Bif (Andrew) Reiser
Henry Gorman
PSU REU

Hyperbolic Outer Billiards

9/18/2011

Introduction

For a convex subset P of a planar domain, we can define the outer billiards map T as follow- pick a point x outside P. There are two lines of support from x to P- make a consistent choice of which one to use (we will assume in the paper that we've chosen the one "on the right" from x's perspective), and then define $T(x)$ to be the reflection of x in the support point. Note that this is map is not well-defined if the support point is not unique- in this case, we call the point x a singular point and the map is not defined on this point.

This system was introduced by B. Neumann (CITE), and is sometimes referred to as the "dual billiards map". In any case, it has been shown that all polygons in the Euclidean plane admit periodic orbits (CITE), and that in the hyperbolic plane, orbits may (and often do!) escape to infinity (Dogru and Tabachnikov). We will now embark on a study of dual billiards in the hyperbolic plane.

Recall the hyperbolic plane is a surface of constant Gaussian curvature negative 1. This brings difficulties- no such thing as similarity.

Background

Common models that we will use in this paper include the upper half plane, Poincare disk, and Klein disc models of hyperbolic geometry. The upper half plane model is given by the open upper half plane in the complex plane, with the real axis and the point at infinity in the Complex plane acting as the points at infinity on the Hyperbolic plane. Geodesics in this space are circles and lines perpendicular to the real axis, and isometries of the half plane are given by Mobius transformations with coefficients in the reals. A very nice thing about such transformations is that they can be represented as elements of $PSL(2,\mathbb{R})$ with unit determinant, and we say that an element $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in PSL(2,\mathbb{R})$ acts on $z \in \mathbb{C}$ by $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) z \mapsto \frac{az+b}{cz+d}$. An extremely important fact is that the matrix of reflection about a point $p+iq$ is given by

$$\begin{pmatrix} p/q & \frac{-p^2-q^2}{q} \\ 1/q & -p/q \end{pmatrix}$$

The Poincare disk model is obtained by taking the upper halfplane conformally into the interior of the unit disc via the map $z \mapsto \frac{z-i}{z+i}$ and in this model geodesics are still arcs of circles or diameters intersecting the boundary orthogonally. The Klein model is then obtained by projecting the Poincare model up onto the surface of a unit sphere sitting at $(0,0,1)$ and then stereographicall projecting down and then scaling the image to the unit circle. The Klein model has the nice property that geodesics are straight lines, but unfortunately the transformation process is not conformal and thus the angular data is not preserved well.

Definitions and Tools

First, fix a convex polygon P in the hyperbolic plane for our outer billiard map T to act on (assume it contains the origin in the Poincare model- this is possible WLOG by just moving it there). Then, define the singularity set Δ of T to be the points where T^n for some $n \in \mathbb{Z}$ is undefined. Note this is the set of preimages and images of the extensions of the sides of the polygon. (It's nice to define Δ_i to be the i^{th} preimages and images of the extensions of the sides). It can be quite complex- SEE PICTURE. The complement of the singularity set is a union of convex regions- call it Σ. Define an invariant region to be a maximal connected region in Σ.

Next, we'll define the coding of a point in Σ. Label the vertices of the polgon P in counter-clockwise order $1, 2, \cdots, n$. Call the coding of a point the bi-infinte sequence of vertices it reflects off of. It is clear that any two points in the same invariant region have the same coding. (Strictly reducing this problem to one about coding is not very helpful- there are actually an infinite number of banned blocks, ie if you're given a coding as a bi-infinite sequence there's not a guarantee that you'll be able to determine whether it is allowed or not subject to the relations. We'll see a detailed explanation of this fact later.) Also note that we can associate a matrix of reflection to each in the halfplane model. For a point $re^{2\pi\lambda i}$ in the Poincare model, the matrix of reflection in the upper half plane model is

$$\begin{pmatrix} \frac{2r\sin(2\pi\lambda)}{1-r^2} & \frac{r^2+2r\cos(2\pi\lambda)+1}{1-r^2} \\ -\frac{r^2-2r\cos(2\pi\lambda)+1}{1-r^2} & -\frac{2r\sin(2\pi\lambda)}{1-r^2} \end{pmatrix}$$

From this association, we can determine that the matrix of the translation associated to a coding of an orbit by simply multiplying the matrices of each of the vertices together in the order specified by the coding. Why is this useful? Mostly because the isometries of the upper half plane model can be classified by properties of their matrix in $PSL(2, \mathbb{R})$. For a matrix with unit determinant, the (non-identity) isometry given by such a matrix is hyperbolic if $|tr| > 2$, parabolic if $|tr| = 2$, and elliptic if $|tr| < 2$. This corresponds to 2 fixed points on the boundary, 1 fixed point on the boundary, and 1 fixed point on the interior of the hyperbolic space, respectively.

