The Desargues-Hilbert Billiard

Kevin Casto

April 29, 2013

Given two distinct conics C_1 and C_2 in $\mathbb{R}P^2$, defined by equations $p_1(x,y,z) = 0$ and $p_2(x,y,z) = 0$, so that p_1 and p_2 are homogeneous quadratic polynomials, we can form their pencil, which is the set of all their linear combinations; i.e., the set of all solutions to $\lambda p_1 + \mu p_2 = 0$ for each $\lambda, \mu \in \mathbb{R}$. It is clear that scaling (λ, μ) leads to the same solution set, so that the elements of the pencil are parameterized by points on the projective line with homogeneous coordinates $[\lambda : \mu]$. In fact, if we write the conics in standard form as

$$p_1(x,y,z) = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} a_1 & b_1/2 & c_1/2 \\ b_1/2 & d_1 & e_1/2 \\ c_1/2 & e_1/2 & p_1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = a_1x^2 + b_1xy + c_1xz + d_1y^2 + e_1yz + f_1z^2 = 0$$

$$p_2(x,y,z) = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} a_2 & b_2/2 & c_2/2 \\ b_2/2 & d_2 & e_2/2 \\ c_2/2 & e_2/2 & p_2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = a_2x^2 + b_2xy + c_2xz + d_2y^2 + e_2yz + f_2z^2 = 0$$

then, again, scaling either gives the same solution set, so that we can consider each conic as a point $[a : b : c : d : e : f]$ in five-dimensional projective space. Thus a pencil of conics is just a (projective) line in this space. Let us abuse notation slightly and denote the pencil $\overline{C_1C_2}$

There is a theorem of Desargues that says the following: given two conics C_0 and C_1, consider a tangent line ℓ to C_0. ℓ intersects each conic in $\overline{C_0C_1}$ at two points (counting multiplicity); thus the pencil induces a map on ℓ, which we shall denote $\mathrm{Des}_\ell(\overline{C_0C_1})$, sending one point of intersection to the other (so that a point of tangency is fixed). Then the theorem says that this map is a projective transformation of ℓ, in particular a projective involution.

Clearly for any other conic C_2 in $\overline{C_0C_1}$, we have $\overline{C_0C_2} = \overline{C_0C_1}$, so that $\mathrm{Des}_\ell(\overline{C_0C_2}) = \mathrm{Des}_\ell(\overline{C_0C_1})$ for each tangent ℓ to C_0. Our first theorem investigates a converse to this:

Theorem (1)**.** *Given three distinct nondegenerate conics C_0, C_1 and C_2, if for seven distinct tangent lines $\ell_1, \dots, \ell_7$ to C_0, we have $\mathrm{Des}_{\ell_i}(\overline{C_0C_1}) = \mathrm{Des}_{\ell_i}(\overline{C_0C_2})$, then the three conics lie in a pencil. Furthermore, this bound is sharp.*

Letting the conics be the zero-sets of homogeneous quadratics $p_0(x,y,z)$, $p_1(x,y,z)$, and $p_2(x,y,z)$, we know that there is a projective transformation sending C_0 to the unit circle in the standard affine plane, so that WLOG we may assume $p_0(x,y,z) = x^2 + y^2 - z^2$. For a given tangent line ℓ to C_0, if we pick two points P and Q on ℓ, we may parameterize ℓ by $(x_1 : x_2) \mapsto x_1P + x_2Q$. One natural point is the point of tangency with C_0; and since C_0 is the unit circle, we know that the "line at infinity" $z = 0$ does not intersect C_0 over $\mathbb{R}$, so the intersection point of ℓ with this line will always give us another distinct point. We let P be this point at infinity and Q be the point of tangency, so that the point of tangency is given by $(0 : 1)$.

Now, we have a projective involution $\mathrm{Des}_\ell(\overline{C_0C_1}) = \mathrm{Des}_\ell(\overline{C_0C_2})$ of ℓ, which we can write under the projective parameterization as just a 2×2 matrix A, considered up to scalar multiples.

