

THE CONFORMAL GEOMETRY OF BILLIARDS

LAURA DE MARCO

In the absence of friction and other external forces, a billiard ball on a rectangular billiard table follows a predictable path. As it reflects off the sides of the table, the ball will either (1) return to its original position and direction, or (2) travel all over the table, spending equal time in regions with equal area. Because of this dichotomy, we say the rectangular table is optimal: billiard trajectories define an easily understood dynamical system on the phase space of positions and directions.

In this article, we ask:

Which polygonal shapes make optimal billiard tables?

Our primary goal is to tell a story about billiards which provides an entry point to a large and active area of research. We will explore some recent results on the search for, and the classification of, the optimal billiard tables. We will also discuss some general results about the Teichmüller geometry of the moduli space of Riemann surfaces and its connection with billiard dynamical systems.

We concentrate on the billiard dynamics in polygons $T \subset \mathbb{R}^2$ with all interior angles equal to rational multiples of π . Each such billiard table gives rise to

- a compact translation surface, that is, a pair (X, ω) consisting of a compact Riemann surface X and a holomorphic 1-form ω on X , via a process called unfolding; and
- a discrete group of 2-by-2 matrices $\mathrm{SL}(X, \omega) \subset \mathrm{SL}_2\mathbb{R}$, consisting of the linear parts of affine automorphisms of the surface (X, ω) .

The path of a billiard ball, which bounces around the table by reflecting off the sides, unfolds into a straight line on the surface X , with respect to the Euclidean metric determined by ω . Thus, the study of billiard dynamics on a table becomes the study of the geodesic flow on $(X, |\omega|)$. The structure of the group $\mathrm{SL}(X, \omega)$ will determine properties of the geodesic flow.

We will introduce three types of optimal billiard tables; namely, the lattice polygons, the ergodically optimal tables, and the topologically optimal tables. Of the three, the lattice polygon is the most restrictive type. By definition, the billiard table T is a lattice polygon if $\mathrm{SL}(X, \omega)$ forms a lattice in $\mathrm{SL}_2\mathbb{R}$; this means that the quotient $\mathrm{SL}_2\mathbb{R}/\mathrm{SL}(X, \omega)$ has finite volume. Traditional rectangular tables and the pentagon of Figure 1 are examples.

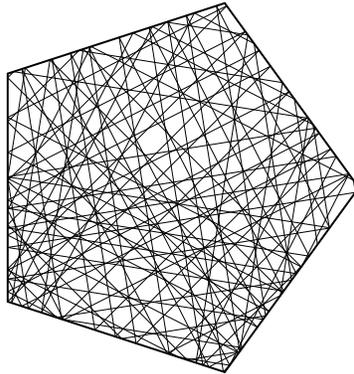


FIGURE 1. A billiard trajectory on the regular pentagon.

The ergodically optimal and topologically optimal tables are defined purely in terms of their billiard dynamical systems. On an ergodically optimal table, any billiard trajectory (which avoids the corners) is either periodic or uniformly distributed, while topologically optimal tables satisfy a weaker dichotomy: trajectories are either periodic or dense. Lattice polygons are both ergodically and topologically optimal, but a characterization of lattice polygons only in terms of their billiard dynamics remains open.

There is an additional dynamical system, an action of $\mathrm{SL}_2\mathbb{R}$ on the moduli space $\Omega\mathcal{M}_g$ of all compact translation surfaces of genus g . When the surface (X, ω) arises from a billiard table T , it is remarkable that properties of its orbit under the $\mathrm{SL}_2\mathbb{R}$ action determine properties of the original billiard dynamical system on T . In fact, the group $\mathrm{SL}(X, \omega)$, defined by symmetries of (X, ω) , coincides with the stabilizer of (X, ω) under this $\mathrm{SL}_2\mathbb{R}$ action. The question of which shapes make optimal billiard tables thus broadens into questions about global invariants of this action on the moduli space.

We will see a few of the general results about this $\mathrm{SL}_2\mathbb{R}$ action on $\Omega\mathcal{M}_g$, but we will keep billiard dynamics as the main focus of this article. We give no proofs, but we provide references throughout.

Acknowledgements. I am greatly indebted to the experts in translation surfaces for helping me prepare this article. Special thanks go to Yitwah Cheung, Howard Masur, and Curt McMullen for lengthy conversations and numerous email exchanges. Curt McMullen generated the images of Figures 1 and 3. I would also like to thank Jayadev Athreya, Matt Bainbridge, Alex Eskin, Pat Hooper, Chris Judge, Ben McReynolds, and John Smillie for helpful discussions. This article does not do justice to their beautiful results; the emphasis here reflects my own bias and background.

This article is an expanded version of my notes for the Current Events session at the January 2010 meeting of the American Mathematical Society in San Francisco. I thank the AMS and David Eisenbud for inviting me to learn about billiards. My research is supported by the National Science Foundation and the Sloan Foundation.

CONTENTS

1. Billiard tables and optimal dynamics	4
1.1. The square table	4
1.2. Ergodically optimal billiard dynamics	4
1.3. First examples and non-examples	5
1.4. Dense but non-uniformly distributed trajectories	6
1.5. Topologically optimal billiard tables	7
2. The translation surface of a billiard table	7
2.1. Translation structure on a Riemann surface	7
2.2. Unfolding a billiard table into a translation surface	8
2.3. Surfaces from tables, but not vice versa	9
2.4. Geodesic flow on a translation surface	9
2.5. The affine group of symmetries and its linear part	9
2.6. Lattice polygons and surfaces	9
3. An $SL_2\mathbb{R}$ action on moduli space	10
3.1. Stretching the Euclidean coordinates	10
3.2. The moduli space of translation surfaces	11
3.3. Teichmüller geometry and Teichmüller curves	11
3.4. Masur's theorem and the Veech dichotomy	12
3.5. Genus 1	13
3.6. Genus 2 and geometric primitivity	13
3.7. Higher genus and algebraic primitivity	14
3.8. Invariant measures and orbit closures	15
4. Towards a characterization of optimal dynamics or the lattice condition	15
4.1. Characterization in genus 2	15
4.2. Coverings and branched coverings	17
4.3. The triangle of Theorem 1.3	18
4.4. Possible characterization of the lattice condition: rational ratios of moduli	18
5. Examples of groups $SL(X, \omega)$ and the search for optimal billiard tables	19
5.1. Arithmetic groups	19
5.2. Triangle groups	19
5.3. Infinitely generated groups	20

5.4. Triangles and the continued search	21
References	22

1. BILLIARD TABLES AND OPTIMAL DYNAMICS

For this article, a *billiard table* means a connected polygon in \mathbb{R}^2 with all angles equal to rational multiples of π . A *billiard trajectory* in direction θ is a straight-line path which begins at some point in the interior of the table, at angle θ as measured from the positive real axis, and bounces off the edges with angle of reflection equal to the angle of incidence. As the angles are rational multiples of π , the billiard path will travel again in direction θ after finitely many reflections off the sides of the table. If a billiard trajectory hits a vertex of the polygon, it stops. See Figure 2.

In this section, we define two notions of optimal billiard dynamics: ergodically optimal and topologically optimal.

1.1. The square table. The simplest example is the square table of side length 1, with sides parallel to the coordinate axes. In this table, it is easy to see that any billiard trajectory of rational slope will either hit a vertex or eventually return to its original configuration (position and angle). On the other hand, for all irrational slopes and any initial point, a trajectory will either encounter a vertex or it will bounce around the table spending equal time in parts with equal area. In particular, the closure of any infinite trajectory with irrational slope will be the whole table.

1.2. Ergodically optimal billiard dynamics. We say a billiard table is *ergodically optimal* if for each direction θ , one of the following holds:

- (1) every trajectory that avoids the vertices is periodic; or
- (2) every trajectory that avoids the vertices is uniformly distributed.