Define the process of scaling a convex n-gon by a factor of k as follows: For a convex n-gon containing the origin in its interior (in the Poincare disk model) with vertices $\{r_1e^{2\pi\lambda_1}, \cdots, r_ne^{2\pi\lambda_n}\}$, let the new set of vertices be $\{kr_1e^{2\pi\lambda_1}, \cdots, kr_ne^{2\pi\lambda_n}\}$. Note that this is not similarity in hyperbolic space.

It's known that every $n-$gon has something called a rotation number which lies in the interval $[1/n, 1/2)$. If it has rotation number $1/n$, we call it large and all outer billiards orbits escape to infinity (see Dogru and Tabachnikov). If not all the orbits escape to infinity, interesting things seem to happen.

Define the rotation number for an n-periodic orbit on the interior of the hyperbolic plane as follows: consider the curve γ_p traced out by connecting our n-periodic point p with $T(p)$ and then with $T^2(p)$ and so on until we connect $T^{n-2}(p)$ to $T^{n-1}(p)$ to p. Define the rotation number to be the winding number of γ_p divided by n. (Note n must be finite, as a simple consequence of the orbit being periodic.)

Results

Well, here goes.

Lemma 1.1. If a bounded invariant region is mapped into itself, it has a fixed point.

Proof. If an invariant region is mapped into itself, it must be mapped entirely into itself. This is clear- if not, there would be some part of Δ between points in the image or preimage of the set. From this as well as the boundedness and the fact that it is mapped isometrically into itself, we can make use of the Brauer fixed point theorem to get that the region has a fixed point.

Lemma 1.2 Periodic points have open neighborhoods that travel with them for all time.

Proof. Let our point p have period n. Define the set $Im(p)$ to be all images and preimages of p under iterations of the billiards map. Then the set $Im(p)$ in this case is just $\{p, T(p), \cdots, T^{n-1}(p)\}$.

Since this set is closed and compact and the extensions of the sides of the triangles are closed, we can take the minimum distance between I and Δ_1 and we're guaranteed to get a positive number r. Then an open connected neighborhood of nonsingular points of radius r surrounds each of the points in $Im(p)$ for all time. (Why? Suppose not. Then we can get to the situation when the offending point of Δ is in Δ_1 by taking a suitable number of iterations of the map. But this is an isometric process and if this distance is smaller, then we've contradicted out method of choosing r).

Lemma 1.3 The only connected open sets σ of Σ which do not have periodic points escape to infinity under both the forwards and backwards billiards maps.

Proof. Suppose we have a region σ with no periodic points. By Lemma 1.1, this is equivalent to it not intersecting any of its images or preimages. But if it does not intersect its own images and preimages, we can construct a set of arbitrarly large measure by taking unions of it's images under either the forwards or backwards billiards map. But then these images cannot all lie in a bounded set, so it must escape to infinity.

Now observe that Σ can be partitioned into 5 subsets: periodic regions, unbounded regions, open regions which escape to infinity, line-like non-open regions, and points with no neighborhood. Ideally, we'd like to collapse this classification somehow.

Theorem 1. The rotation number of a periodic orbit inside the model is bounded above by the rotation number of the polygon on the boundary.

Proof. Consider the curve γ_p in the Klein model traced by connecting p to $T(p)$ to $T^2(p)$, etc, until connecting $T^{n-1}(p)$ back to p. This can be projected out onto the boundary of the circle by picking a point q inside our chosen polygon P such that no two points of $Im(p) = \{p, T(p), T^2(p), \cdots, T^{n-1}(p)\}$ are collinear with q and then drawing a ray from q to each of these such points. Define V_i to be the intersection of $\overrightarrow{qT^{i-1}(p)}$ with the unit circle.

Define an orientation-preserving homeomorphism of the circle Q in the following manner: if the arc from V_i to V_j in the positive direction contains no other V_k, then send the interval $[V_i, V_j] \in S^1$ to $[V_{i+1}, V_{j+1}]$ (with indicies taken modulo n) so that its lift into $\mathbb{R}$ is a linear map. This map clearly has the same rotation number as the orbit of p. On the other hand, for all lifts of T and Q, $\tilde{T} \geq \tilde{Q}$, so we have that the rotation number of the circle map must be at least that of Q, and thus the theorem follows. ■

Unfortunately, just because an orbit has rotation number inside the given range does not mean it must be achieved. For example, the 8 periodic orbit with skip size 3 (rotation number 3/8) is prohibited from existing for all suitably small equilateral triangles. It exists for the equilateral triangle with vertices $re^{2\pi ik/3}$ for $r = 0123456789$ and $k = 0, 1, 2$ in the Poincare disk model, but then the trace of the matrix for this transformation is larger than 2 for all equilateral triangles with vertices $re^{2\pi ik/3}$ with $0 < r < K \approx 0123456789$ in the Poincare disk model.