Claim. *Under this parameterization of ℓ, we have $A \equiv \begin{pmatrix} 1 & 0 \\ c & -1 \end{pmatrix}$, which in standard affine coordinates where $x_2 \neq 0$ is the linear-fractional transformation $t \mapsto \dfrac{t}{ct-1}$*

Proof. Since Q, given by $(0:1)$, is a fixed point of the transformation, we must have

$$A\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}b\\d\end{pmatrix} = \begin{pmatrix}M\\0\end{pmatrix}$$

for some constant M, so that we must have $b = 0$. Furthermore, since we have an involution, A^2 must act as the identity, so that $A^2 = \lambda I$ for a scalar λ. Thus

$$A^2 = \begin{pmatrix}a^2 & 0\\ ac+cd & d^2\end{pmatrix} = \begin{pmatrix}\lambda & 0\\ 0 & \lambda\end{pmatrix} \implies d = \pm a$$

If $d = a$, then since $ac + cd = 2ab = 0$, either $a = 0$ and A is the zero map, which is impossible, or $c = 0$ and A the identity, which is impossible since the map interchanges two distinct points. Thus we must have $d = -a$, and so since a is nonzero, we can write A as $\begin{pmatrix}1 & 0\\ c & -1\end{pmatrix}$. □

We notice that on the affine line, for $t \neq 0$ we have that

$$t^{-1} + \left(\frac{t}{ct-1}\right)^{-1} = c$$

We can extend this result to all of ℓ with the following "reciprocal addition" law:

Definition Let $\oplus : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ be given by $(a,b)\oplus(c,d) = (ac, bc+ad)$. Then $\oplus$ is invariant under scalar multiplication, and so descends to a map $\mathbb{R}P^2 \times \mathbb{R}P^2 \to \mathbb{R}P^2$.

For any point $(x_1 : x_2)$ in ℓ, we then have

$$\begin{pmatrix}x_1\\x_2\end{pmatrix} \oplus A\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}x_1\\x_2\end{pmatrix} \oplus \begin{pmatrix}x_1\\ cx_1 - x_2\end{pmatrix} = \begin{pmatrix}x_1^2\\ x_1x_2 + cx_1^2 - x_1x_2\end{pmatrix} = \begin{pmatrix}1\\c\end{pmatrix}$$

as long as $x_1 \neq 0$, i.e., as long as $(x_1 : x_2) \neq (0:1) = Q$. This allows us to establish the equivalence of the two projective involutions in a simple way:

Claim. $\mathrm{Des}_\ell(\overline{C_0C_1}) = \mathrm{Des}_\ell(\overline{C_0C_2})$ *iff for every tangent line ℓ to C_0, we have $P_1 \oplus P_1' = P_2 \oplus P_2'$, where P_1 and P_1' are the points of intersection of ℓ with C_1, and P_2 and P_2' with C_2.*

Proof. One direction is trivial. For the other direction, we note that for a given tangent line ℓ, the two involutions A_1 and A_2 are uniquely determined by their values c_1 and c_2. Thus if $P_1 \oplus P_1' = c_1 = P_2 \oplus P_2' = c_2$, we must have $A_1 = A_2$. If this is true for every tangent line, the two transformations are necessarily identical. □

So if we can solve algebraically for the generic points of intersection of ℓ with C_1 and C_2, we can express the fact that $\mathrm{Des}_\ell(\overline{C_0C_1}) = \mathrm{Des}_\ell(\overline{C_0C_2})$ in a purely algebraic way. We do this as follows: we have a standard isomorphism $\mathbb{R}P^1 \to C_0$ given by $(x_1 : x_2) \mapsto (x_1^2 - x_2^2 : 2x_1x_2 : x_1^2 + x_2^2)$. If ℓ, given in dual coordinates by $[a:b:c]$, intersects C_0 at a point parameterized by $(x_1 : x_2)$, we must have

$$a(x_1^2 - x_2^2) + b(2x_1x_2) + c(x_1 + x^2) = (c+a)x_1^2 + (2b)x_1x_2 + (c-a)x_2^2 = 0$$