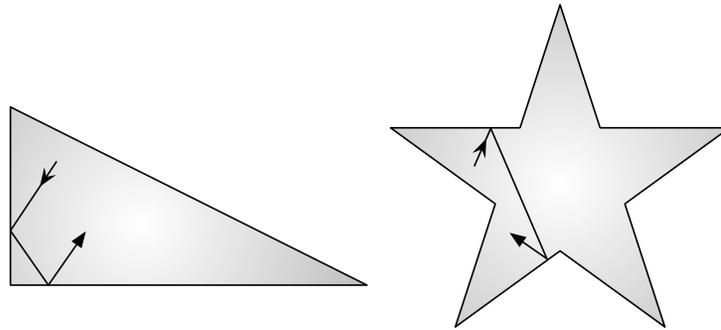


FIGURE 2. Polygonal billiard tables and trajectories.

We must take care in our meaning of uniform distribution: the trajectory determines a closed, invariant surface in the phase space (of all possible positions and directions), and the trajectory must spend equal time in regions of the surface with equal area. To make this more precise, note that every trajectory points in only finitely many different directions under reflections off the sides of the tables, because all angles are rational multiples of π . For each direction θ , we take multiple copies of the polygonal table, one for each direction arising by reflection of trajectories in direction θ . For any trajectory of infinite length that starts in direction θ , let $\gamma(t)$, $t \geq 0$, be a parametrization of this trajectory with unit speed (so with each reflection in a side, $\gamma(t)$ jumps to another copy of the table). For each time $s > 0$, we can define a probability measure on the union of tables by

$$\frac{1}{s} \gamma_* m_s$$

where m_s is arc-length measure on the interval $[0, s]$ in \mathbb{R} . Uniform distribution means that this family of measures converges weakly as $s \rightarrow \infty$ to normalized area measure on the finite union of polygonal tables.

The property of ergodically optimal dynamics is also called *Veech dichotomy* in the literature. The uniform distribution of all (infinite) trajectories in some direction is called *unique ergodicity*.

1.3. First examples and non-examples. As explained in §1.1, the unit square table is ergodically optimal. By similar reasoning, any rectangular table will be ergodically optimal. With more work, it is also possible to show:

Theorem 1.1. [Gu] *Any billiard table which is tiled by squares is ergodically optimal.*

A tiling by squares requires that all tiles be translates of a single square, and all vertices of the polygon lie on vertices of the tiles. Using different methods, Veech showed:

Theorem 1.2. [Ve] *For every $n \geq 3$, the regular n -gon is ergodically optimal.*

On the other hand, it is easy to construct tables which are not ergodically optimal. For example, begin with a square table and attach a rectangle to one side with side lengths a and b where $a/b \notin \mathbb{Q}$. See Figure 3. Taking the direction $\theta = \pi/4$, we see that the trajectories in direction θ which enter the smaller rectangle are neither closed nor dense.

We shall see many more ergodically optimal examples throughout this article. In Section 3, we relate this dynamical property of being ergodically optimal to the Teichmüller geometry of moduli space and the stronger condition of being a lattice polygon.

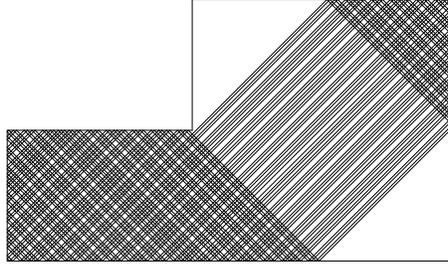


FIGURE 3. A billiard trajectory on an L -shaped table, neither closed nor dense.

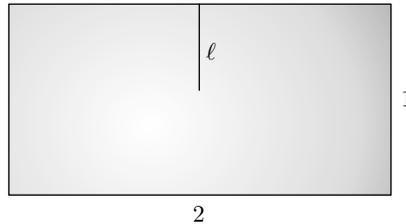


FIGURE 4. A rectangular billiard table with barrier of length ℓ .

1.4. Dense but non-uniformly distributed trajectories. It is a non-trivial task to find billiard trajectories which are dense but non-uniformly distributed, but they do exist. The following examples were studied by Masur and Smillie, following a construction of Veech; see the discussion in [MT, §3.1]. Consider a rectangular table with barrier as in Figure 4: begin with a rectangular table of side lengths 1 and 2 and build a perpendicular wall in the middle of the long side of length $\ell < 1$. When ℓ is rational, the table is ergodically optimal. When ℓ is irrational, there are orbits which are neither dense nor closed (as for the trajectory of Figure 3), and there are dense orbits that are non-uniformly distributed. In fact, when ℓ is Diophantine (so it is not too closely approximated by rationals), the set of directions $\theta \in [0, 2\pi)$ with dense but non-uniformly distributed trajectories is as large as possible, having Hausdorff dimension $1/2$ [Ch]. In the recent preprint [CHM2], the authors provide an arithmetic condition on the barrier length ℓ that characterizes the tables with a large set of minimal, non-uniquely ergodic directions. Minimal means that all (infinite) trajectories in the given direction are dense.

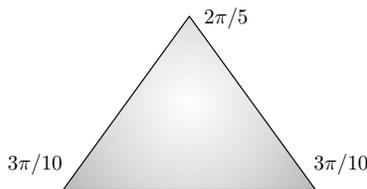


FIGURE 5. A triangular table with topologically optimal dynamics but not ergodically optimal; there are dense trajectories that are not uniformly distributed.

1.5. Topologically optimal billiard tables. A billiard table is said to be *topologically optimal* if for each direction θ ,

- (1) every trajectory that avoids the vertices is periodic; or
- (2) every trajectory that avoids the vertices is dense.

As with uniform distribution, we require that the trajectory be dense on the finite union of tables corresponding to different directions under reflection.

Ergodically optimal clearly implies topologically optimal. The following recent result shows that the converse is not true:

Theorem 1.3. [CHM1] *The billiard dynamics on the isosceles triangle with angles $(2\pi/5, 3\pi/10, 3\pi/10)$ are topologically optimal but not ergodically optimal: for each direction, either all infinite trajectories are closed or all are dense, but there exist trajectories which are dense but not uniformly distributed.*

See Figure 5. In standard dynamical language, the billiard flow in each direction is either completely periodic or minimal, while certain minimal directions are not uniquely ergodic. In fact, this is the only known billiard table to be topologically optimal but not ergodically optimal. More will be said about this example in §4.3.

2. THE TRANSLATION SURFACE OF A BILLIARD TABLE

In this section we describe the process of *unfolding*, i.e. passing from a polygonal billiard table to a Riemann surface equipped with a holomorphic 1-form. In this way, a billiard trajectory which reflects off the sides of the table unfolds into a straight line on the surface. We also define the symmetry group $SL(X, \omega)$ associated to any translation surface. We extend the notions of ergodically optimal and topologically optimal (from Section 1) to general translation surfaces. Finally, we introduce the notion of a lattice surface and compare it to an ergodically optimal surface.

2.1. Translation structure on a Riemann surface. By definition, a holomorphic 1-form ω on a Riemann surface X can be expressed in local coordinates as

$$\omega = f(z) dz$$

for some holomorphic function f . Any such pair (X, ω) with $\omega \not\equiv 0$ is called a *translation surface*. The 1-form ω gives a natural way to choose local Euclidean coordinates on X away from the zeroes of ω . Namely, for any point $z_0 \in X$ with $\omega(z_0) \neq 0$, we can integrate $\omega = f(z) dz$ to define a coordinate chart near z_0 by

$$(2.1) \quad \varphi(z) = \int_{z_0}^z f(z) dz$$

which is locally invertible and locally independent of path. In fact, the transition functions for these coordinate charts are given by translations, which explains the term “translation” surface. The Euclidean charts induce a flat metric on the surface, away from the zeroes of ω , and the geodesics in this metric are simply the straight lines in these coordinates. The charts glue up at the zeroes of ω to form cone-like singularities of the metric structure. The total angle at a zero of order k will be $2(k+1)\pi$.

