Rational Billiards

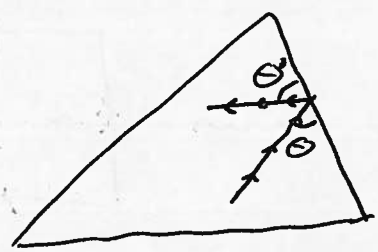
Bradley Zykoski, Student Dynamics Seminar, 9/16/19

Outline:

1. Introduction: Billiards & Unfolding
2. Behavior of generic trajectories
3. Illumination & the Magic Wand theorem

1. Introduction: Billiards & Unfolding

Consider a polygon P whose angles are rational multiples of π. We want to study the straight-line trajectories on P that bounce off the sides with an angle of reflection equal to the angle of incidence.



$\theta' = \theta$

We call such trajectories "billiards trajectories," because we can view this setup as an idealized version of a billiards table, where the trajectory is the path a billiard ball takes after it is hit.

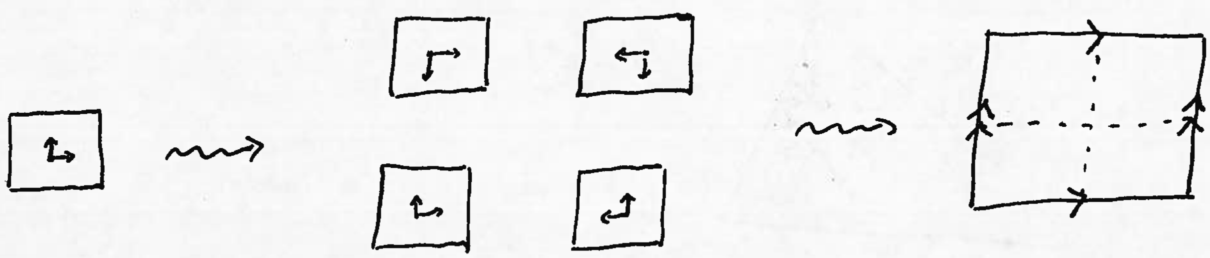
There is no canonical way of continuing such a trajectory after it meets a vertex of P, and so we make no such choice. Not all trajectories can be extended indefinitely. Continuing with our billiards analogy, we can think of the vertices of P as pockets of the billiards table.

This setup is pretty hard to work with, and so we pass to an equivalent construction that is easier to analyze: we unfold the billiards table P.

<u>Definition</u>: Let P be a polygon whose angles are rational multiples of π, and let G be the subgroup of $O(2)$ generated by the linear parts of the reflections through the the sides of P. Since the angles of P are rational multiples of π, the group G is finite.

Consider the standard framing $(e_1, e_2) \in T_x P \times T_x P$ at every interior point x of P. The <u>unfolding</u> of P is the surface obtained by taking the disjoint union $\coprod_{g \in G} g.P$ and identifying a side S of $g.P$ with a side S' of $h.P$ if gh^{-1} is the linear part of the reflection through S, the induced map $g.P \to h.P$ takes S to S', and this map carries the framing $(g.e_1, g.e_2)$ to the framing $(h.e_1, h.e_2)$.

<u>Example 1</u>: (Unfolding the square)

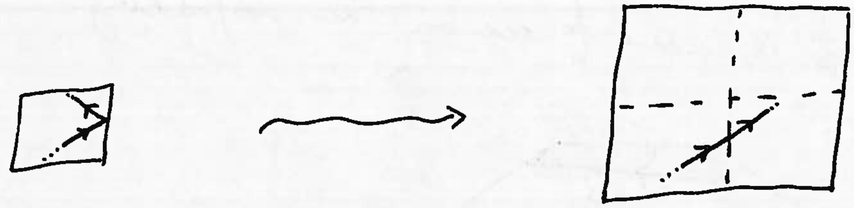


P $\qquad$ $\coprod_{g \in P} g.P$ $\qquad$ unfolding of P

The unfolding of the square P is the torus $T^2 = S^1 \times S^1$. One question about billiards that interests us is: How can we classify billiards trajectories in terms of their long-term behavior? In this example, a simple answer presents itself:

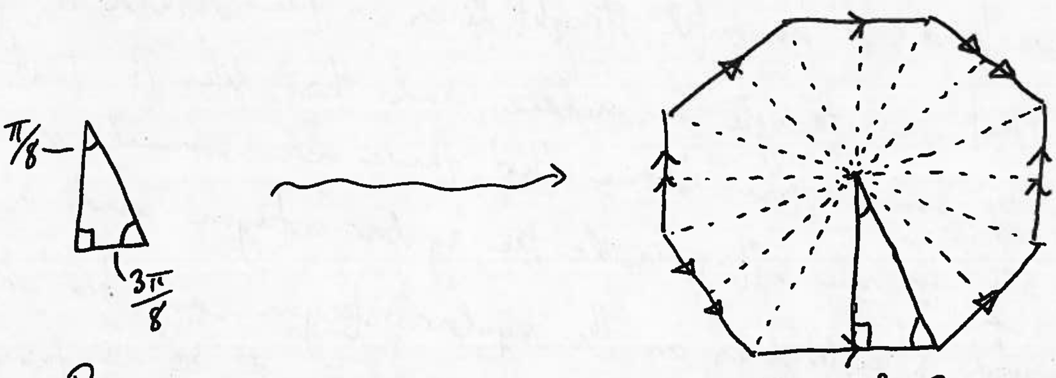
<u>Fact</u>: On the square, a trajectory that never hits a vertex is either periodic or dense (in the sense of topology).

This fact is far easier to see on the unfolding: A billiards trajectory on P corresponds to a straight line on T^2:



Since $T^2 = \mathbb{R}^2/\mathbb{Z}^2$, we see that straight lines of rational slope are periodic, while those of irrational slope are dense.

<u>Example 2</u>: (Unfolding the $\frac{\pi}{8}$-$\frac{3\pi}{8}$-$\frac{\pi}{2}$ triangle)



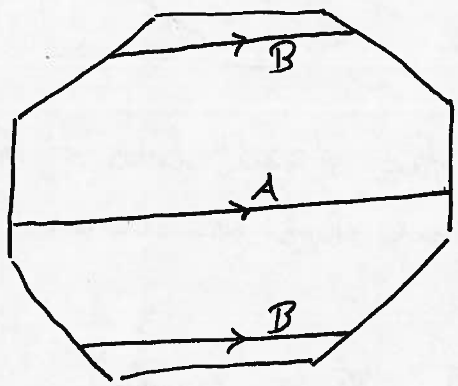
P unfolding of P

The unfolding of the $\frac{\pi}{8}$-$\frac{3\pi}{8}$-$\frac{\pi}{2}$ triangle is the regular octagon with opposite sides identified. One can easy verify (for example by computing the Euler characteristic) that this surface is the genus 2 oriented surface S_2.

It is true that the $\frac{\pi}{8}$-$\frac{3\pi}{8}$-$\frac{\pi}{2}$ triangle admits a classification of billiards trajectories similar to that of the square, but this is harder to prove. Let us make the following observation, though.

Fact: On the $\frac{\pi}{8}$ - $\frac{3\pi}{8}$ - $\frac{\pi}{2}$ triangle, every horizontal and vertical trajectory that does not hit a vertex is periodic.

Once again, this fact is easier to see on the unfolding:



There are two kinds of horizontal straight lines on the regular octagon: those, like A, that pass through the middle, and those, like B, that pass through the top and bottom. Notice that there are several copies of the $\frac{\pi}{8}$ - $\frac{3\pi}{8}$ - $\frac{\pi}{2}$ triangle sitting inside the regular octagon, and in particular, horizontal straight lines on the regular octagon can be obtained by unfolding horizontal and vertical trajectories on the $\frac{\pi}{8}$ - $\frac{3\pi}{8}$ - $\frac{\pi}{2}$ triangle.

Notice that in every case, unfolding a rational polygon P produces a topological surface, and this surface inherits a flat Riemannian metric by virtue of the fact that P sits in the Euclidean plane (though, the metric is undefined at finitely many points, e.g. the vertices of the regular octagon above). Notice also that this metric always has trivial holonomy (equivalently, transition functions between coordinate patches are Euclidean translations).