For ℓ to be tangent, the discrimant $b^2 - (c+a)(c-a) = a^2 + b^2 - c^2$ must vanish, so that lines tangent to C_0 lie on the "dual circle" $a^2 + b^2 - c^2 = 0$. So we likewise have an isomorphism from $\mathbb{R}P^1$ to the space of lines tangent to C_0; here it makes sense to take modified form $(\lambda : \mu) \mapsto [-\lambda^2 + \mu^2 : -2\lambda\mu : \lambda^2 + \mu^2]$, so that we have

$$(\lambda^2 - \mu^2)(x_1^2 - x_2^2) + (-2\lambda\mu)(2x_1x_2) + (\lambda^2 + \mu^2)(x_1^2 + x_2^2) = 2(\mu x_1 - \lambda x_2)^2 = 0$$

and the point of tangency Q is $(x_1 : x_2) = (\lambda : \mu)$, and is given in planar coordinates by $(\lambda^2 - \mu^2 : 2\lambda\mu : \lambda^2 + \mu^2)$. That is, the two maps simultaneously parameterize a point on C_0 and the line tangent to it at that point.

To then find P, the point of intersection of ℓ with the line $z = 0$, we note that the line $ax+by+cz = 0$ and $z = 0$ will intersect at the point $(-b : a : 0)$, so we have $P = (2\lambda\mu : -\lambda^2 + \mu^2 : 0)$. So our parameterization of ℓ is given by

$$(x_1 : x_2) \mapsto x_1Q + x_2P = (2\lambda\mu x_1 + (\lambda^2 - \mu^2)x_2 : (\mu^2 - \lambda^2)x_1 + 2\lambda\mu x_2 : (\lambda^2 + \mu^2)x_2)$$

Note that this is not well-defined over $\mathbb{C}$: at the tangent line parameterized by $(\lambda : \pm i\lambda)$, we have $P = Q = (1 : \pm i : 0)$, one of the so-called circular points at infinity, and the map on tuples is given by $(x_1 : x_2) \mapsto 2\lambda^2(x_2 \pm ix_1)(\pm 1 : i : 0)$, so that $(\pm i : 1) \mapsto (0 : 0 : 0)$.

We can now find the points of intersection of ℓ with C_1 and C_2 by plugging the parameterized equation of ℓ into the conics: if a conic C has equation $ax^2 + bxy + cxz + dy^2 + eyz + fz^2 = 0$, then the points of intersection are the roots of

$$\begin{gathered}(d\lambda^4 - 2b\lambda^3\mu + (4a - 2d)\lambda^2\mu^2 + 2b\lambda\mu^3 + d\mu^4)x_1^2 + \\ (-(b + e)\lambda^4 + (4a + 2c - 4d)\lambda^3\mu + 6b\lambda^2\mu^2 + (-4a + 2c + 4d)\lambda\mu^3 + (e - b)\mu^4)x_1x_2 + \\ ((a + c + f)\lambda^4 + 2(b + e)\lambda^3\mu + (-2a + 4d + 2f)\lambda^2\mu^2 + 2(e - b)\lambda\mu^3 + (a - c + f)\mu^4)x_2^2 = 0\end{gathered}$$

If we write this as $\alpha x_1^2 + \beta x_1x_2 + \gamma x_2^2 = 0$, then its roots are $r_1 = (2\gamma : -\beta + \sqrt{\beta^2 - 4\alpha\gamma})$ and $r_2 = (2\gamma : -\beta - \sqrt{\beta^2 - 4\alpha\gamma})$, and we have that $r_1 \oplus r_2 = (\gamma : -\beta)$. But as we proved in the lemma, this quantity is the same for both C_1 and C_2, so that $(\gamma_1 : -\beta_1) \equiv (\gamma_2 : -\beta_2)$, and therefore $\gamma_1\beta_2 - \gamma_2\beta_1 = 0$. Expanding this out, we get a degree eight homogeneous polynomial $P(\lambda, \mu)$ in λ and μ.