2.2. Unfolding a billiard table into a translation surface. Fix a polygon T in \mathbb{R}^2 and assume that all of its angles are rational multiples of π . Let $G \subset O_2(\mathbb{R})$ be the group generated by reflections in the sides of T . Because of the rational angles, the group G is finite; let $N = |G|$. If the interior angles of T are expressed as $m_i\pi/n_i$, where the integers m_i and n_i have no common factors, then the number N is equal to twice the least common multiple of the n_i .

Take N copies of T , one for each reflected image gT with $g \in G$. Glue edges of distinct copies according to the reflection rules: if $h \in G$ is represented by reflection across an edge e of gT , then e is glued to its image in hgT . The genus of the resulting compact surface X is given by the formula

$$(2.2) \quad g(X) = 1 + \frac{N}{4} \left(k - 2 - \sum_{i=1}^k \frac{1}{n_i} \right)$$

where k is the number of vertices of T ; see [MT].

It is easy to see that the unit square unfolds into a torus, as does an equilateral triangle. For the regular pentagon depicted in Figure 1, the reflection group has 10 elements, and the table unfolds into a surface of genus 6. On the other hand, the $(2\pi/5, 3\pi/10, 3\pi/10)$ triangle shown in Figure 5 tiles the regular pentagon, but it unfolds into a surface of genus 4.

The Euclidean coordinates on the polygon T induce a flat conformal structure on the resulting surface, together with a finite collection of cone-point singularities (where the total angle at a point exceeds 2π). This structure can be recorded by the holomorphic 1-form dz on T , glued up to define a 1-form ω on the unfolded surface X . The cone points are simply the zeroes of ω . The pair (X, ω) defines the translation surface associated to the table T .

2.3. Surfaces from tables, but not vice versa. Every translation surface (X, ω) with X compact can be obtained by gluing a finite collection of polygons, via translations in the plane, taking the 1-form dz on the each of the polygons; see the discussion in [Ma2]. In general, though, the polygons will not be reflections of a single polygonal shape T . In any given genus, the subset of billiard table surfaces within the space of all translation surfaces (X, ω) forms a small (measure 0) subset.

2.4. Geodesic flow on a translation surface. The notion of optimal dynamics can be defined on general translation surfaces (X, ω) , in the context of the geodesic flow. Specifically, given a point $x \in X$ with $\omega(x) \neq 0$ and a vector v of unit length in the tangent space at x , the image of x under the time t geodesic flow is the unique point at distance t from x in direction v . The flow is only well-defined where straight lines avoid the zeroes of ω . We say a translation surface (X, ω) is *ergodically optimal* if its geodesics satisfy the dichotomy of §1.2. We say the surface is *topologically optimal* if its geodesics satisfy the weaker dichotomy of §1.5.

The idea of unfolding seems to have first appeared in [KZ], where the authors studied topological transitivity of the geodesic flow on the associated translation surface, concluding that most directions on a table are minimal (so that all trajectories avoiding the corners of the table are dense). General results about differentials on closed surfaces imply that there exist periodic trajectories in a dense set of directions; see [MT]. On the other hand, it is known that almost every direction gives rise to uniformly distributed geodesics on a translation surface [KMS].

2.5. The affine group of symmetries and its linear part. Fix a compact translation surface (X, ω) . An orientation-preserving diffeomorphism $h : X \rightarrow X$ is an *affine automorphism* if it is affine in the local Euclidean coordinates from ω . Because h must preserve area and orientation, its linear part (i.e. its derivative Dh) is a 2 by 2 matrix of determinant 1. We let $\text{SL}(X, \omega) \subset \text{SL}_2\mathbb{R}$ denote the group of linear parts, formed from all affine automorphisms of (X, ω) . The group $\text{SL}(X, \omega)$ is always a discrete subgroup of $\text{SL}_2\mathbb{R}$.

2.6. Lattice polygons and surfaces. It can happen that $\text{SL}(X, \omega)$ forms a lattice in $\text{SL}_2\mathbb{R}$, meaning that the quotient $\text{SL}_2\mathbb{R}/\text{SL}(X, \omega)$ has finite volume. In this case, we say the surface (X, ω) is a *lattice surface*. When a billiard table T unfolds into a lattice surface, we say that T is a *lattice polygon*. As an example, the group $\text{SL}(X, \omega)$ for the unfolding of the square table is the lattice $\text{SL}_2\mathbb{Z}$, consisting of all 2×2 matrices with integer entries and determinant 1. In fact, the square-tiled tables of Theorem 1.1 are all lattice polygons. In general, we have:

Theorem 2.1. [Ve] *If $\text{SL}(X, \omega) \subset \text{SL}_2\mathbb{R}$ is a lattice, then (X, ω) is ergodically optimal.*

In §3.4, we will sketch the proof of this result. By computing the group $\mathrm{SL}(X, \omega)$ explicitly for the translation surface associated to a regular n -gon billiard table, Veech used this Theorem 2.1 to deduce Theorem 1.2.

Only recently, Smillie and Weiss have shown that the converse of Theorem 2.1 does not hold:

Theorem 2.2. [SW] *There exist translation surfaces (X, ω) which are ergodically optimal but for which $\mathrm{SL}(X, \omega)$ is not a lattice.*

More will be said about this theorem in §4.2.

In the following section, we will explore some of the ideas which enter into the search for lattice surfaces and their relation to the Teichmüller geometry of moduli space.

3. AN $\mathrm{SL}_2\mathbb{R}$ ACTION ON MODULI SPACE

In this section, we describe a linear stretching action on the space of translation surfaces. We introduce Teichmüller curves, algebraic curves in the moduli space \mathcal{M}_g of genus g compact Riemann surfaces which are isometrically immersed with respect to the Teichmüller metric. The translation surfaces (X, ω) which generate Teichmüller curves are precisely the lattice surfaces, introduced in §2.6. We describe McMullen's classification of Teichmüller curves in genus 2, and we mention a few results about the search for Teichmüller curves in higher genus. We conclude with a description of the main goal in this direction of study: a complete analysis of orbits and invariant measures for the action of $\mathrm{SL}_2\mathbb{R}$ on the moduli space $\Omega\mathcal{M}_g$.

3.1. Stretching the Euclidean coordinates. Fix a translation surface (X, ω) . We can define a new translation surface $A \cdot (X, \omega)$ for each $A \in \mathrm{SL}_2\mathbb{R}$ by composing the natural Euclidean charts for ω with A . Specifically, recall that charts for X are defined by the φ of equation (2.1) away from the zeroes of ω . We replace each such φ with $A \circ \varphi$. These charts $\{A \circ \varphi\}$ glue to form a new Riemann surface Y , homeomorphic to X , and the 1-form dz in the plane defines a holomorphic 1-form η on Y . We set $A \cdot (X, \omega) = (Y, \eta)$.

In the context of billiard tables, the action may be viewed as follows. Apply the matrix A to the billiard table T and each of its reflected copies to obtain a new collection of polygons. Maintain the identification of edges that form the unfolded surface for T , and glue corresponding pairs of edges of the stretched polygons by translations. See Figure 6.

Recall that every translation surface (X, ω) can be represented by a finite set of polygons $\{P_i\}$ in the plane, with edges identified by translations. The new surface $A \cdot (X, \omega)$ can be viewed as the set of polygons $\{A(P_i)\}$ with edges identified according to the same combinatorial rules.