Lemma 1.4 The trace of the matrix of a transformation approaches 0 or 2 as the side lengths approach 0.

Proof. As the polygon approaches a single point, we have either an even or odd number of reflections through a point- this will be conjugate to either the identity or the rotation by π matrix, with trace 0 or 2 respectively. Since the entries of the matrix depend continuously on the polygon, the lemma is proven.

Lemma 1.5 For a periodic orbit in the hyperbolic plane, the rotation of the open region around it is given by $2\cos^{-1}(\frac{tr}{2})$ where tr is the trace of the sequence of matrices that give this transformation.

Proof. An elliptic transformation (ie has fixed point in hyperbolic plane) is conjugate to a rotation matrix of form $\begin{pmatrix} \cos(\theta/2) & \sin(\theta/2) \\ -\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}$, where it describes a rotation through $\theta/2$ about a point. But since trace is an invariant of conjugacy class, the lemma is obvious.

Lemma 1.6 All achieved periodic orbits inside the hyperbolic plane can be forced into an infinite number of rational and irrational rotations by suitable "scaling" of the polygon.

Proof. This is clear from the definition of our matrices and lemma 1.5- our polynomial expression for trace in a scaling factor of r is going to be continuous, and since the rotation is obtained continuously from this polynomial and the whole thing is a function of just one variable, apply the mean value theorem.

Theorem 2. Each realized orbit at infinity can be made to enter the hyperbolic plane.

Proof. Begin by taking a point p on the unit circle with a given coding. It reflects off some vertex, in the manner described in the drawing. See drawing. Fixed angle α and side length r.

By law of sines, $\sin(\pi - \alpha - x) = \frac{\sin(x)}{r}$, which simplifies to $\tan x = \frac{r \sin\alpha}{1 - r\cos\alpha}$. Take ArcTan and derive, see $\frac{dx}{dr} = \frac{\sin\alpha}{1 + r^2 - 2r\cos\alpha}$ which is obviously greater than 0 since $0 < \alpha < \pi$ and $1 + r^2 - 2r\cos\alpha > 1 + r^2 - 2r = (1-r)^2 > 0$. So $\frac{dx}{dr}$ is positive, and can be bounded from below by ϵ uniformly. The arc traveled in the positive direction from p to $T(p)$ is given by $\pi - 2x$, and so as r decreases, this arc lengthens.

Now, consider the case when p is the attractor (respectively repellor) of an n-periodic orbit on the boundary given by a hyperbolic isometry. Shrinking vertices uniformly will cause $T^n(p)$ to process in the positive direction past p. This implies that the attractor of the orbit has moved in the positive direction by a positive amount, and the repellor has also moved, but in the negative direction by a positive amount. This implies that they have gotten closer together. Note that this process occurs on a compact subset of S^1 and thus must eventually stop. However, if there is a positive distance between the attractor and repellor, then we can repeat this process. But this implies that the process only stops when the attractor and repellor merge to become the fixed point of a parabolic transformation on the boundary.

If the process is begun with a parabolic transformation, the orbit on the boundary is destroyed, as there are no points which map to "behind" themselves,

Shifting points of view to the Mobius transformation picture, what we've shown is that as the polygon is scaled down, the trace decreases uniformly, and the limiting value of $\frac{a}{2c}$ lies in the correctly-coded region. As the polygon continues to be scaled down, the two fixed points of the Mobius transformation, $\frac{a-d\pm\sqrt{tr^2-4}}{2c}$, one enters the upper half plane and in fact enters into the correctly coded region, by a continuity argument. Unfortunately, this continuity argument does not yet establish that such motion continues in a manner that preserves the coding. ■

Interesting Observations and Directions for Future Research

We have discovered a tiling of the hyperbolic plane by the outer billiards map. The triangle with vertices $re^{2\pi k/3}$ for $k = 0, 1, 2$ and $r \approx .275798$ (value that sets trace of matrix for 3-periodic orbit around equilateral triangle in Klein model equal to 1) in the Klein Disk model gives the triheptagonal tiling of the Hyperbolic plane, with many of the same properties that the tilings studied by Dogru and Tabachnikov (CITE). Notably:

1. The notion of rank is still well-defined and useful.
2. The tiles of a fixed rank are fixed by the outer billiards map. The proof is analogous to that of Dogru and Tabachnikov.
3. However, the tiles of a fixed rank are NOT necessarily permuted cyclically. Note that 3 distinct orbits of rotation number 2/5 exists as the tiles of rank 2.
4. The rotation number of the map on the circle at infinity is irrational, and equal to $\frac{1}{2} - \frac{1}{6\sqrt{5}}$. Proof of this fact is given below.