Claim. *$P(\lambda, \mu)$ is divisible by $\lambda^2 + \mu^2$, i.e., $P(\lambda, \pm i\lambda) = 0$*

Proof. As was mentioned, if we try to extend our notions to $\mathbb{C}P^2$ and look at the tangent line parameterized by $(\lambda : \pm i\lambda)$, the map $\mathbb{R}P^1 \to \ell$ degenerates to $(x_1 : x_2) \mapsto 2\lambda^2(x_2 \pm ix_1)(\pm 1 : i : 0)$, which is not even well-defined at $(\pm i : 1)$. Since we have only been working over $\mathbb{R}$, this does not bother us, but it does mean that if we formally carry out our procedure, well-defined over $\mathbb{R}$, of plugging in the parameterized equation for ℓ into the conic equation, we get as our homogeneous polynomial in $(x_1 : x_2)$ simply $4\lambda^4(a+bi-d)(x_2+ix_1)^2$. Thus γ_k and β_k are just $(a_k + b_k i - d_k)$ times a term without any of the conic coefficients, so that $\gamma_1\beta_2 - \gamma_2\beta_1 = 0$, and $P(\lambda, \pm i\lambda) = 0$. □

Dividing P through by $\lambda^2 + \mu^2$, we get a degree six homogeneous polynomial $Q(\lambda, \mu)$. Since $Q(\lambda, \mu) = (\gamma_1\beta_2 - \gamma_2\beta_1)/(\lambda^2 + \mu^2)$, all the coefficients consist of terms of the form $v_1w_2 - w_2v_1 = (v_1, v_2) \times (w_1, w_2)$, the two-dimensional scalar cross product. This motivates us to define $\mathbf{a} = (a_1, a_2), \mathbf{b} = (b_1, b_2), \dots, \mathbf{f} = (f_1, f_2)$. Having done this, the coefficients are further simplified by making the substition $\mathbf{A} = \mathbf{a} + \mathbf{f}$, $\mathbf{D} = \mathbf{d} + \mathbf{f}$, which eliminates $\mathbf{f}$. We can then write

$$\begin{gathered}Q(\lambda, \mu) = (\mathbf{b} + \mathbf{e}) \times (\mathbf{A} + \mathbf{c})\,\lambda^6 + \\ 2((2\mathbf{D} - \mathbf{A}) \times \mathbf{c} - 2\mathbf{A} \times \mathbf{D})\,\lambda^5\mu + \\ ((\mathbf{A} + 3\mathbf{c} + 4\mathbf{D}) \times b + (-5\mathbf{A} - 3\mathbf{c} + 4\mathbf{D}) \times \mathbf{e})\,\lambda^4\mu^2 + \\ 4(\mathbf{A} \times \mathbf{c} + 2\mathbf{e} \times \mathbf{b})\,\lambda^3\mu^3 + \\ ((\mathbf{A} - 3\mathbf{c} + 4\mathbf{D}) \times \mathbf{b} + (5\mathbf{A} - 3\mathbf{c} - 4\mathbf{D}) \times \mathbf{e})\,\lambda^2\mu^4 + \\ 2((2\mathbf{D} - \mathbf{A}) \times \mathbf{c} + 2\mathbf{A} \times \mathbf{D})\,\lambda\mu^5 + \\ (\mathbf{b} - \mathbf{e}) \times (\mathbf{A} - \mathbf{c})\,\mu^6\end{gathered}$$

If Q is nonzero, it has at most six real roots in $\mathbb{R}P^1$; but it vanishes for each $(\lambda : \mu)$ parameterizing a point on the unit circle with a projective tangent line; thus if there are seven distinct such tangent lines, it must be identically zero, so that all of its coefficients vanish. We thus get a system of seven equations. By adding and subtracting opposite pairs of these equations (first and last, second and second-to-last, etc.), and combining pairs of equations with one side equal, these can be reduced to the following:

$$\mathbf{A} \times \mathbf{D} = 0$$
$$(2\mathbf{D} - \mathbf{A}) \times \mathbf{c} = 0$$
$$(\mathbf{A} + \mathbf{D}) \times \mathbf{b} = 0$$
$$(-2\mathbf{A} + \mathbf{D}) \times \mathbf{e} = 0$$