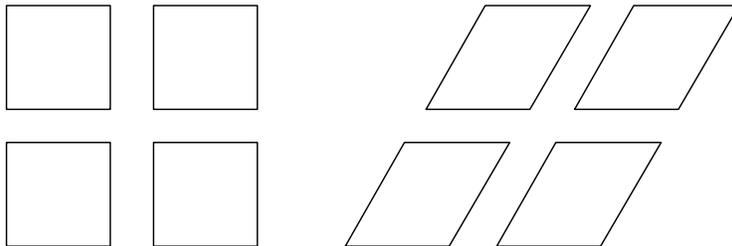


FIGURE 6. At left, the square billiard table with its reflected copies which glue to form the square torus; at right, their images under a linear shear, which glue to form a torus that does not generally come from a billiard table.

3.2. The moduli space of translation surfaces. We let \mathcal{M}_g denote the moduli space of compact Riemann surfaces of genus $g \geq 1$. The vector spaces of holomorphic 1-forms on the Riemann surfaces $X \in \mathcal{M}_g$ form a rank g bundle over the moduli space \mathcal{M}_g . The *moduli space of translation surfaces* is the subset $\Omega\mathcal{M}_g$ of nonzero holomorphic 1-forms, the bundle minus its zero section over \mathcal{M}_g . The linear stretching of translation surfaces defines a continuous action of $\mathrm{SL}_2\mathbb{R}$ on all of $\Omega\mathcal{M}_g$.

The stabilizer of a point (X, ω) under the $\mathrm{SL}_2\mathbb{R}$ action coincides with the group $\mathrm{SL}(X, \omega)$ defined in §2.5. Indeed, an element $A \in \mathrm{SL}_2\mathbb{R}$ lies in the stabilizer of a translation surface (X, ω) if the stretched surface $A \cdot (X, \omega)$ is isomorphic to (X, ω) . Specifically, the underlying surfaces are conformally isomorphic via a biholomorphism which pulls back the 1-form. This can be seen in terms of the polygons, via a certain cut and paste transformation. If the polygons forming (X, ω) can be cut into smaller polygons and reglued (without violating any of the initial gluing rules) to form the translation surface $A \cdot (X, \omega)$, then the surfaces are isomorphic.

It is important to observe that the $\mathrm{SL}_2\mathbb{R}$ action does not preserve the subset of surfaces coming from billiard tables in any genus. In genus 1, only two tori in \mathcal{M}_1 can be the unfolded surface of a billiard table; see §3.5. However, every point in $\Omega\mathcal{M}_1$ lies in the orbit of a billiard table translation surface. For each genus $g \geq 2$, the union of all $\mathrm{SL}_2\mathbb{R}$ -orbits containing billiard table surfaces forms a measure 0 subset of $\Omega\mathcal{M}_g$. This can be seen by observing that the stabilizer $\mathrm{SL}(X, \omega)$ is always nontrivial for a billiard table surface, while the stabilizer $\mathrm{SL}(X, \omega)$ is generally trivial in $\Omega\mathcal{M}_g$ (except in genus 2 where it will contain $\{\pm I\}$ because of the hyperelliptic involution); see [Mö2].

3.3. Teichmüller geometry and Teichmüller curves. The diagonal action of $\mathrm{SL}_2\mathbb{R}$, the action of the 1-parameter family of matrices

$$g_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$$

on the space of translation surfaces $\Omega\mathcal{M}_g$, plays a key role. It parametrizes the *Teichmüller geodesic flow* on the moduli space \mathcal{M}_g . Via these flow lines, a 1-form ω on X gives rise to a tangent vector to $X \in \mathcal{M}_g$; its norm in the Teichmüller metric is the integral

$$\int_X |\omega|^2.$$

See [Hu, Chapter 6] for background on Teichmüller spaces and the Teichmüller metric.

Recall that the quotient $\mathrm{SO}_2\mathbb{R}\backslash\mathrm{SL}_2\mathbb{R}$ can be naturally identified with the hyperbolic plane \mathbb{H} . For a given translation surface (X, ω) , the rotations preserve the underlying surface X , and the orbit

$$\mathrm{SL}_2\mathbb{R} \ni A \mapsto A \cdot (X, \omega) \in \Omega\mathcal{M}_g$$

descends to a holomorphic map

$$\mathbb{H} \rightarrow \mathcal{M}_g$$

which is a local isometry with respect to the Poincaré metric on the upper half-plane \mathbb{H} and the Teichmüller metric on \mathcal{M}_g ; see e.g. [KMS], [Ve], [Mc1]. When the stabilizer $\mathrm{SL}(X, \omega)$ forms a lattice, the image of \mathbb{H} forms an algebraic curve in \mathcal{M}_g . Such curves are called *Teichmüller curves*; these are precisely the isometrically embedded algebraic curves in \mathcal{M}_g . Recall that each Teichmüller curve gives rise to a family of ergodically optimal translation surfaces, by Theorem 2.1. Using this lattice condition to locate ergodically optimal surfaces, we see how an understanding of Teichmüller geometry of \mathcal{M}_g , in addition to the properties of this $\mathrm{SL}_2\mathbb{R}$ action, might help us classify such surfaces and ergodically optimal billiard tables.

3.4. Masur’s theorem and the Veech dichotomy. In Theorem 2.1, we have seen that any surface with $\mathrm{SL}(X, \omega)$ a lattice will be ergodically optimal. The proof of this result follows from a theorem of Masur:

Theorem 3.1. [Ma1] *If the orbit $g_t \cdot (X, \omega)$ is recurrent in \mathcal{M}_g , then the vertical direction on (X, ω) is uniquely ergodic.*

Recall that a direction is uniquely ergodic if it satisfies condition (2) in the ergodically optimal dichotomy of §1.2. The vertical trajectories are those in direction $\theta = \pi/2$ with respect to the natural Euclidean coordinates on (X, ω) . As described in §3.3, the image of $g_t \cdot (X, \omega)$ in \mathcal{M}_g is a geodesic in the Teichmüller metric.

Any geodesic on a Teichmüller curve is either recurrent or it heads out a cusp. In the first case, the direction is uniquely ergodic by Theorem 3.1. Veech’s Theorem 2.1 follows by analyzing the case where a geodesic goes out the cusp. In that case, the geodesic $g_t \cdot (X, \omega)$ must be orthogonal to a closed horocycle on the Teichmüller curve. There is then a parabolic element of $\mathrm{SL}(X, \omega)$ which preserves the vertical foliation, acting as the identity on leaves connecting zeroes. These closed leaves decompose the surface (X, ω) into cylinders. See details in [Ma2].

3.5. Genus 1. There is a unique holomorphic 1-form (up to scaling) on a torus, coming from the form dz on \mathbb{C} , when representing the torus as the quotient of \mathbb{C} by a lattice. The translation structure from dz is just the usual flat metric from the plane with no singularities. It is well-known that geodesics on a flat torus satisfy the optimal dichotomy described in §1.2. As there are no zeroes of the 1-form, every geodesic is infinite; so in a given direction, all are closed or all are dense. Thus any billiard table which unfolds into a genus 1 translation surface must be ergodically optimal. In fact, using the genus formula (2.2), we see that there are only 4 such tables: the three triangles with angles $(\pi/3, \pi/3, \pi/3)$, $(\pi/2, \pi/4, \pi/4)$, $(\pi/2, \pi/3, \pi/6)$, and the square. These unfold into the tori from the square lattice and the hexagonal lattice in \mathbb{R}^2 . There is a unique Teichmüller curve in genus 1 (because \mathcal{M}_1 is one-dimensional), and any (X, ω) in $\Omega\mathcal{M}_1$ has ergodically optimal geodesic flow.

