

FROM RATIONAL BILLIARDS TO DYNAMICS ON MODULI SPACES

ALEX WRIGHT

ABSTRACT. This short expository note gives an elementary introduction to the study of dynamics on certain moduli spaces, and in particular the recent breakthrough result of Eskin, Mirzakhani, and Mohammadi. We also discuss the context and applications of this result, and connections to other areas of mathematics such as algebraic geometry, Teichmüller theory, and ergodic theory on homogeneous spaces.

CONTENTS

1. Rational billiards	1
2. Translation surfaces	3
3. The $GL(2, \mathbb{R})$ action	7
4. Renormalization	9
5. Eskin-Mirzakhani-Mohammadi's breakthrough	10
6. Applications of Eskin-Mirzakhani-Mohammadi's Theorem	11
7. Context from homogeneous spaces	12
8. The structure of the proof	14
9. Relation to Teichmüller theory and algebraic geometry	15
10. What to read next	17
References	17

1. RATIONAL BILLIARDS

Consider a point bouncing around in a polygon. Away from the edges, the point moves at unit speed. At the edges, the point bounces according to the usual rule that angle of incidence equals angle of reflection. If the point hits a vertex, it stops moving. The path of the point is called a billiard trajectory.

The study of billiard trajectories is a basic problem in dynamical systems and arises naturally in physics. For example, consider two points of different masses moving on a interval, making elastic collisions

with each other and with the endpoints. This system can be modeled by billiard trajectories in a right angled triangle [MT02].

A rational polygon is a polygon all of whose angles are rational multiples of π . Many mathematicians are especially interested in billiards in rational polygons for the following three reasons.

First, without the rationality assumption, few tools are available, and not much is known. For example, it is not even known if every triangle has a periodic billiard trajectory. With the rationality assumption, quite a lot can be proven.

Second, even with the rationality assumption a wide range of interesting behavior is possible, depending on the choice of polygon.

Third, the rationality assumption leads to surprising and beautiful connections to algebraic geometry, Teichmüller theory, ergodic theory on homogenous spaces, and other areas of mathematics.

The assumption of rationality first arose from the following simple thought experiment. What if, instead of letting a billiard trajectory bounce off an edge of a polygon, we instead allowed the trajectory to continue straight, into a reflected copy of the polygon?

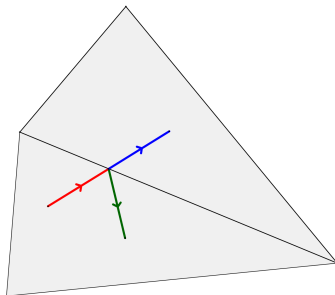


FIGURE 1.1. A billiard trajectory in a polygon P . Instead of allowing the trajectory to bounce off the edge of P , we may allow it to continue straight into a reflected copy of P . A key observation is that the trajectory that continues into the reflected copy of P is in fact the reflection of the trajectory in P that bounces off of the edge.

This leads us to define the “unfolding” of a polygon P as follows: Let G be the subgroup of $O(2)$ (linear isometries of \mathbb{R}^2) generated by the derivatives of reflections in the sides of P . The group G is finite if and only if the polygon P is rational (in which case G is a dihedral group). For each $g \in G$, consider the polygon gP . These polygons gP can be translated so that they are all disjoint in the plane. We identify the edges in pairs in the following way. Suppose r is the derivative of the reflection in one of the edges of hP . Then this edge of hP is identified with the corresponding edge of rhP .

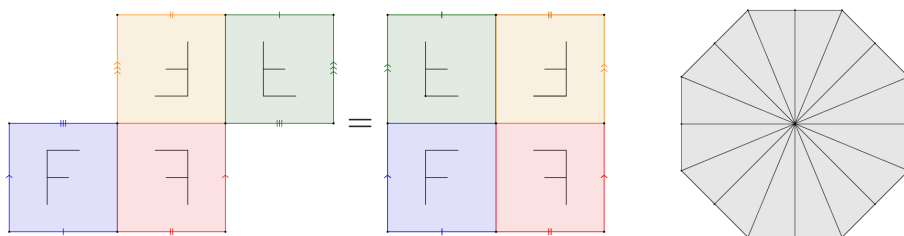


FIGURE 1.2. Left: The unit square unfolds to four squares, with opposite edges identified (a flat torus). (By “opposite edges” in these pictures, we mean pairs of boundary edges that are perpendicularly across from each other, so for example the top left and top right vertical edges on the unfolding of the square are opposite.) When two polygons are drawn with an adjacent edge, by convention this means these two adjacent edges are identified. Here each square has been decorated with the letter F, to illustrate which squares are reflections of other squares. Right: Unfolding the right angled triangle with smallest angle $\pi/8$ gives the regular octagon with opposite sides identified.

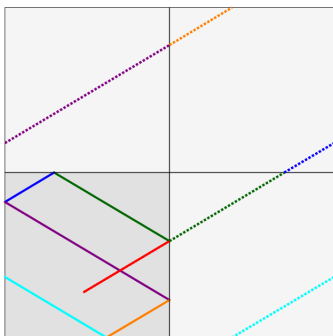


FIGURE 1.3. A billiard trajectory on a rational polygon unfolds to a straight line on the unfolding of the polygon. In this illustration, we have unfolded a billiard trajectory on square (bottom left) to a straight line on a flat torus. The square and its unfolding are superimposed, the billiard trajectory is drawn with a solid line, and the unfolded straight line is drawn with a dotted line.

The unfolding construction is most easily understood through examples: see Figures 1.2 and 1.3.