$\mathbf{A} \times \mathbf{D} = 0$ tells us that $\mathbf{A}$ and $\mathbf{D}$ are parallel. They can thus be written as scalar multiples of the same vector $\mathbf{v}$; so say $\mathbf{A} = A\mathbf{v}$ and $\mathbf{D} = D\mathbf{v}$. But then $2\mathbf{D} - \mathbf{A}$, $\mathbf{A} + \mathbf{D}$ and $-2\mathbf{A} + \mathbf{D}$ are also parallel to $\mathbf{v}$, meaning that, by the other three equations, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{e}$ are as well, so that we can also write $\mathbf{b} = b\mathbf{v}$, $\mathbf{c} = c\mathbf{v}$, and $\mathbf{e} = e\mathbf{v}$. But then, if $\mathbf{v} = (v_1, v_2)$, we have

$$\begin{pmatrix} A_1 \\ b_1 \\ c_1 \\ D_1 \\ e_1 \end{pmatrix} = v_1 \begin{pmatrix} A \\ b \\ c \\ D \\ e \end{pmatrix}; \quad \begin{pmatrix} A_2 \\ b_2 \\ c_2 \\ D_2 \\ e_2 \end{pmatrix} = v_2 \begin{pmatrix} A \\ b \\ c \\ D \\ e \end{pmatrix}$$

In our original coordinates, we have $\mathbf{a} = \mathbf{A} - \mathbf{f}$ and $\mathbf{d} = \mathbf{D} - \mathbf{f}$, and thus

$$C_1 = \begin{pmatrix} a_1 \\ b_1 \\ c_1 \\ d_1 \\ e_1 \\ f_1 \end{pmatrix} = \begin{pmatrix} v_1 A - f_1 \\ v_1 b \\ v_1 c \\ v_1 D - f_1 \\ v_1 e \\ f_1 \end{pmatrix}; \; C_2 = \begin{pmatrix} a_2 \\ b_2 \\ c_2 \\ d_2 \\ e_2 \\ f_2 \end{pmatrix} = \begin{pmatrix} v_2 A - f_2 \\ v_2 b \\ v_2 c \\ v_2 D - f_2 \\ v_2 e \\ f_2 \end{pmatrix}$$

Furthermore, we know that $v_1 \neq 0$, $v_2 \neq 0$, because otherwise we would just have $C_1 = C_0$ or $C_2 = C_0$ (recall that C_0 is parameterized by $(1 : 0 : 0 : 1 : 0 : -1)$). But then it is clear that

$$v_2 C_1 - v_1 C_2 + (v_2 f_1 - v_1 f_2) C_0 = 0$$

so that the three conics are collinear in projective space, i.e., lie in a pencil.

Finally, to see that the bound is sharp, consider the conics C_0 the unit circle and C_i given by the equation $a_i x^2 + (3a_i + 2f_i)y^2 + e_i yz + f_i z^2 = 0$ for $i = 1, 2$, so that $C_i = (a_i : 0 : 0 : 3a_i + 2f_i : e_i : f_i)$. Then one can confirm that, in the terminology above, $Q(\lambda, \mu) = \mathbf{e} \times (\mathbf{a} + \mathbf{f})(\lambda^6 - 7\lambda^4\mu^2 + 7\lambda^2\mu^4) = \mathbf{e} \times (\mathbf{a} + \mathbf{f})(\lambda - \mu)(\lambda + \mu)(\lambda - (\sqrt{2} - 1)\mu)(\lambda + (\sqrt{2} - 1)\mu)(\lambda - (1 + \sqrt{2})\mu)(\lambda + (1 + \sqrt{2})\mu)$. Since we can clearly choose $\mathbf{a}$, $\mathbf{e}$, and $\mathbf{f}$ to make $\mathbf{e} \times (\mathbf{a} + \mathbf{f})$ nonzero, the three are not in the same pencil, but for the six tangent lines to C_0 parameterized by the roots of Q, we have $\mathrm{Des}_\ell(\overline{C_0 C_1}) = \mathrm{Des}_\ell(\overline{C_0 C_2})$. As a concrete example, we may take the ellipses $4x^2 + 2y^2 - 5 = 0$ and $3x^2 + y^2 - 2y - 4 = 0$, both of which lie entirely outside C_0. Since they intersect one another but not C_0, it is clear that they cannot all lie in the same pencil, but again, we have $Q(\lambda, \mu) = 2(-\lambda^6 + 7\lambda^4\mu^2 - 7\lambda^2\mu^4 + \mu^6)$, so that it has the full set of distinct real roots listed above. □