Definition: A translation surface is a closed oriented topological surface S, along with a flat Riemannian metric on $S \setminus \Sigma$ that has trivial holonomy, where $\Sigma \subseteq S$ is finite.

Consider a holomorphic 1-form ω on a closed Riemann surface X. Away from the zeroes of ω, the coordinate transition functions are of the form $z \mapsto z + C$, $C \in \mathbb{C}$. Therefore such a pair (X, ω) is equivalent to the above data of a translation surface, and it is not hard to see that we have an equivalence

$$\left\{ \begin{matrix} S \text{ a translation} \\ \text{surface} \end{matrix} \right\} \longleftrightarrow \left\{ \begin{matrix} X \text{ a Riemann surface} \\ \text{with } \omega \text{ a nonzero holo. 1-form} \end{matrix} \right\}$$

where the pointset Σ corresponds to the set of zeroes of ω. We therefore refer also to pairs (X, ω) as translation surfaces.

Final Note: In the literature, holomorphic 1-forms are often called abelian differentials.

---

2. Behavior of generic trajectories

We will from now on only work with translation surfaces; results about billiards can be obtained by applying the following results to translation surfaces obtained via unfolding.

In the previous examples, we remarked on the periodicity of certain billiards trajectories. On unfoldings, these corresponded to straight lines that closed back up on themselves. Notice that given any such straight closed loop, there will be some $\varepsilon > 0$ so that we can homotope this loop, through straight closed loops, all throughout the ε-neighborhood of this loop. This realizes the ε-neighborhood as a topological cylinder (i.e. annulus). Note that the only obstruction to making ε bigger is if such a homotopy would make our loop bump into one of the points on our translation surface at which the metric is undefined (we call these points <u>singularities</u>). When our loop bumps into singularities, they subdivide it into a union of straight arcs whose endpoints are singularities (we call such straight arcs <u>saddle connections</u>). Note however that a saddle connection need not arise via such a family of loops. However, we have two nice theorems about saddle connections and cylinders (by which we mean cylinders arising via straight closed loops):

<u>Theorem</u>: Any translation surface has countably many saddle connections.

<u>Proof</u>: This is straightforward: It is an immediate consequence of the fact that there are finitely many singularities and countably many homotopy classes of arcs between them. $\square$

(* Well-defined on (X,ω) by virtue of trivial holonomy)

<u>Theorem:</u> There is a countable dense set of angles* $\theta \in S^1$, depending on a translation surface (X,ω), such that (X,ω) has a straight line at angle θ that forms a closed loop, and hence also a cylinder.

<u>Proof</u>: This is hard! See Theorem 4.1 of [MT02]. $\square$

We now consider some deeper dynamical properties:

Definition: A flow is *minimal* if each of its orbits is dense.

Theorem: For all but countably many directions Θ, the straight-line flow on a translation surface is minimal.

Proof: Not super hard. See theorem 1.8 of [MT02]. $\square$

Definition: A flow is *uniquely ergodic* if it has only one invariant measure.

We now come to a cornerstone of the theory: *Masur's criterion for unique ergodicity*. Let $\mathcal{M}_g$ denote the moduli space of genus g Riemann surfaces, and let $\mathcal{H}$ denote the moduli space of genus g translation surfaces. Then there is an obvious forgetful map

$$\begin{array}{ccc} \mathcal{H} & & (X,\omega) \\ \downarrow & & \downarrow \\ \mathcal{M}_g & & X \end{array}$$

and in fact this map is proper with respect to the relevant topologies on $\mathcal{H}$ and $\mathcal{M}_g$, which we do not discuss here.

$\mathcal{H}$ is a dynamical system in its own right: Any translation surface can be drawn as some collection of polygons in the plane, with sides identified by Euclidean translations. $\mathcal{H}$ therefore admits a natural action by $GL_2^+(\mathbb{R})$:

e.g.