2. TRANSLATION SURFACES

Unfoldings of rational polygons are special examples of translation surfaces. There are several equivalent definitions of translation surface,

the most elementary of which is a finite union of polygons in the plane with edge identifications, obeying certain rules, up to a certain equivalence relation. The rules are:

- (1) The interiors of the polygons must be disjoint, and if two edges overlap then they must be identified.
- (2) Each edge is identified with exactly one other edge, which must be a translation of the first. The identification is via this translation.
- (3) When an edge of one polygon is identified with an edge of a different polygon, the polygons must be on “different sides” of the edge. For example, if a pair of vertical edges are identified, one must be on the left of one of the polygons, and the other must be on the right of the other polygon.

Two such families of polygons are considered to be equivalent if they can be related via a string of the following “cut and paste” moves.

- (1) A polygon can be translated.
- (2) A polygon can be cut in two along a straight line, to give two adjacent polygons.
- (3) Two adjacent polygons that share an edge can be glued to form a single polygon.

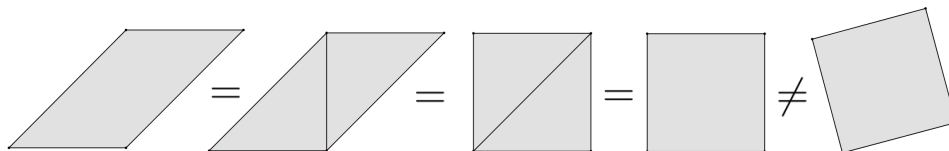


FIGURE 2.1. In all five translation surfaces above, opposite edges are identified. In the leftmost four, each adjacent pair of translation surfaces differs by one of the above three moves, so all four of these pictures give the same translation surface. The rightmost rotated surface is (presumably) not equal to the other four, since rotation is not one of the three allowed moves.

In general, it is difficult to decide if two collections of polygons as above are equivalent (describe the same translation surface), because each collection of polygons is equivalent to infinitely many others.

The requirements above ensure that the union of the polygons modulo edge identifications gives a closed surface. This surface has flat metric, given by the flat metric on the plane, away from a finite number of singularities. The singularities arise from the corners of the

polygons. For example, in the regular 8-gon with opposite sides identified, the 8 vertices give rise to a single point with cone angle. See Figure 2.2.

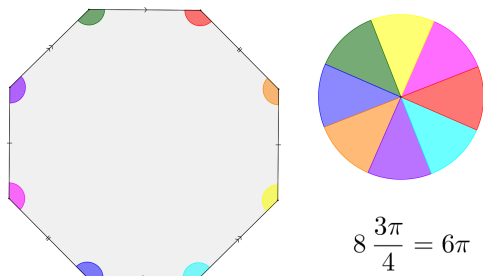


FIGURE 2.2. The edge identifications imply that the 8 corners of the octagon are in fact all identified, and give a single point on the translation surface. Around this point there is 6π total angle, since at each of the 8 corners of the polygon there is $\frac{3}{4}\pi$ interior angle.

The singularities of the flat metric on a translation surface are always of a very similar conical form, and the total angle around a singularity on a translation surface is always an integral multiple of 2π . Note that, although the flat metric is singular at these points, the underlying topological surface is not singular at any point. (That is, at every single point, including the singularities of the flat metric, the surface is locally homeomorphic to \mathbb{R}^2 .)

Most translation surfaces do not arise from unfoldings of rational polygons. This is because unfoldings of polygons are exceptionally symmetric, in that they are tiled by isometric copies of the polygon.

Translation surfaces satisfy a Gauss-Bonnet type theorem. If a translation surface has s singularities with cone angles

$$(1 + k_1)2\pi, (1 + k_2)2\pi, \dots, (1 + k_s)2\pi,$$

then the genus g is given by the formula $2g - 2 = \sum k_i$. (So, in a formal comparison to the usual Gauss-Bonnet formula, one might say that each extra 2π of angle on a translation surface counts for 1 unit of negative curvature.)

Consider now the question of how a given translation surface can be deformed to give other translation surfaces. The polygons, up to translation, can be recorded by their edge vectors in \mathbb{C} (plus some finite amount of combinatorial data, for example the cyclic order of edges around the polygons). Not all edge vectors need be recorded, since some are determined by the rest. Changing the edge vectors (subject

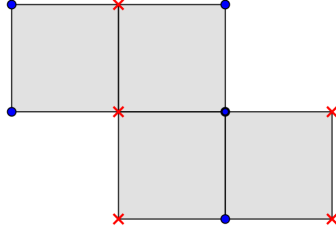


FIGURE 2.3. Consider the translation surface described by the above polygon, with opposite edges identified. This surface has two singularities, each with total angle 4π . One singularity has been labelled with a dot, and the other with an x. An Euler characteristic calculation ($V - E + F = 2 - 2g$) show that it has genus 2. The regular octagon with opposite sides identified also has genus 2, but it has only a single singularity, with total angle 6π .

to the condition that identified edges should remain parallel and of the same length) gives a deformation of the translation surface.