While the rather delicate algebro-geometric questions restricted to pencils of conics are interesting, there is a rather different direction we can go in: namely, we can consider the Desargues map not just operating on individual lines at a time, but on the whole plane simultaneously. This opens the subject up to a dynamical systems perspective which proves quite fruitful.

So, given a (nondegenerate) conic C and a pencil $\mathcal{P}$ that it lies in, for any point p there are by duality two lines through p tangent to C; the lines are equal if $p \in C$, both complex if p is "inside" C (in the component of $\mathbb{R}P^2 \setminus C$ homeomorphic to a disc) and both real if is "outside" C (in the component of $\mathbb{R}P^2 \setminus C$ homeomorphic to a Möbius strip). Suppose we pick one of the two and call it R; then $\mathrm{Des}_R(\mathcal{P})$ takes p to another point p'. Clearly one of the tangent lines through p' will be R, so call the other R': we then continue by using $\mathrm{Des}_{R'}(\mathcal{P})$ to map p' to p'', and so on. Letting $p = p_0$, we get an infinite sequence of points $(p_0, p_1, p_2, \dots)$. Likewise we can use the other tangent line—call it L—to go "backwards", and thus get a doubly-infinite sequence of points $(\dots, p_{-2}, p_{-1}, p_0, p_1, p_2, \dots)$.

The names L and R are chosen because we are distinguishing left from right, from the point of view of p looking at C. Unfortunately, since $\mathbb{R}P^2$ is non-orientable, we cannot consistently do this on the entire plane. However, any affine neighborhood U of C is clearly orientable, so that if U^+ is the component of U "outside" of C (homeomorphic to an annulus), we get a well-defined map $\mathrm{Des}_U(\mathcal{P}) : U^+ \to U^+$. By definition, each conic in $\mathcal{P}$ (or its intersection with U^+) is invariant under this map. Our last theorem thus proves a very rigid kind of invariance of conics under this map: namely that if $\mathrm{Des}(\mathcal{P})$ takes any seven points of a conic into itself, then in fact the whole conic is invariant under the map.

This leads us to consider the invariance of other curves under the map. Are there other curves besides conics in $\mathcal{P}$ that are invariant? If there are, the invariance becomes less interesting: after all, conics in a pencil are invariant under the identity map, but that has nothing to do with their being conics. What we are really after is a minimal set of invariant curves, so that our dynamical system is integrable. Integrability is a famously vague and variously-defind notion, but here we mean that, at least in a neighborhood U^+ of C, we have a foliation of minimal invariant curves. All we have left is to establish minimality; but to answer that question, it's worth recognizing the map $\mathrm{Des}(\mathcal{P})$ in a different guise.

In fact, this map has been relatively well-studied: it is the so-called "dual billiards", or "outer billiards" map, in the hyperbolic plane, as discussed and proven in [BILL]. The idea of outer billiards is to take a convex shape in the plane, and then maps points outside of it to their reflection through tangent lines; this turns out to be spherically dual to the standard billiards map of points (or really, directed lines) bouncing around inside a convex shape. Euclidean outer billiards obviously uses the Euclidean metric to measure distance—the idea being that the point goes twice the distance to the point of tangency of one of its tangent lines to the shape—but we can just as well play this game in the hyperbolic plane. If we take an element of the pencil around C which is still "around" C, then we can map this outer ellipse to the unit circle, and use this as a Beltrami-Klein model of hyperbolic geometry, with Cayley's cross-ratio metric placed on it. Since $\mathrm{Des}(\mathcal{P})$ is defined projectively, it's not too hard to see that in fact it's exactly the map sending a point to its outer billiard reflection through the point of tangency.