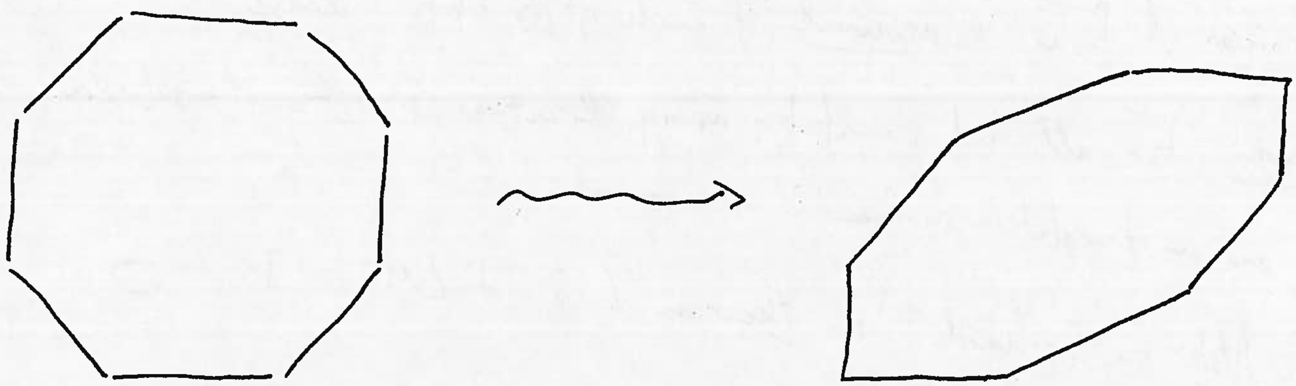


Fig. The action of $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in GL_2^+(\mathbb{R})$ on the regular-octagon translation surface

Let $\pi : \mathcal{H} \to \mathcal{M}_g$ denote the forgetful map.

<u>Theorem (Masur's criterion)</u>: Suppose the straight-line flow on a translation surface (X, ω) is minimal but not uniquely ergodic in the direction θ. Then, as $t \to \infty$, $\pi\left(\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}\begin{pmatrix} e^{-t} & 0 \\ 0 & e^{t} \end{pmatrix}\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \cdot (X, \omega)\right)$ leaves every compact subset of $\mathcal{M}_g$. (e.g.  in $\mathcal{M}_g$, $g=2$)

Note that since π is proper, this is stronger than the argument above leaving every compact subset of $\mathcal{H}$.

Furthermore, we have the following theorem:

<u>Theorem:</u> The set of directions θ for which the above divergence occurs has <u>Lebesgue measure</u> 0.

Therefore we conclude:

<u>Theorem</u>: The straight-line flow on a translation surface is minimal but not uniquely ergodic in almost no directions.

To summarize the results from this section, we give the following table:

On a translation surface, the set of directions $\theta \in [0, 2\pi]$ in which the straight-line flow ...	... is ...
forms a cylinder	countable & dense
is non-minimal	countable
is minimal, but not uniquely ergodic	of measure 0
is uniquely ergodic	of full measure

3. Illumination & the Magic Wand theorem

Suppose we replace the "billiards" metaphor with an equivalent one: instead of P being thought of as a billiards table and the straight-line trajectories as billiards trajectories, we think of P as a polygonal room whose sides are walls made of mirrors, and of the straight-line trajectories as rays of light.

<u>Question:</u> Is it true that in every polygonal room P, a candle placed at any point will illuminate every other point? If not, then to what extent does this property fail?

If P is a <u>convex</u> polygon, then obviously any point illuminates any other:

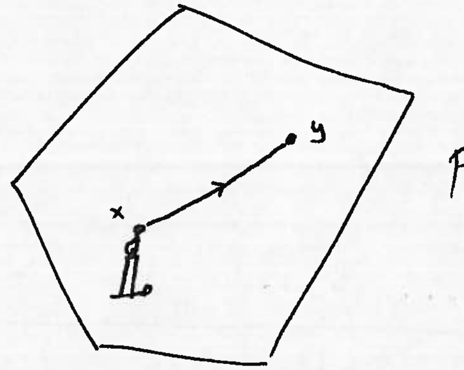


Fig. A candle at x illuminates y

In 1958, Roger Penrose gave an example of a non-polygonal room in which no point illuminates every other.