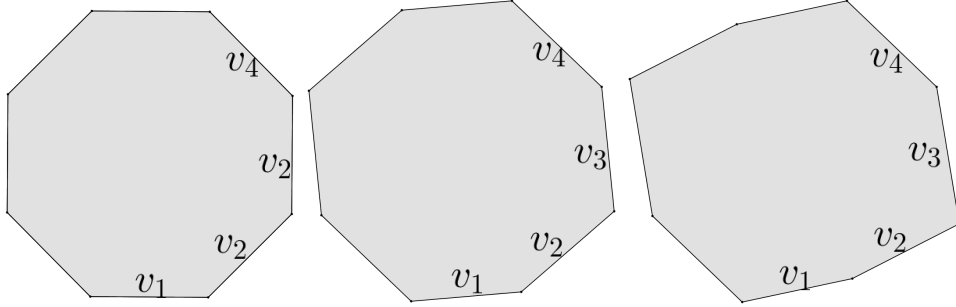


FIGURE 2.4. Any octagon whose opposite edges are parallel can be described by the 4-tuple of its edges vectors $(v_1, v_2, v_3, v_4) \in \mathbb{C}^4$. (Not all choices of v_i give valid octagons.) The coordinates (v_1, v_2, v_3, v_4) are local coordinates for space of deformations of the regular octagon translation surface. These coordinates are not canonical: other equally good coordinates can be obtained by cutting up the octagon and keeping track of different edge vectors.

To formalize this observation, we define moduli spaces of translation surfaces. Given an unordered collection k_1, \dots, k_s of positive integers whose sum is $2g - 2$, the stratum $\mathcal{H}(k_1, \dots, k_s)$ is defined to be the set of all translation surfaces with s singularities, of cone angles $(1 + k_i)2\pi$, $i = 1, \dots, s$. The genus of these surfaces must be g by the Gauss-Bonnet formula above. We have

Lemma 2.1. *Each stratum is a complex orbifold of dimension $n = 2g + s - 1$. Each stratum has a finite cover that is a manifold and has an atlas of charts to \mathbb{C}^n with transition functions in $GL(n, \mathbb{Z})$.*

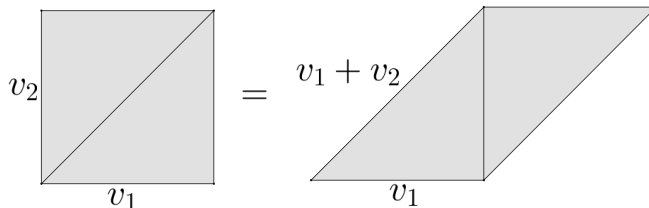


FIGURE 2.5. These two polygons (with opposite sides identified) both describe the same translation surface. Keeping track of the edge vectors in either polygon gives equally good local coordinates for the space of nearby translation surfaces. The two local coordinates thus obtained are related by the linear transformation $(v_1, v_2) \mapsto (v_1, v_1 + v_2)$.

The coordinate charts are called period coordinates. They consist of complex edge vectors of polygons. That strata are orbifolds instead of manifolds is a technical point that should be ignored by non-experts.

Strata are not always connected, but their connected components have been classified by Kontsevich and Zorich [KZ03]. There are always at most three connected components. The topology (and birational geometry) of strata is currently not well understood. Kontsevich has conjectured that strata are $K(\pi, 1)$ spaces.

3. THE $GL(2, \mathbb{R})$ ACTION

There is a $GL(2, \mathbb{R})$ action on each stratum, obtained by acting linearly on polygons and keeping the same identification.

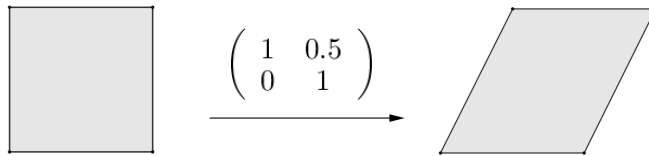


FIGURE 3.1. An example of the $GL(2, \mathbb{R})$ action. In both pictures, opposite edges are identified.

Note that if two edges or polygons differ by translation by a vector v , then their images under the linear map $g \in GL(2, \mathbb{R})$ must differ by translation by gv .

Example 3.1. The stabilizer of the standard flat torus (a unit square with opposite sides identified) is $GL(2, \mathbb{Z})$. For example, Figure 2.1 (near the beginning of the previous section) proves that $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is in the stabilizer. This example illustrates the complexity of the $GL(2, \mathbb{R})$ action: applying a large matrix (say of determinant 1) will yield a collection of very long and thin polygons, but it is hard to know when this collection of polygons is equivalent to a more reasonable one.

Translation surfaces have a well defined area, given by the sum of the areas of the polygons. The action of $SL(2, \mathbb{R})$ of determinant 1 matrices in $GL(2, \mathbb{R})$ preserves the locus of unit area translation surfaces. This locus is not compact, because the polygons can have edges of length going to 0, even while the total area stays constant.

Define

$$g_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} \in SL(2, \mathbb{R}).$$

Suppose one wants to know if a translation surface S has a vertical line joining singularities (cone points) of length e^{10} . This is equivalent to the question of whether $g_{10}(S)$ has a vertical line segment of length 1 joining two singularities.

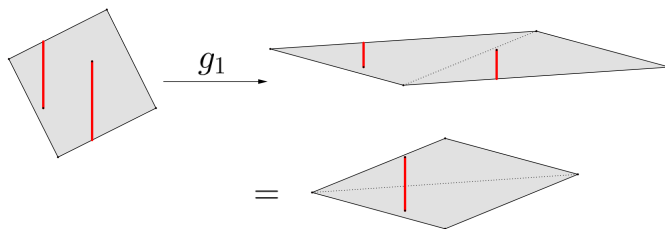


FIGURE 3.2. Applying g_t to a translation surface with a vertical line segment gives a new translation surface with a shorter vertical line segment.