3.6. Genus 2 and geometric primitivity. The story in genus 2 is significantly richer, and a complete characterization of ergodically optimal billiard tables has involved a sophisticated understanding of the $\mathrm{SL}_2\mathbb{R}$ action on $\Omega\mathcal{M}_2$.

First note that it is fairly easy to construct Teichmüller curves in genus > 1 by simply passing to branched covers of tori, branched only over torsion points; see [GJ]. It is more difficult to construct genuinely new Teichmüller curves, not inherited from a lower genus.

A Teichmüller curve is said to be *geometrically primitive* if it does not arise from a branched cover construction from a Teichmüller curve in a moduli space of lower genus. Precisely, suppose (X, ω) generates a Teichmüller curve in genus g , so $\mathrm{SL}(X, \omega)$ is a lattice, and let $S \subset X$ be a finite set which is invariant under $\mathrm{SL}(X, \omega)$. For any branched cover $f : Y \rightarrow X$ branched only over S , the group $\mathrm{SL}(Y, f^*\omega)$ will also form a lattice in $\mathrm{SL}_2\mathbb{R}$. The associated Teichmüller curve, generated by the orbit of $(Y, f^*\omega)$, is then said to be commensurable with the given curve generated by (X, ω) . A Teichmüller curve is geometrically primitive if it lies in the moduli space of minimal genus, with respect to all commensurable Teichmüller curves. See [Mc1, §3].

For some time, only finitely many geometrically primitive Teichmüller curves were known in each genus. The situation changed in 2002 when McMullen showed that there are infinitely many geometrically primitive Teichmüller curves in genus 2 [Mc2].

A key new idea was to bring the endomorphism ring of the Jacobian of X into play. Using this perspective, for each integer $D > 4$, with $D = 0$ or $1 \pmod{4}$, McMullen defined the *Weierstrass curve* $W_D \subset \mathcal{M}_2$ as the locus where the Jacobian $\mathrm{Jac}(X)$ has special properties (real multiplication by the ring O_D and an eigenform with a double zero). The curve W_D is irreducible unless $D = 1 \pmod{8}$ and $D > 9$, in which case it has two components, distinguished by a so-called *spin invariant* in $\mathbb{Z}/2\mathbb{Z}$. The classification of primitive Teichmüller curves in genus 2 was then completed by the following results.

Theorem 3.2. [Mc2] *Every component of W_D is a Teichmüller curve, which is geometrically primitive provided D is not a square.*

Theorem 3.3. [Mc5] *The only other geometrically primitive Teichmüller curve in \mathcal{M}_2 comes from billiards in the triangle $(\pi/2, 2\pi/5, \pi/10)$. It gives the unique example generated by a form ω with two simple zeroes.*

The Weierstrass curves account for all Teichmüller curves in genus 2 generated by forms with double zeros. For example, the curves W_5 and W_8 come from billiards in the triangle $(\pi/2, \pi/5, 3\pi/10)$ (which unfolds into two copies of a regular pentagon) and the triangle $(\pi/2, \pi/8, 3\pi/8)$ (which unfolds into a regular octagon), respectively. Theorem 3.3 shows it is much harder to obtain primitive Teichmüller curves from forms with simple zeros. Note that the triangle $(\pi/2, 2\pi/5, \pi/10)$ unfolds into a regular decagon with opposite sides identified.

3.7. Higher genus and algebraic primitivity. For a lattice surface (X, ω) , its *trace field* is the field extension of \mathbb{Q} generated by adjoining elements $\{\text{tr } \gamma : \gamma \in \text{SL}(X, \omega)\}$. The degree of this extension over \mathbb{Q} is always bounded by the genus of X ; see the discussion in [Mö1]. A Teichmüller curve is *algebraically primitive* if the trace field has degree over \mathbb{Q} equal to the genus. Algebraically primitive implies geometrically primitive, but the converse only holds in genus 1 and 2. In [Mö1, §2], Möller describes an example of Hubert and Schmidt from [HS1] in genus 3, from unfolding the triangular billiard table with angles $(\pi/4, \pi/3, 5\pi/12)$, which is geometrically primitive but not algebraically primitive.

Similar to the result of Theorem 3.3, Bainbridge and Möller have shown recently:

Theorem 3.4. [BM1] *In genus 3, there are only finitely many algebraically primitive Teichmüller curves generated by a form in the stratum $\Omega\mathcal{M}_3(3, 1)$.*

The stratum $\Omega\mathcal{M}_3(3, 1)$ of $\Omega\mathcal{M}_3$ is the subset of translation surfaces (X, ω) where ω has exactly two zeroes, one of multiplicity 3 and the other of multiplicity 1.

There is one known example from the stratum $\Omega\mathcal{M}_3(3, 1)$, due to Kenyon and Smillie, obtained by unfolding the $(2\pi/9, \pi/3, 4\pi/9)$ triangle; see Theorem 5.4 below. In [BM1], the authors also describe a computer algorithm for searching the stratum $\Omega\mathcal{M}_3(4)$ of forms with one zero of multiplicity 4 for algebraically primitive Teichmüller curves. The Teichmüller curve associated to the right triangle $(\pi/2, \pi/7, 5\pi/14)$ is algebraically primitive and lies in this stratum; it may be the only one. Bainbridge and Möller conjecture that there are only finitely many algebraically primitive Teichmüller curves in all of \mathcal{M}_3 .

On the other hand, McMullen found an infinite collection of Teichmüller curves with quadratic trace field:

Theorem 3.5. [Mc4] *There are infinitely many geometrically primitive Teichmüller curves in genus 3 and 4.*

Examples of billiard tables generating these curves include the triangles $(\pi/4, \pi/8, 5\pi/8)$ in genus 3 and $(\pi/5, \pi/10, 7\pi/10)$ in genus 4 [Wa]. It is unknown whether infinitely many geometrically primitive Teichmüller curves can exist in any genus > 4 .

3.8. Invariant measures and orbit closures. The Teichmüller curve is a special type of $\mathrm{SL}_2\mathbb{R}$ orbit on $\Omega\mathcal{M}_g$, projected to \mathcal{M}_g . The image is closed with finite volume. In fact, McMullen gave a complete description of the orbit closures and invariant measures for the action of $\mathrm{SL}_2\mathbb{R}$ on $\Omega\mathcal{M}_g$ in genus $g = 2$:

Theorem 3.6. [Mc6] *Every orbit closure and every ergodic invariant measure for the action of $\mathrm{SL}_2\mathbb{R}$ on $\Omega\mathcal{M}_2$ is algebraic.*

Specifically, the projection of an orbit closure is an algebraic variety. It is either all of \mathcal{M}_2 , a Hilbert modular surface, or a Teichmüller curve. A similar description can be given of the invariant measures. To put these results in context, recall that Ratner's celebrated results imply that any $\mathrm{SL}_2\mathbb{R}$ orbit closure in a homogenous space is algebraic [Ra].

In higher genus, $g > 2$, it remains a challenge to

- (1) show all orbit closures are algebraic, and
- (2) classify the orbit closures by discrete invariants.

These problems are open, even in genus 3.

4. TOWARDS A CHARACTERIZATION OF OPTIMAL DYNAMICS OR THE LATTICE CONDITION

In this section, we return to billiards and the main theme of this article. Recall that we have introduced three types of optimal translation surfaces: lattice surfaces, ergodically optimal surfaces, and topologically optimal surfaces. Figure 7 shows the relative inclusions of these three notions. For genus 2 translation surfaces, we shall see that the three notions coincide. However, we have already seen that the equivalence fails in general (Theorems 1.3 and 2.2). In this section, we discuss some of the ideas which explain the failure of equivalence in higher genus, and we describe a possible characterization of the lattice condition.