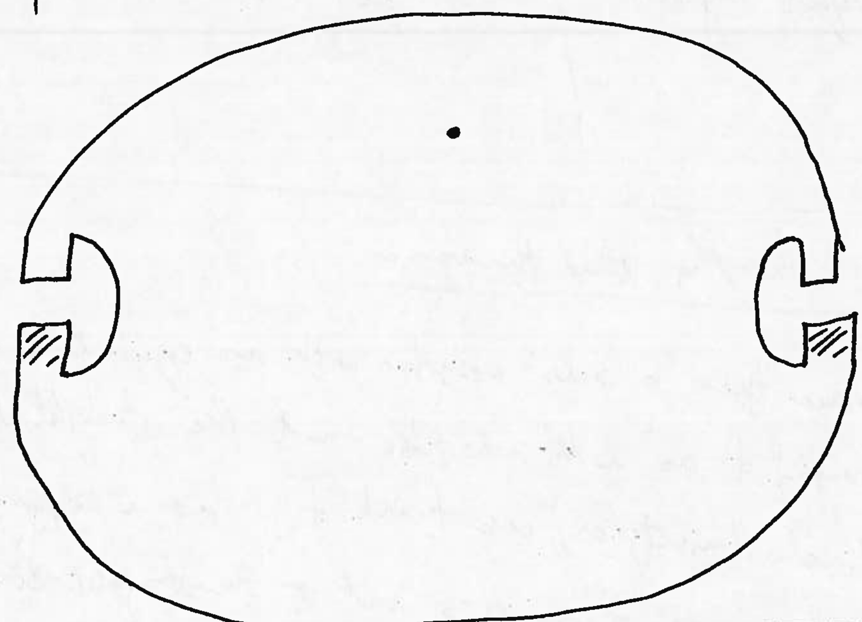
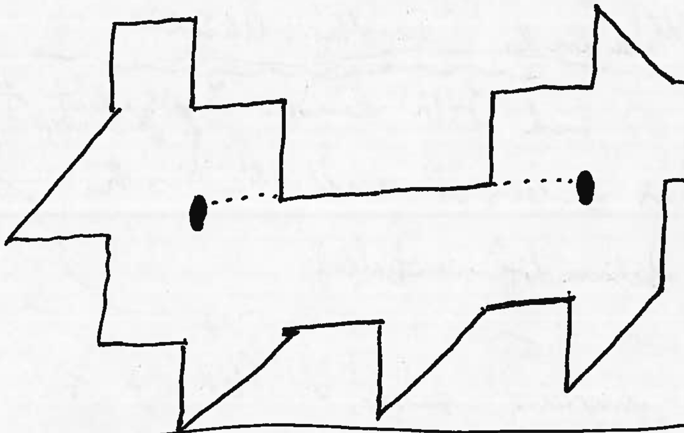


Fig. Rough sketch of Roger Penrose's room. The indicated point fails to illuminate the shaded regions

In 1995, George Tokarsky gave an example of a polygonal room in which at least one point fails to illuminate at least one other.



(The dotted line is a straight horizontal path between the mutually non-illuminating points. It hits the wall.)

Fig. Rough sketch of George Tokarsky's room. The indicated points do not illuminate each other

In 2014*, Lelièvre, Monteil, and Weiss proved that:

Theorem (Lelièvre-Monteil-Weiss): Let (X,ω) be a translation surface, and let $x \in X$. Then $\{y \in X \mid x \text{ does not illuminate } y\}$ is a finite set.

Their proof depends on the "Magic Wand" theorem. Here, we will sketch a proof of the above theorem using the Magic Wand, along with a result from 2016* (and hence not available at the time of L-M-W's proof)*. (I don't know of a source for this proof sketch. It was suggested to me in conversation with Alex Wright.)

In this sketch, we bypass the language of AISs by using Filip's 2016* theorem. See [W15]* for an introduction to AISs.