In fact, for any matrix $g \in GL(2, \mathbb{R})$, the surfaces S and $g(S)$ are closely related, since one can go back and forth between them using g and g^{-1} . Really S and $g(S)$ are just different perspectives on the same object, in which different features are apparent. To understand S from all possible perspectives, we'd like to understand its $GL(2, \mathbb{R})$ orbit. However, just from definitions, it's not really possible to understand the $GL(2, \mathbb{R})$ orbit of a surface. It's really hard to know, given two surfaces S and S' , if there are large matrices g so that the polygons defining $g(S)$ can be cut up and reglued to be almost equal to the polygons defining S' !

4. RENORMALIZATION

This section aims to give some of the early motivations and successes of the field. It can be safely skipped and returned to later by anyone eager to get to the modern breakthrough and its applications and connections to other area of mathematics.

Suppose once again that we have a vertical line segment of length L on a translation surface S . For example, if S is the unfolding of a rational polygon, the vertical line might be the unfolding of a billiard trajectory. If one is interested in a line that is not vertical, one can rotate the whole picture (giving a different translation surface) so that it becomes vertical.

Applying g_t results in a translation surface $g_t(S)$ with a vertical segment of length $e^{-t}L$. We are interested in doing this when L is very large, and $t = \log(L)$ is chosen so the new vertical segment will have length 1. Indeed the point is to do this over and over as L gets longer, giving a family of surfaces $g_t(S)$.

This idea of taking longer and longer trajectories (here vertical lines on the translation surface) and replacing them with bounded length trajectories on new objects is called renormalization, and is a powerful and frequently used tool in the study of dynamical systems. The typical strategy is to transfer some understanding of the sequence of renormalized objects into results on the behavior of the original system. In this case, showing that the geometry of $g_t(S)$ does not degenerate allows good understanding of vertical lines on S .

Theorem 4.1 (Masur’s criterion [Mas92]). *Suppose $\{g_t(S) : t \geq 0\}$ does not diverge to infinity in the stratum. Then every infinite vertical line on S is equidistributed on S .*

“Equidistributed” is a technical term that indicates that the vertical lines becomes dense in S without favoring one part of S over another. Using this, Kerckhoff-Masur-Smillie [KMS86] were able to show

Theorem 4.2. *In every translation surface, for almost every slope, every infinite line of this slope is equidistributed.*

There are some surfaces where much more is true. For example, on the unit square with opposite sides identified, any line of rational slope is periodic, and every line of irrational slope is equidistributed. Genus one translation surfaces are quite special, because $GL(2, \mathbb{R})$ acts transitively on the space of genus one translation surfaces. In particular, the $GL(2, \mathbb{R})$ orbit of any genus one translation surface is closed, in a trivial way, since the orbit is the entire moduli space.

Theorem 4.3 (Veech Dichotomy). *If S is a translation surface with closed $GL(2, \mathbb{R})$ orbit, then for all but countably many slopes, every line with that slope is equidistributed. Moreover every line with slope contained in the countable set is periodic.*

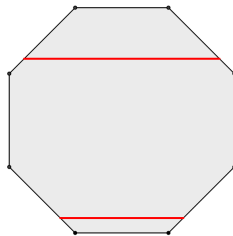


FIGURE 4.1. An example of a periodic line on the regular octagon with opposite sides identified.

Veech also showed that the regular $2n$ -gon with opposite sides identified has closed orbit. However, the property of having a closed orbit is extremely special.

Theorem 4.4 (Masur [Mas82], Veech [Vee82]). *The $GL(2, \mathbb{R})$ orbit of almost every translation surface is dense in a connected component of a stratum.*

For the experts, we remark that in fact Masur and Veech showed the stronger statement that the g_t action on the loci of unit area surfaces in a connected component of a stratum is ergodic, with respect to a Lebesgue class probability measure called the Masur-Veech measure.

This result of Masur and Veech is not satisfactory from the point of view of billiards in rational polygons, since the set of translation surfaces that are unfoldings of polygons has measure 0.

5. ESKIN-MIRZAKHANI-MOHAMMADI'S BREAKTHROUGH

The g_t orbit closure of a translation surface may be a fractal object. While this behavior might at first seem pathological, it is in fact quite typical in dynamical systems. Generally speaking, given a group action it is hugely unrealistic to ask for any understanding of every single orbit, since these may typically be arbitrarily complicated. Thus the following result is quite amazing.

Theorem 5.1 (Eskin-Mirzakhani-Mohammadi [EM, EMM]). *The $GL(2, \mathbb{R})$ orbit closure of a translation surface is always a manifold. Moreover,*

the manifolds that occur are locally defined by linear equations in period coordinates. These linear equations have real coefficients and zero constant term.

Note that although the local period coordinates are not canonical, if a manifold is cut out by linear equations in one choice of period coordinates, it must also be in any other overlapping choice of period coordinates, because the transition map between these two coordinates is a matrix in $GL(n, \mathbb{Z})$.

Previously orbit closures had been classified in genus 2 by McMullen. (One open problem remains in genus two, which is the classification of $SL(2, \mathbb{Z})$ orbits of square-tiled surfaces in $\mathcal{H}(1, 1)$.) The techniques of Eskin-Mirzakhani-Mohammadi's, unlike those of McMullen, are rather abstract, and have surprisingly little to do with translations surfaces. Thus the work of Eskin-Mirzakhani-Mohammadi does not give any information about how many or what sort of submanifolds arise as orbit closures, except for what is given in the theorem statement.