4.1. Characterization in genus 2. Theorem 2.1 of Veech states that any lattice translation surface is ergodically optimal. McMullen provided the following converse to Theorem 2.1:

Theorem 4.1. [Mc3] *If X has genus 2 and $\mathrm{SL}(X, \omega)$ is not a lattice in $\mathrm{SL}_2\mathbb{R}$, then there exist geodesics on (X, ω) which are neither dense nor closed.*

In particular, any billiard table T which unfolds into a surface of genus 2 with non-lattice group $\mathrm{SL}(X, \omega)$ will have a trajectory similar to the trajectory in Figure 3.

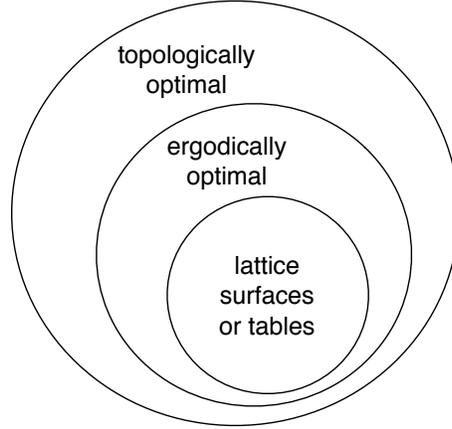


FIGURE 7. A diagram of inclusions for billiard tables or translation surfaces.

Calta gave a different proof for the case when ω has a double zero; see [Ca]. Consequently, we have:

Corollary 4.2. *If X has genus 2, then the following are equivalent:*

- (1) *the translation surface (X, ω) is ergodically optimal;*
- (2) *the translation surface (X, ω) is topologically optimal; and*
- (3) *the group $\mathrm{SL}(X, \omega)$ is a lattice in $\mathrm{SL}_2\mathbb{R}$.*

Building on the work of McMullen, Cheung and Masur recently proved a converse of a different sort to Theorem 2.1:

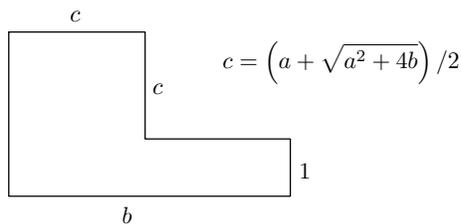
Theorem 4.3. [CM] *If X has genus 2, and $\mathrm{SL}(X, \omega)$ is not a lattice in $\mathrm{SL}_2\mathbb{R}$, then there exist geodesics on (X, ω) which are dense but non-uniformly distributed.*

In particular, any billiard table T which unfolds into a surface of genus 2 with non-lattice group $\mathrm{SL}(X, \omega)$ has dynamical features similar to the rectangular table with barrier of Figure 4, for irrational ℓ .

Using his classification of Teichmüller curves (see §3.6), McMullen obtains the following classification of lattice polygons. For any pair of integers a and b with $b > 0$, the billiard table $L(a, b)$ is shown in Figure 8. Two tables are *equivalent* if their unfolded surfaces lie in the same $\mathrm{SL}_2\mathbb{R}$ orbit.

Theorem 4.4. [Mc5] *Let T be a billiard table which unfolds into a translation surface (X, ω) of genus 2. Then T is a lattice polygon if and only if it is equivalent to*

- (1) *a table tiled by congruent triangles of angles $(\pi/2, \pi/3, \pi/6)$ or $(\pi/2, \pi/4, \pi/4)$;*
- (2) *an L -shaped table $L(a, b)$ for some $a, b \in \mathbb{Z}$; or*
- (3) *the triangle $(\pi/2, 2\pi/5, \pi/10)$.*

FIGURE 8. The L -shaped table $L(a, b)$.

The triangle $(\pi/2, 2\pi/5, \pi/10)$ unfolds into a decagon with opposite sides identified. The corresponding 1-form has two simple zeroes, and it generates the unique such Teichmüller curve as stated in Theorem 3.3.

4.2. Coverings and branched coverings. When a translation surface (Y, η) is a covering of another translation surface (X, ω) , so that there is a covering map $p : Y \rightarrow X$ such that $\eta = p^*\omega$, then the groups $\mathrm{SL}(X, \omega)$ and $\mathrm{SL}(Y, \eta)$ are commensurable [GJ]; that is, after conjugating in $\mathrm{SL}_2\mathbb{R}$, the two groups share a finite-index subgroup. We have also seen that commensurability holds when branching only over points with finite $\mathrm{SL}(X, \omega)$ orbit in X . Gutkin and Judge used this covering property to characterize the translation surfaces with $\mathrm{SL}(X, \omega)$ commensurable to $\mathrm{SL}_2\mathbb{Z}$ (see Theorem 5.1 below).

For more general branched covers, the structure of $\mathrm{SL}(X, \omega)$ is not necessarily preserved. Hubert and Schmidt considered branched covers of lattice surfaces (X, ω) branched over points of a special type. A point p in X is *periodic* if its orbit under the group $\mathrm{SL}(X, \omega)$ is finite in X . A point p is a *connection point* if any geodesic from a zero of ω through the point p again encounters a zero of ω . Hubert and Schmidt showed [HS2]:

- (1) If p is a non-periodic connection point on a lattice surface (X, ω) , then the subgroup $\mathrm{SL}(X, \omega, p) := \{\gamma \in \mathrm{SL}(X, \omega) : \gamma(p) = p\}$ is infinitely generated; and
- (2) If a branched cover $(Y, \eta) \rightarrow (X, \omega)$ is branched only over a connection point p , and if $\mathrm{SL}(X, \omega)$ is a lattice, then $\mathrm{SL}(Y, \eta)$ is commensurable with $\mathrm{SL}(X, \omega, p)$, and the surface (Y, η) is topologically optimal.

They showed further that the second statement holds when branching over more than one connection point if the base surface (X, ω) has a property they call *strong holonomy type*.

The Smillie-Weiss examples in Theorem 2.2 rely on the covering construction of Hubert and Schmidt in [HS2]. In certain cases, the dynamical properties of the geodesic flow on the surface (X, ω) are preserved when passing to the branched cover. Smillie and Weiss concentrated on branched covers with a single ramification point in the base translation space of genus $g \geq 2$. They cleverly combined two facts: one

is the simple observation that the forgetful map from $\mathcal{M}_{g,1}$ down to \mathcal{M}_g has compact fibers. The second is Masur's Theorem 3.1.

4.3. The triangle of Theorem 1.3. Recall from Theorem 1.3 that the triangular billiard table with angles $(2\pi/5, 3\pi/10, 3\pi/10)$ is topologically optimal but not ergodically optimal. In [CHM1], where Theorem 1.3 is proved, the authors again use one of the Hubert-Schmidt branched covers, but branched over *two* points. The special example of Theorem 1.3 arises in the following way. Begin with the translation surface of the triangle with angles $(\pi/2, \pi/5, 3\pi/10)$. This triangle unfolds into two reflected copies of the regular pentagon, forming a surface of genus 2. It is a lattice polygon and therefore ergodically optimal [Ve]; in fact, the triangle is equivalent to one of the L -shaped tables in McMullen's classification Theorem 4.4 [Mc1, §9]. The centers of the two pentagons are connection points (see definition in §4.2). Taking a double cover of this genus 2 surface, branched over the two centers, produces the translation surface of genus 4 which is the unfolding of the triangle $(2\pi/5, 3\pi/10, 3\pi/10)$. The new surface is topologically optimal by the arguments of [HS2], but Cheung, Hubert, and Masur show that it has non-uniformly distributed, dense geodesics.

