

[2] W. Barth, Th. Bauer. *Poncelet theorems.* Expos. Math. **14** (1996), 125–144.

[3] M. Berger. *Geometry.* Springer-Verlag, 1987.

[4] H. Bos, C. Kers, F. Oort, D. Raven. *Poncelet's closure theorem.* Expos. Math. **5** (1987), 289–364.

[5] S.-J. Chang, R. Friedberg. *Elliptical billiards and Poncelet's theorem.* J. Math. Phys. **29** (1988), 1537–1550.

[6] D. Elliott. *M. L. Urquhart.* J. Aust. Math. Soc. **8** (1968), 129–133.

[7] P. Griffith, J. Harris. *A Poncelet theorem in space.* Comm. Math. Helv. **52** (1977), 145–160.

[8] V. Gutenmacher, N. Vasilyev. *Lines and curves.* Birkhauser, 2004.

[9] A. Katok, B. Hasselblatt. *Introduction to the modern theory of dynamical systems.* Camb. Univ. Press, 1995.

[10] J. King. *Three problems in search of a measure.* Amer. Math. Monthly **101** (1994), 609–628.

[11] V. Kozlov, D. Treshchev. *Billiards. A genetic introduction to the dynamics of systems with impacts.* Amer. Math. Soc., Providence, RI, 1991.

[12] B. Mirman. *Numerical ranges and Poncelet curves.* Linear Algebra Appl. **281** (1998), 59–85.

[13] B. Mirman. *Sufficient conditions for Poncelet polygons not to close.* Amer. Math. Monthly **112** (2005), 351–356.

[14] J. Moser. *Geometry of quadrics and spectral theory.* Chern Symp., Springer-Verlag, 1980, pp. 147–188.

[15] D. Pedoe. *The most "elementary" theorem of Euclidean geometry.* Math. Mag. **49** (1976), 40–42.

[16] K. Poorrezaei. *Two proofs of Graves's theorem.* Amer. Math. Monthly **110** (2003), 826–830.

[17] L. Santalo. *Integral geometry and geometric probability.* Addison-Wesley, 1976.

[18] R. Schwartz. *The Poncelet grid.* Preprint.

[19] Yu. Suris. *The problem of integrable discretization: Hamiltonian approach.* Birkhauser, 2003.

[20] S. Tabachnikov. *Poncelet's theorem and dual billiards.* L'Enseign. Math. **39** (1993), 189–194.

[21] S. Tabachnikov. *Billiards*, Soc. Math. France, "Panoramas et Syntheses", No. 1, 1995.

[22] S. Tabachnikov. *Geometry and billiards*, Amer. Math. Soc., Providence, 2005.

[23] A. Veselov. *Integrable mappings.* Russ. Math. Surv. **46** No. 5 (1991), 1–51.