6. APPLICATIONS OF ESKIN-MIRZAKHANI-MOHAMMADI'S THEOREM

There are many applications of Theorem 5.1 to translation surfaces, rational billiards, and other related dynamics systems, for example interval exchange transformations. Here we list just a few of the most easily understood applications.

Generalized diagonals in rational polygons. Let P be a rational polygon. A generalized diagonal is a billiard trajectory that begins and ends at a corner of P . If P is a square, an example is a diagonal of P . Let $N_P(L)$ be the number of generalized diagonals in P of length at most L . It is a folklore conjecture that

$$\lim_{L \rightarrow \infty} \frac{N_P(L)}{L^2}$$

exists for every P and is non-zero. Previously, Masur had shown that the limsup and liminf are non-zero and finite [Mas90, Mas88]. Eskin-Mirzakhani-Mohammadi give the best result to date: with some additional Cesaro type averaging, the conjecture is true, and furthermore only countably many real numbers may occur as such a limit.

The illumination problem. Given a polygon P and two points x and y , say that y is illuminated by x if there is a billiard trajectory going from y to x . This terminology is motivated by thinking of P as a polygonal room whose walls are mirrors, and thinking of a candle placed

at x . The light rays travel along billiard trajectories. We emphasize that the polygon need not be convex.

Lelièvre, Monteil and Weiss have shown that if P is a rational polygon, for every x there are at most finitely many y not illuminated by x [LMW].

The Wind Tree Model. This model arose from physics, and is sometimes called the Ehrenfest model. Consider the plane with periodically shaped rectangular barriers (“trees”). Consider a particle (of “wind”) which moves at unit speed and collides elastically with the barriers.

Delecroix-Hubert-Lelièvre have determined the divergence rate of the particle for all choices of size of the rectangular barriers [DHL14]. Without Eskin-Mirzakhani-Mohammadi (and work of Chaika-Eskin [CE]), the best that could be proven was the existence of a unspecified full measure set of choices of sizes for which such a result holds.

There are many other examples along these lines, where previously results were known to hold for almost all examples without being known to hold in any particular example, and now with Eskin-Mirzakhani-Mohammadi can be upgraded to hold in all cases.

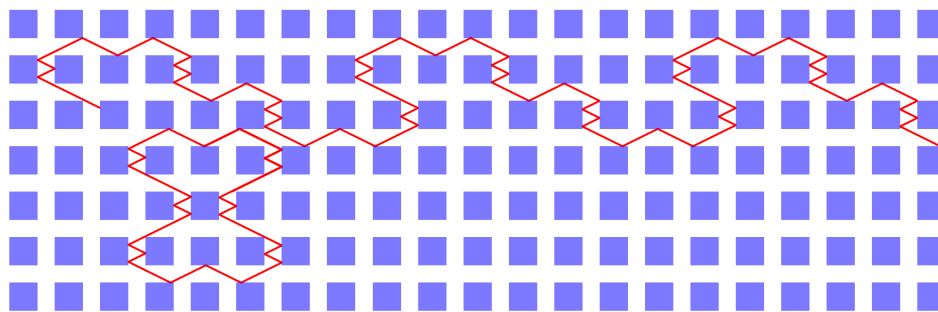


FIGURE 6.1. An example of a trajectory in the Wind Tree Model. Figure courtesy of Vincent Delecroix.

Applications of Eskin-Mirzakhani’s proof. The ideas that Eskin-Mirzakhani developed have applications beyond moduli spaces of translation surfaces. They are currently being used by Rodriguez-Hertz and Brown to study random diffeomorphisms on surfaces [BRH] and the Zimmer program (lattice actions on manifolds), and are also expected to have applications in ergodic theory on homogeneous spaces.

7. CONTEXT FROM HOMOGENEOUS SPACES

The primary motivation for Theorem 5.1 is the following theorem.

Theorem 7.1 (Ratner’s Theorem). *Let G be a Lie group, and let $\Gamma \subset G$ be a lattice. Let $H \subset G$ be a subgroup generated by unipotent one parameter groups. Then every H orbit closure in G/Γ is a manifold, and moreover is a sub-homogeneous space.*

For example, the theorem applies if $G = SL(3, \mathbb{R})$, $\Gamma = SL(3, \mathbb{Z})$, and $H = \{h_t : t \in \mathbb{R}\}$, where

$$h_t = \begin{pmatrix} 1 & t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{or} \quad h_t = \begin{pmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}.$$

Ratner’s work confirmed a conjecture of Raghunathan. Special cases of this conjecture had previously been verified by Dani and Margulis, for example for the second choice of h_t above [DM90].

The basic idea behind such proofs is the strategy of additional invariance. Given a closed H -invariant set, one starts with two points x and y very close together, and applies h_t until the points drift apart. The direction of drift is controlled by another one parameter subgroup, and one tries to show that the closed H -invariant set is in fact also invariant under the one parameter group that gives the direction of drift. One continues this argument inductively, each time producing another one parameter group the set is invariant under, until one shows that the closed H invariant set is in fact invariant under a larger group L , and is contained in (and hence equal to) a single L orbit. This gives the set in question is homogenous, and in particular a manifold.

Of course, this is in fact very difficult, and complete proofs of Ratner’s Theorem are very long and technical. For one thing, as is the case in the work of Eskin-Mirzakhani, it is in fact too difficult to work directly with closed invariant sets, as we have just suggested. Rather, one first must classify invariant measures. Thus the argument takes place in the realm of ergodic theory, which exactly studies group actions on spaces with invariant measures. See [Mor05, Ein06] for an introduction.