4.4. Possible characterization of the lattice condition: rational ratios of moduli. The schematic of Figure 7 indicates the relative inclusions of translation surfaces (or billiard tables) which have lattice group $\mathrm{SL}(X, \omega)$, those which are ergodically optimal, and those which are topologically optimal. By McMullen's theorem (Corollary 4.2), the sets coincide for translation surfaces of genus 2. The examples of Smillie and Weiss (Theorem 2.2) and of Cheung, Hubert, and Masur (Theorem 1.3) show that the containments are strict in the setting of translation surfaces of arbitrary genus. Note, however, that Theorem 2.2 leaves open the existence of ergodically optimal billiard tables which unfold into surfaces with non-lattice $\mathrm{SL}(X, \omega)$, but a search is under way.

Further investigations are also in progress about possible characterizations of the lattice condition, because it is the geometry of the $\mathrm{SL}_2\mathbb{R}$ action on the moduli space of translation surfaces which drives most of the interest in this class of billiards. In particular, there are many consequences of the lattice condition, other than ergodically optimal geodesic flow, described in Veech's article [Ve].

For example, suppose we fix a direction θ in which all geodesics on (X, ω) are closed. The periodic trajectories decompose X into a finite collection of open cylinders, foliated by the closed geodesics. Recall that the *modulus* of a Euclidean cylinder is its height over its circumference. When $\mathrm{SL}(X, \omega)$ forms a lattice in $\mathrm{SL}_2\mathbb{R}$, the ratio of moduli for any pair of these cylinders is rational. This is because each periodic direction on (X, ω) corresponds to a parabolic element of $\mathrm{SL}(X, \omega)$ which preserves the foliation, acting by a Dehn twist on the foliated cylinders. It is possible that

$\mathrm{SL}(X, \omega)$ is always a lattice when (X, ω) is ergodically optimal and the moduli of cylinders in periodic directions have rational ratios.

5. EXAMPLES OF GROUPS $\mathrm{SL}(X, \omega)$ AND THE SEARCH FOR OPTIMAL BILLIARD TABLES

In this final section, we include a collection of recent results describing groups that arise as $\mathrm{SL}(X, \omega)$ and some lists of billiard tables known to be lattice polygons. We conclude with Schwartz's recent result on periodic trajectories outside the setting of rational polygons, as an example of the new challenges that arise in general.

5.1. Arithmetic groups. Recall from Theorem 1.1 that any billiard table which is tiled by squares must be ergodically optimal. In fact, the unfolded surface (X, ω) is a branched cover of a square torus, branched over a single point; consequently the group $\mathrm{SL}(X, \omega)$ is commensurable to the lattice $\mathrm{SL}_2\mathbb{Z}$. Such subgroups of $\mathrm{SL}_2\mathbb{R}$ are called *arithmetic*. Gutkin and Judge provided a converse to this result, characterizing translation surfaces with arithmetic $\mathrm{SL}(X, \omega)$.

Theorem 5.1. [GJ] *The group $\mathrm{SL}(X, \omega)$ is commensurable with $\mathrm{SL}_2\mathbb{Z}$ if and only if the translation surface (X, ω) is tiled by a Euclidean parallelogram.*

Recently, Ellenberg and McReynolds took this result further, by addressing which arithmetic groups arise as the group $\mathrm{SL}(X, \omega)$ for some translation surface.

Theorem 5.2. [EMc] *Let Γ be any finite index subgroup of $\mathrm{SL}_2\mathbb{Z}$ which contains $\{\pm I\}$ and is contained in the level two congruence subgroup*

$$\Gamma(2) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{2} \right\}.$$

Then there is a compact translation surface (X, ω) with $\mathrm{SL}(X, \omega)$ isomorphic to Γ .

In their article, Ellenberg and McReynolds also show that any algebraic curve X defined over $\bar{\mathbb{Q}}$ is birationally equivalent to a Teichmüller curve. By Belyi's theorem, such an X is birationally equivalent to an unramified cover \hat{X} of $\Sigma = \mathbb{CP}^1 \setminus \{0, 1, \infty\}$. Therefore, $\Gamma = \pi_1(\hat{X})$ may be viewed as a subgroup of $\mathbb{P}\Gamma(2) = \Gamma(2)/\{\pm I\}$. The geometric idea of their argument is to realize \hat{X} as a Teichmüller curve, or, in algebro-geometric terms, a moduli space of covers of genus-1 algebraic curves branched at 1 point.

5.2. Triangle groups. The classical *triangle group*, depending on integer parameters $m, n, l \geq 2$, has the finite presentation

$$\Delta(m, n, l) = \langle a, b, c \mid a^2 = b^2 = c^2 = (ac)^m = (bc)^n = (ca)^l = 1 \rangle.$$

For different choices of the parameters, the group can be realized as the reflection group in the three sides of a Euclidean, spherical, or hyperbolic triangle.

For $l = \infty$, with $2 \leq m \leq n$ and $mn \geq 6$, the group

$$\Delta(m, n, \infty) = \langle a, b, c \mid a^2 = b^2 = c^2 = (ac)^m = (bc)^n = 1 \rangle$$

may be realized as a subgroup of $\mathrm{PGL}(2, \mathbb{R})$, by reflections in a hyperbolic triangle with one ideal vertex, a vertex of angle π/m and a vertex of angle π/n . Restricting to the elements that preserve orientation of the hyperbolic plane, we obtain the index two subgroup

$$\Delta^+(m, n, \infty) = \Delta(m, n, \infty) \cap \mathrm{PSL}(2, \mathbb{R})$$

forming a lattice that we will call a *Fuchsian triangle group*. Among the Fuchsian triangle groups, only the pairs $(m, n) \in \{(2, 3), (2, 4), (2, 6), (3, 3), (4, 4), (6, 6)\}$ are arithmetic [Ta].

The first non-arithmetic lattice groups $\mathrm{SL}(X, \omega)$ associated to billiard tables were provided by Veech, in his seminal 1989 article [Ve]. He showed that for each $n \geq 3$, the Fuchsian triangle group $\Delta^+(2, n, \infty)$ (or an index 2 subgroup, depending on the parity of n) is the group $\mathrm{PSL}(X, \omega)$ associated to billiards in the isosceles triangle $(\pi/n, \pi/n, (n-2)\pi/n)$. Here, $\mathrm{PSL}(X, \omega)$ is the image of $\mathrm{SL}(X, \omega)$ under the projection $\mathrm{SL}_2\mathbb{R} \rightarrow \mathrm{PSL}_2\mathbb{R} = \mathrm{SL}_2\mathbb{R}/\{\pm I\}$. In particular, he showed that these triangular billiard tables are ergodically optimal.

Building on Veech's work, Ward studied billiards in the triangle $(\pi/2n, \pi/n, (2n-3)\pi/2n)$. He concluded that for all $n \geq 4$, this triangular table is a lattice polygon. The associated group is $\mathrm{PSL}(X, \omega) = \Delta^+(3, n, \infty)$ [Wa].

By defining explicit Teichmüller curves in the moduli space, Bouw and Möller provided the following generalization of Veech and Ward's constructions:

Theorem 5.3. [BM2] *For each pair of integers (m, n) , with $1 < m < n < \infty$ and m and n not both even, there exists a translation surface (X, ω) with $\mathrm{PSL}(X, \omega) = \Delta^+(m, n, \infty)$.*

For the case $m = n$, the construction of Bouw and Möller yields a translation surface with $\mathrm{PSL}(X, \omega)$ isomorphic to either $\Delta^+(m, m, \infty)$ or $\Delta^+(2, m, \infty)$. The ambiguity comes from the inclusion $\Delta^+(m, m, \infty) \subset \Delta^+(2, m, \infty)$. For the cases $2 \leq m \leq 5$ and $m < n$ not both even, Bouw and Möller show that their examples come from billiard tables: they construct a triangle or quadrilateral for each such pair (m, n) . For $m = 2, 3$, they recover the triangles of Veech and Ward.