* What with these dates being so recent, and the time between appearance on the arXiv and publication so variable, I'm not very confident in any of these dates. Perhaps L-M-V were indeed aware of this proof sketch; the methods in their paper prove more that just the above theorem.

~~What with these dates being so recent, and the time between appearance on the arXiv and publication, I'm not very confident in any of these dates. Perhaps L-M-W were indeed aware of this proof sketch.~~

Theorem (Eskin-Mirzakhani-Mohammadi; "The Magic Wand"):

Let κ be a partition of $2g-2$, and $\mathcal{H}(\kappa)$ denote the subset of $\mathcal{H}$ of (X,ω) such that ω has zeroes of multiplicities κ. (We call this a "stratum" of abelian differentials).

Any closed $GL_2^+(\mathbb{R})$-invariant subset of $\mathcal{H}(\kappa)$ is a finite union of AISs.

Theorem (Filip): Every AIS is a quasi-projective variety over $\mathbb{C}$.

Proof sketch of finite illumination: Indeed, we may allow our partition to include any number of 0's. Then $\mathcal{H}(\kappa)$ is simply a collection of abelian differentials with certain marked points, the "zeroes of [redacted] multiplicity 0" for ω. Where κ is a partition containing only positive integers, let [redacted] us consider $\mathcal{H}(\kappa,0,0)$, a stratum of surfaces with two distinct marked points.

Consider the subset $N \subseteq \mathcal{H}(\kappa,0,0)$ of $((X,\omega),\{p_1,p_2\})$ such that p_1 does not illuminate p_2. It is evident that N is closed and $GL_2^+(\mathbb{R})$-invariant, and hence by the Magic Wand and the theorem of Filip, N is a quasiprojective complex variety.

~~Consider the subset~~

Fix $((X_0,\omega_0),p_1) \in \mathcal{H}(\kappa,0)$, and consider the [redacted] subset $V = \{((X_0,\omega_0),\{p_1,p_2\}) \in \mathcal{H}(\kappa,0,0) \mid p_2 \in X_0 \setminus p_1\}$.

It is not hard to see that V is an ~~embedding~~ embedded copy of the variety $X_0 \setminus p_1$ inside of $\mathcal{H}(k,0,0)$. Note that

$$V \cap N = \{((X_0,\omega_0),\{p_1,p_2\}) \in \mathcal{H}(k,0,0) \mid p_1 \text{ does not illuminate } p_2\}$$

It remains to show that $V \cap N$ is a finite set. Notice that $V \not\subseteq N$, and that $\dim_{\mathbb{C}} V = 1$. But intersections in algebraic geometry are nice! That is to say, $V \cap N$ is a 0-dimensional subvariety of V... a finite set! $\square$

References

§1. [W15] · Wright, Alex. Translation surfaces and their orbit closures: An introduction for a broad audience. arXiv: 1411.1827

§2. [MT02] · Masur, H. & Tabachnikov, S. Rational billiards and Flat Structures. Handbook of dynamical Systems, vol. 1A, ch. 13. Edited by B. Hasselblatt and A. Katok. Elsevier Science B.V., 2002.

§3. · Two part video series featuring Howard Masur, created by Brady Haran for the YouTube channel Numberphile.
- Part 1: The Illumination Problem. URL: youtube.com/watch?v=xhj5er1k6GQ
- Part 2: Problems with Periodic Orbits URL: youtube.com/watch?v=AGX0cLbHaog

· Illumination Problem, Wolfram MathWorld. URL: mathworld.wolfram.com/IlluminationProblem.html (and the references therein)

- Zorich, Anton. The Magic Wand Theorem of A. Eskin and M. Mirzakhani. arXiv: 1502.05654.
- Lelièvre, S.; Monteil, T.; Weiss, B. Everything is illuminated. Geom. Topol. Vol. 20, No. 3 (2016), 1737 - 1762. (arXiv: 1407.2975).
- Filip, Simion. Splitting mixed Hodge structures over affine invariant submanifolds. Ann. of Math. (2) 183 (2016), no. 2, 681 - 713 (arXiv: 1311.2350).