The fundamental requirement of the proof is that orbits of nearby points drift apart slowly and in a controlled way. This is intimately tied to the fact that unipotent one parameter groups are polynomial, as can be seen in the h_t above. Contrast this to the one parameter group

$$\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix},$$

whose orbit closures may be fractal sets.

One might hope to study $GL(2, \mathbb{R})$ orbit closures of translation surfaces using the action of

$$u_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

on strata, in analogy to the proof of Ratner's Theorem. Unfortunately, the dynamics of the u_t action on strata is currently much too poorly understood for this.

8. THE STRUCTURE OF THE PROOF

The proof of Theorem 5.1 builds on many ideas in homogeneous space dynamics, including the work of Benoist-Quint [BQ09], and the high and low entropy methods of Einsiedler-Lindenstrauss-Katok [EKL06] (see also the beautiful introduction for a general audience by Venkatesh [Ven08]). Lindenstrauss won the Fields medal in 2010 partially for the development of the low entropy method, and Benoist and Quint won the Clay prize for their work.

Entropy measures the unpredictability of a system that evolves over time.

Define P to be the upper triangular subgroup of $SL(2, \mathbb{R})$. The proof of Theorem 5.1 proceeds in two main stages.

In the first, Eskin-Mirzakhani show that any ergodic P invariant measure is in fact a Lebesgue class measure on a manifold cut out by linear equations, and must be $SL(2, \mathbb{R})$ invariant. (An ergodic measure is an invariant measure which is not the average of two other invariant measures in a nontrivial way. Thus the ergodic measures are the building blocks for all other invariant measures.)

In the second stage, Eskin-Mirzakhani-Mohammadi use this to prove Theorem 5.1, by constructing a P -invariant measure on every P -orbit closure. By contrast, it is not possible to directly construct an $SL(2, \mathbb{R})$ invariant measure on each $SL(2, \mathbb{R})$ orbit closure, and this is why the use of P is crucial. The algebraic structure of P makes it possible to average over larger and larger subsets of P and thus produce P invariant measures, whereas the more complicated algebraic structure of $SL(2, \mathbb{R})$ does not allow this. (The relevant property is that P is amenable, while $SL(2, \mathbb{R})$ is not.)

In the paper of Eskin-Mirzakhani, which carries out the first stage, the most difficult part is in fact to show P -invariant measures are $SL(2, \mathbb{R})$ invariant. To do this, extensive entropy arguments are used, partially inspired by the Margulis-Tomanov proof of Ratner's Theorem [MT94] and to a lesser extent the high and low entropy methods. This part is the technical heart of the argument, and takes almost 100 pages of

delicate arguments. One of the morals is that entropy arguments are surprisingly effective in this context, and can be made to work without the use of an ergodic theorem.

Once Eskin-Mirzakhani show that P -invariant measures are $SL(2, \mathbb{R})$ invariant, they build upon ideas of Benoist-Quint to conclude that every such measure is a Lebesgue class measure on a manifold cut out by linear equations

All together, the proof is remarkably abstract. The only facts used about translation surfaces are formulas for Lyapunov exponents due to Forni [For02]. (The Lyapunov exponents of a smooth dynamical system, in this case the action of g_t on a stratum, measure the rates of expansion and contraction in different directions.) Forni was awarded the Brin prize partially for these formulas, which are of a remarkably analytic nature and arose from an insight of Kontsevich [Kon97]. Eskin-Mirzakhani also make use of a result of Filip [Filb] to handle a volume normalization issue at the end of the proof.

Results which classify invariant measures are rare gems. The arguments are abstract, and their purpose is to rule out nonexistent objects, and thus they cannot be guided by examples. To obtain a truly new measure classification result, one must find truly new ideas from among the sea of ideas which don't quite work, and the devil is in the details. This is very much the case in the paper of Eskin-Mirzakhani.

9. RELATION TO TEICHMÜLLER THEORY AND ALGEBRAIC GEOMETRY

Every translation surface has, in particular, the structure of a Riemann surface X . The extra flat structure is determined by additionally specifying an Abelian differential ω (a.k.a. holomorphic one form, or global section of the canonical bundle of X). The holomorphic one form is dz on the polygons, where z is the usual coordinate on the plane $\mathbb{C} \simeq \mathbb{R}^2$. It has zeros at the singularities of the flat metric.

Thus every translation surface can be given as a pair (X, ω) . For example, the translation surface given by a square with unit area and opposite edges identified is $(\mathbb{C}/\mathbb{Z}[i], dz)$.

There is a projection map $(X, \omega) \mapsto X$ from a stratum of translation surfaces of genus g to the moduli space \mathcal{M}_g of Riemann surfaces of genus g . Under this map, g_t orbits of translation surfaces project to geodesics for the Teichmüller metric. But it is important to note that there is no $GL(2, \mathbb{R})$ or g_t action on \mathcal{M}_g itself, only on (the strata) of the bundle of Abelian differentials over \mathcal{M}_g . This is somewhat analogous to the fact that, given a Riemannian manifold, the geodesic flow is

defined on the tangent bundle, and there is no naturally related flow on the manifold itself.

The map $(X, \omega) \mapsto X$ has fibers of real dimension two (given by multiplying ω by any complex number), and thus the projection of a four dimensional $GL(2, \mathbb{R})$ orbit to \mathcal{M}_g is a two real dimensional object. It turns out that this object is an isometrically immersed copy of the upper half plane in \mathbb{C} with its hyperbolic metric: such objects are called complex geodesics or Teichmüller disks. Note that Royden showed that the Teichmüller metric is equal to the Kobayashi metric on \mathcal{M}_g .