While we will continue to make use of all the work that has already been done on outer billiards, the only thing we really want right now is the existence of an area form ω invariant under our map. This is the tool we need to prove our theorem:

Theorem (2)**.** *In sufficiently small affine neighborhoods U^+ of C, any smooth curve $\Gamma \subset U^+$ invariant under $\mathrm{Des}_U(\mathcal{P})$ is a conic and thus an element of $\mathcal{P}$.*

Proof. The "sufficiently small" clause is just to assure us that the pencil is still a foliation of nested ellipses, so that we still use the hyperbolic construction. (We will try to get beyond this to more general situations shortly.) For each conic D in the pencil, the area form ω gives us a length element dx on D which is also invariant under $\mathrm{Des}(\mathcal{P})$: namely, to measure a portion $L \subset D$, take an annular strip L_ϵ of height ϵ, and then take the limit of $\omega(L_\epsilon)/\epsilon$ as $\epsilon \to 0$. The invariance of dx then just follows from the invariance of ω by construction. This means that, from the point of view of dx, the map $\mathrm{Des}(\mathcal{P})|_D$ is just a rotation by a fixed amount, because otherwise dx wouldn't be invariant. Thus each element of the pencil has some rotation number θ_D. It is clear that $\theta_C = 0$ and, at least close to C as we are assuming, θ increases as we move away, since points outside C are not fixed, and θ varies continuously by the continuity of ω. Now, suppose $\Gamma \not\subset D$ for any $D \subset U^+$; then since the conics cover U^+, it must meet at least two, say C_1 and C_2, where WLOG

C_1 is the smaller one enclosed by C_2. But then, since the region between C_1 and C_2 is an annulus, Γ must intersect all the conics between them. And since θ varies continuously, this means it intersects at least one conic C_3 with irrational rotation number. If we call the point of intersection p_0, then since Γ is invariant, all the images of p_0 under $\mathrm{Des}(\mathcal{P})$ must lie in it, but likewise for C_3. And since $\mathrm{Des}(\mathcal{P})|_{C_3}$ is just given by $t \mapsto t + \theta \pmod 1$, with θ irrational, we get an infinite sequence $p_0, p_1, \ldots$ which gets arbitrarily close to every point of C_3. So since Γ is a smooth curve, we must have $\Gamma = C_3$. □

Thinking about it from the outer billiards perspective, we have proven that when playing outer billiards in the hyperbolic plane with a conic table, that any invariant curve must be a conic. We might also wonder whether (again, still in the hyperbolic plane) any continuous convex closed curve that we take as an outer billiards table, which has a conic as an invariant curve, must itself be a conic. To prove this, we use another result from [BILL], based on the existence of the invariant area form ω, which in fact is much stronger:

Theorem (3)**.** *Given a convex curve in the hyperbolic plane, there is a foliation of outer billiards tables inside the curve that have it as an invariant curve under the outer billiards map.*

The basic idea is an area construction, dual to the "string construction" used for an analagous theorem in standard billiards. If the outer billiards table C_0 has an invariant curve C_1, then for each tangent line ℓ to C_0, the region bounded by ℓ and C_1 has constant area under ω as ℓ varies around C_0. So if we are given C_1, we can reconstruct C_0 as the envelope of the family of lines cutting off a fixed amount of area out of C_1. Clearly if the area is 0, this is just C_1, and as we increase the area from this, we get a foliation of nested curves, which do not intersect one another.

This immediately gives us our desired statement:

Corollary. *If a billiard table C_0 in the hyperbolic plane has a conic C_1 as an invariant curve, C_0 must be a conic.*

Proof. This is simply because the foliation of billiard tables with C_1 as an invariant curve is just the pencil of conics generated by C_1 and the circle at infinity. In particular, there is some fixed area α carved out by the tangent lines to C_0; but we already know that the envelope of all lines carving out area α from C_1 is a conic, so that C_0 must be a conic. □