Proof of rotation number result.
First, note that the number of tiles of a given rank can be found recursively by the formulas $f(n) = 3f(n-2) - f(n-4)$ for odd n and $f(n) = 4f(n-1) - f(n-3)$ for even n. By the standard method of solving linear recurrences, we find that the number of tiles of rank n is $\frac{3\sqrt{5}-3}{2}\left(\frac{3+\sqrt{5}}{2}\right)^k - \frac{3\sqrt{5}+3}{2}\left(\frac{3-\sqrt{5}}{2}\right)^k$ for odd n with $k = \frac{n+1}{2}$, and $3\sqrt{5}\left((\frac{3+\sqrt{5}}{2})^{n/2} - (\frac{3-\sqrt{5}}{2})^{n/2}\right)$ for even n.

Now, we must deduce the rotation number from the given information. It is then clear that when the outer billiards map has been applied to a tile of rank n a total of $f(n)f(n+1)$ times, the tile of rank n and all neighboring tiles of rank $n+1$ will be back to their starting points. Then, the rotation number of the tile of rank $n+1$, denoted $r(n+1)$ will be $r(n) + \frac{\gcd(f(n), f(n+1))}{f(n)f(n+1)}$. It's easy to see, via an application of the Euclidean algorithm to the iterative relationships for $f(n)$, that this gcd is always 3. Therefore, the rotation number at infinity, ie $\lim_{n\to\infty} r(n)$ is equal to the sum of the series

$$\frac{1}{3} + 3\sum_{n=1}^{\infty} \frac{1}{f(n)f(n+1)}$$

But we know that $f(2k-1) + f(2k+1) = f(2k)$, so we can rewrite this series as

$$\frac{1}{3} + 3\sum_{k=1}^{\infty} \frac{1}{f(2k-1)f(2k+1)}$$

and then plugging in the formula for $f(2k+1)$, get

$$\frac{1}{3} + 3\sum_{k=1}^{\infty} \frac{1}{\left(\frac{3\sqrt{5}-3}{2}\left(\frac{3+\sqrt{5}}{2}\right)^k - \frac{3\sqrt{5}+3}{2}\left(\frac{3-\sqrt{5}}{2}\right)^k\right)\left(\frac{3\sqrt{5}-3}{2}\left(\frac{3+\sqrt{5}}{2}\right)^{k+1} - \frac{3\sqrt{5}+3}{2}\left(\frac{3-\sqrt{5}}{2}\right)^{k+1}\right)}$$

, or

$$\frac{1}{3} + \sum_{k=1}^{\infty} \frac{1}{3\left((\frac{3+\sqrt{5}}{2})^{2k} + (\frac{3-\sqrt{5}}{2})^{2k} - 3\right)}$$

By a standard trick, we can replace this with a q-Gamma function, and the given sum is $1/3 + \frac{1}{6\sqrt{5}\log(3-\sqrt{5})}\left(\Psi_{(\frac{3-\sqrt{5}}{2})^2}(1-\frac{\log((3-\sqrt{5})/2)}{2\log((3-\sqrt{5})/2)}) - \Psi_{(\frac{3-\sqrt{5}}{2})^2}(1-\frac{\log((3+\sqrt{5})/2)}{2\log((3-\sqrt{5})/2)})\right)$. But the quantity inside the parenthesis is simply $\Psi_{(\frac{3-\sqrt{5}}{2})^2}(\frac{1}{2}) - \Psi_{(\frac{3-\sqrt{5}}{2})^2}(\frac{3}{2})$, and the whole expression can be simplified to $\frac{1}{2} - \frac{1}{6\sqrt{5}}$. An interesting fact about this number is that it has continued fraction expansion

$$\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{1+\cfrac{1}{5+\cfrac{1}{1+\cfrac{1}{5\cdots}}}}}}$$

, which eventually falls into the same pattern as that demonstrated by Dogru and Tabachnikov, where each coefficient in the continued fraction is two less than the number of sides on one of the tiles ($3-2=1$ or $7-2=5$).

Other things we would like to know about tilings include which ones can we make? When do they occur? We have to simultaneously make a lot of polynomials of different degrees rational, so figuring out when this can happen would tell us a lot about how the trace polynomials of the different transformations were constructed.

This brings us to our next point of interest. We would love to figure out how exactly the entries of the matrix of the transformation associated to an orbit, and especially the trace polynomial of such matrix, depend on the coordinates of the vertices of the polygon in a simpler way. This would tell us a comparatively large amount about how the transformations behave, and how the fixed points of the transformations behave.

Our last direction we'd be interested in exploring and probably finishing up would be to characterize the behavior of orbits as they move further away from infinity. It's known via theorem 2 that all orbits that exist at infinity eventually move inside the hyperbolic plane. On the other hand, we don't know much about orbits away from infinity. Showing that all achieved orbits come from infinity would be nice, and telling exactly which orbits fall into the wrong regions via having their fixed points moved into fake regions would be really nice.