McMullen showed that every $GL(2, \mathbb{R})$ orbit closure in genus 2 is either a closed orbit, or an eigenform locus, or a stratum [McM03]. In particular, every $GL(2, \mathbb{R})$ orbit closure of genus 2 translation surfaces is a quasi-projective variety. The corresponding statement for \mathcal{M}_2 is that every complex geodesic is either closed, or dense in a Hilbert modular surface, or dense in \mathcal{M}_2 .

The appearance of algebraic geometry in the study of orbit closures was unexpected, and arose in very different ways from work of McMullen and Kontsevich. A recent success in this direction is the following, which builds upon Theorem 5.1 and work of Möller [Möl06b, Möl06a].

Theorem 9.1 (Filip [Fila]). *In every genus, every $GL(2, \mathbb{R})$ orbit closure is an algebraic variety that parameterizes pairs (X, ω) with special algebro-geometric properties, such as $\text{Jac}(X)$ having real multiplication.*

Furthermore, Filip's work gives an algebro-geometric characterization of orbit closures: They are loci of (X, ω) with special algebraic properties, when these loci have maximal possible dimension. (When these loci do not have maximum possible dimension, they do not give orbit closures are not of interest in dynamics.) It is of yet unclear how to make use of this characterization, since calculating the dimension of the relevant sub-varieties of \mathcal{M}_g seems to be an incredibly difficult problem, very close to the Schottky problem. Nonetheless Filip's work means that an equivalent definition of orbit closure can be given to an algebraic geometer, without mentioning flat geometry or even the $GL(2, \mathbb{R})$ action!

It is at present a major open problem to classify $GL(2, \mathbb{R})$ orbit closures. Progress is ongoing using techniques based on flat geometry and dynamics, see for example [Wri14, NW14, ANW]. One of the key tools is the author's Cylinder Deformation Theorem, which allows one to produce deformations of a translation surface that remain in the

$GL(2, \mathbb{R})$ orbit closure, without actually computing any surfaces in the $GL(2, \mathbb{R})$ orbit [Wri15a].

10. WHAT TO READ NEXT

We recommend the two page “What is . . . measure rigidity” article by Einsiedler [Ein09], as well as the eight page note “The mathematical work of Maryam Mirzakhani” by McMullen [McM14] and the 13 page note “The magic wand theorem of A. Eskin and M. Mirzakhani” by Zorich [Zor14, Zor]. There are a large number of surveys on translation surfaces, for example [MT02, Zor06] and the author’s recent introduction aimed at a broad audience [Wri15b]. Alex Eskin has given a mini-course on his paper with Mirzakhani, and notes are available on his website [Esk].

Acknowledgements. I am very grateful to Yiwei She, Max Engelstein, Weston Ungemach, Aaron Pollack, Amir Mohammadi, Barak Weiss, Anton Zorich, and Howard Masur for making helpful comments and suggestions, which lead to significant improvements. This expository article grew out of notes for a talk in the Current Events Bulletin at the Joint Meetings in January 2015. I thank David Eisenbud for organizing the Current Events Bulletin and inviting me to participate, and Susan Friedlander for encouraging me to publish this article.

REFERENCES

- [ANW] David Aulicino, Duc-Manh Nguyen, and Alex Wright, *Classification of higher rank orbit closures in $\mathcal{H}^{\text{odd}}(4)$* , preprint, arXiv 1308.5879 (2013), to appear in *Geom. Top.*
- [BQ09] Yves Benoist and Jean-François Quint, *Mesures stationnaires et fermés invariants des espaces homogènes*, *C. R. Math. Acad. Sci. Paris* **347** (2009), no. 1-2, 9–13.
- [BRH] Aaron Brown and Federico Rodriguez Hertz, *Measure rigidity for random dynamics on surfaces and related skew products*, preprint, arXiv 1406.7201 (2014).
- [CE] Jon Chaika and Alex Eskin, *Every flat surface is Birkhoff and Oseledets generic in almost every direction*, preprint, arXiv 1305.1104 (2014).
- [DHL14] Vincent Delecroix, Pascal Hubert, and Samuel Lelièvre, *Diffusion for the periodic wind-tree model*, *Ann. Sci. Éc. Norm. Supér. (4)* **47** (2014), no. 6, 1085–1110.
- [DM90] S. G. Dani and G. A. Margulis, *Orbit closures of generic unipotent flows on homogeneous spaces of $SL(3, \mathbb{R})$* , *Math. Ann.* **286** (1990), no. 1-3, 101–128.
- [Ein06] Manfred Einsiedler, *Ratner’s theorem on $SL(2, \mathbb{R})$ -invariant measures*, *Jahresber. Deutsch. Math.-Verein.* **108** (2006), no. 3, 143–164.
- [Ein09] ———, *What is . . . measure rigidity?*, *Notices Amer. Math. Soc.* **56** (2009), no. 5, 600–601.

- [EKL06] Manfred Einsiedler, Anatole Katok, and Elon Lindenstrauss, *Invariant measures and the set of exceptions to Littlewood's conjecture*, Ann. of Math. (2) **164** (2006), no. 2, 513–560.
- [EM] Alex Eskin and Maryam Mirzakhani, *Invariant and stationary measures for the $SL(2, \mathbb{R})$ action on moduli space*, preprint, arXiv 1302.3320 (2013).
- [EMM] Alex Eskin, Maryam Mirzakhani, and Amir Mohammadi, *Isolation theorems for $SL(2, \mathbb{R})$ -invariant submanifolds in moduli space*, preprint, arXiv:1305.3015 (2013).
- [Esk] Alex Eskin, *Lectures on the $SL(2, \mathbb{R})$ action on moduli space*, math.uchicago.edu/~eskin/luminy2012/lectures.pdf.
- [Fila] Simion Filip, *Semisimplicity and rigidity of the Kontsevich-Zorich cocycle*, preprint, arXiv:1307.7314 (2013).
- [Filb] ———, *Splitting mixed Hodge structures over affine invariant manifolds*, preprint, arXiv:1311.2350 (2013).
- [For02] Giovanni Forni, *Deviation of ergodic averages for area-preserving flows on surfaces of higher genus*, Ann. of Math. (2) **155** (2002), no. 1, 1–103.
- [KMS86] Steven Kerckhoff, Howard Masur, and John Smillie, *Ergodicity of billiard flows and quadratic differentials*, Ann. of Math. (2) **124** (1986), no. 2, 293–311.
- [Kon97] M. Kontsevich, *Lyapunov exponents and Hodge theory*, The mathematical beauty of physics (Saclay, 1996), Adv. Ser. Math. Phys., vol. 24, World Sci. Publ., River Edge, NJ, 1997, pp. 318–332.
- [KZ03] Maxim Kontsevich and Anton Zorich, *Connected components of the moduli spaces of Abelian differentials with prescribed singularities*, Invent. Math. **153** (2003), no. 3, 631–678.
- [LMW] Samuel Lelièvre, Thierry Monteil, and Barak Weiss, *Everything is illuminated*, preprint, arXiv 1407.2975 (2014).
- [Mas82] Howard Masur, *Interval exchange transformations and measured foliations*, Ann. of Math. (2) **115** (1982), no. 1, 169–200.
- [Mas88] ———, *Lower bounds for the number of saddle connections and closed trajectories of a quadratic differential*, Holomorphic functions and moduli, Vol. I (Berkeley, CA, 1986), Math. Sci. Res. Inst. Publ., vol. 10, Springer, New York, 1988, pp. 215–228.
- [Mas90] ———, *The growth rate of trajectories of a quadratic differential*, Ergodic Theory Dynam. Systems **10** (1990), no. 1, 151–176.
- [Mas92] ———, *Hausdorff dimension of the set of nonergodic foliations of a quadratic differential*, Duke Math. J. **66** (1992), no. 3, 387–442.
- [McM03] Curtis T. McMullen, *Billiards and Teichmüller curves on Hilbert modular surfaces*, J. Amer. Math. Soc. **16** (2003), no. 4, 857–885 (electronic).
- [McM14] Curtis McMullen, *The mathematical work of Maryam Mirzakhani*, Proceedings of the International Congress of Mathematicians (Seoul 2014), vol. 1, Kyung Moon SA Co. Ltd., Berlin, 2014, pp. 73–79.
- [Möl06a] Martin Möller, *Periodic points on Veech surfaces and the Mordell-Weil group over a Teichmüller curve*, Invent. Math. **165** (2006), no. 3, 633–649.
- [Möl06b] ———, *Variations of Hodge structures of a Teichmüller curve*, J. Amer. Math. Soc. **19** (2006), no. 2, 327–344 (electronic).

- [Mor05] Dave Witte Morris, *Ratner's theorems on unipotent flows*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 2005.
- [MT94] G. A. Margulis and G. M. Tomanov, *Invariant measures for actions of unipotent groups over local fields on homogeneous spaces*, Invent. Math. **116** (1994), no. 1-3, 347–392.
- [MT02] Howard Masur and Serge Tabachnikov, *Rational billiards and flat structures*, Handbook of dynamical systems, Vol. 1A, North-Holland, Amsterdam, 2002, pp. 1015–1089.
- [NW14] Duc-Manh Nguyen and Alex Wright, *Non-Veech surfaces in $\mathcal{H}^{\text{hyp}}(4)$ are generic*, Geom. Funct. Anal. **24** (2014), no. 4, 1316–1335.
- [Vee82] William A. Veech, *Gauss measures for transformations on the space of interval exchange maps*, Ann. of Math. (2) **115** (1982), no. 1, 201–242.
- [Ven08] Akshay Venkatesh, *The work of Einsiedler, Katok and Lindenstrauss on the Littlewood conjecture*, Bull. Amer. Math. Soc. (N.S.) **45** (2008), no. 1, 117–134.
- [Wri14] Alex Wright, *The field of definition of affine invariant submanifolds of the moduli space of abelian differentials*, Geom. Topol. **18** (2014), no. 3, 1323–1341.
- [Wri15a] ———, *Cylinder deformations in orbit closures of translation surfaces*, Geom. Topol. **19** (2015), no. 1, 413–438.
- [Wri15b] ———, *Translation surfaces and their orbit closures: An introduction for a broad audience*, EMS Surv. Math. Sci. **2** (2015), no. 1, 63–108.
- [Zor] Anton Zorich, *The magic wand theorem of A. Eskin and M. Mirzakhani*, English translation of paper in Gaz. Math., arXiv 1502.05654 (2015).
- [Zor06] ———, *Flat surfaces*, Frontiers in number theory, physics, and geometry. I, Springer, Berlin, 2006, pp. 437–583.
- [Zor14] ———, *Le théorème de la baguette magique de A. Eskin et M. Mirzakhani*, Gaz. Math. (2014), no. 142, 39–54.

E-mail address: alexmurraywright@gmail.com