In [Ho2], Hooper gave an alternate construction of translation surfaces satisfying the conclusion of Theorem 5.3. He showed further that the triangle groups $\Delta^+(m, n, \infty)$ can *not* be contained in the group $\mathrm{PSL}(X, \omega)$ of any translation surface when either $\mathrm{gcd}(m, n) = 2$ or both m and n are even while $m/\mathrm{gcd}(m, n)$ and $n/\mathrm{gcd}(m, n)$ are both odd.

5.3. Infinitely generated groups. While the group $\mathrm{SL}(X, \omega)$ is known to always be discrete, it can happen that it is infinitely generated. The first such examples were

seen by Hubert and Schmidt using their branched cover construction; see §4.2. On the other hand, McMullen showed that $\mathrm{SL}(X, \omega)$ is often infinitely generated, at least in genus 2. He showed that when the 1-form ω has two simple zeroes, and if $\mathrm{SL}(X, \omega)$ contains an element with irrational trace, then either (X, ω) lies in the $\mathrm{SL}_2\mathbb{R}$ orbit of the triangle $(\pi/2, 2\pi/5, \pi/10)$ or $\mathrm{SL}(X, \omega)$ is infinitely generated.

5.4. Triangles and the continued search. While a great deal of progress has been made towards the classification of lattice polygons, we are far from a complete understanding, even in the restricted setting of triangles. We conclude the article with two interesting results about these seemingly simple three-sided billiard tables.

After Veech studied the isosceles triangles mentioned in §5.2, a search began for more examples of triangular lattice polygons. Kenyon and Smillie provided a striking breakthrough with the following result:

Theorem 5.4. [KS] *Let T be a triangle with all angles $\leq \pi/2$. Then T is a lattice polygon if and only if its angles are in the following list:*

- $(\pi/2, \pi/n, (n-2)\pi/2n)$ for $n \geq 4$
- $(\pi/n, (n-2)\pi/4n, (n-2)\pi/4n)$ for $n \geq 3$
- $(\pi/4, \pi/3, 5\pi/12)$
- $(\pi/5, \pi/3, 7\pi/15)$
- $(2\pi/9, \pi/3, 4\pi/9)$

In [KS], the theorem for acute non-isosceles triangles was stated with an extra hypothesis; the proof was completed in [Pu]. It has not yet been established exactly which acute or right triangles are ergodically optimal.

Finally, we mention that little is known once the billiard tables are allowed to have irrational angles. Recall that all polygonal billiard tables with angles equal to rational multiples of π have many periodic trajectories; see §2.4. By contrast, it is an open question whether every triangular billiard table, allowing arbitrary angles, has a periodic trajectory. In a recent article, Schwartz has shown:

Theorem 5.5. [Sch] *Let T be an obtuse triangle (with rational or irrational angles) with its largest angle no greater than 100 degrees. Then T has a stable periodic billiard trajectory.*

Stable means that all nearby triangles have a periodic trajectory of the same combinatorial type; see [Sch] and the references there. We may compare Theorem 5.5 to the fact that every right triangle has periodic trajectories, but Hooper has shown that these trajectories can never be stable [Ho1].

REFERENCES

- [BM1] M. Bainbridge and M. Möller. Deligne-Mumford compactification of the real multiplication locus and Teichmüller curves in genus three. Preprint, 2009.
- [BM2] I. Bouw and M. Möller. Teichmüller curves, triangle groups, and Lyapunov exponents. To appear, *Ann. of Math. (2)*.
- [Ca] K. Calta. Veech surfaces and complete periodicity in genus two. *J. Amer. Math. Soc.* **17**(2004), 871–908.
- [Ch] Y. Cheung. Hausdorff dimension of the set of nonergodic directions. *Ann. of Math. (2)* **158**(2003), 661–678. With an appendix by M. Boshernitzan.
- [CHM1] Y. Cheung, P. Hubert, and H. Masur. Topological dichotomy and strict ergodicity for translation surfaces. *Ergodic Theory Dynam. Systems* **28**(2008), 1729–1748.
- [CHM2] Y. Cheung, P. Hubert, and H. Masur. Dichotomy for the Hausdorff dimension of the set of nonergodic directions. Preprint, 2009.
- [CM] Y. Cheung and H. Masur. Minimal non-ergodic directions on genus-2 translation surfaces. *Ergodic Theory Dynam. Systems* **26**(2006), 341–351.
- [EMc] J. Ellenberg and D. B. McReynolds. Every curve is a Teichmüller curve. Preprint, 2009.
- [Gu] E. Gutkin. Billiards on almost integrable polyhedral surfaces. *Ergodic Theory Dynam. Systems* **4**(1984), 569–584.
- [GJ] E. Gutkin and C. Judge. Affine mappings of translation surfaces: geometry and arithmetic. *Duke Math. J.* **103**(2000), 191–213.
- [Ho1] W. P. Hooper. Periodic billiard paths in right triangles are unstable. *Geom. Dedicata* **125**(2007), 39–46.
- [Ho2] W. P. Hooper. Grid graphs and lattice surfaces. Preprint, 2009.
- [Hu] J. H. Hubbard. *Teichmüller theory and applications to geometry, topology, and dynamics. Vol. 1*. Matrix Editions, Ithaca, NY, 2006.
- [HS1] P. Hubert and T. A. Schmidt. Invariants of translation surfaces. *Ann. Inst. Fourier (Grenoble)* **51**(2001), 461–495.
- [HS2] P. Hubert and T. A. Schmidt. Infinitely generated Veech groups. *Duke Math. J.* **123**(2004), 49–69.
- [KZ] A. B. Katok and A. N. Zemljakov. Topological transitivity of billiards in polygons. *Mat. Zametki* **18**(1975), 291–300.
- [KS] R. Kenyon and J. Smillie. Billiards on rational-angled triangles. *Comment. Math. Helv.* **75**(2000), 65–108.
- [KMS] S. Kerckhoff, H. Masur, and J. Smillie. Ergodicity of billiard flows and quadratic differentials. *Ann. of Math. (2)* **124**(1986), 293–311.
- [Ma1] H. Masur. Hausdorff dimension of the set of nonergodic foliations of a quadratic differential. *Duke Math. J.* **66**(1992), 387–442.
- [Ma2] H. Masur. Ergodic theory of translation surfaces. In *Handbook of dynamical systems. Vol. 1B*, pages 527–547. Elsevier B. V., Amsterdam, 2006.
- [MT] H. Masur and S. Tabachnikov. Rational billiards and flat structures. In *Handbook of dynamical systems, Vol. 1A*, pages 1015–1089. North-Holland, Amsterdam, 2002.
- [Mc1] C. T. McMullen. Billiards and Teichmüller curves on Hilbert modular surfaces. *J. Amer. Math. Soc.* **16**(2003), 857–885.
- [Mc2] C. T. McMullen. Teichmüller curves in genus two: discriminant and spin. *Math. Ann.* **333**(2005), 87–130.

- [Mc3] C. T. McMullen. Teichmüller curves in genus two: the decagon and beyond. *J. Reine Angew. Math.* **582**(2005), 173–199.
- [Mc4] C. T. McMullen. Prym varieties and Teichmüller curves. *Duke Math. J.* **133**(2006), 569–590.
- [Mc5] C. T. McMullen. Teichmüller curves in genus two: torsion divisors and ratios of sines. *Invent. Math.* **165**(2006), 651–672.
- [Mc6] C. T. McMullen. Dynamics of $SL_2(\mathbb{R})$ over moduli space in genus two. *Ann. of Math. (2)* **165**(2007), 397–456.
- [Mö1] M. Möller. Periodic points on Veech surfaces and the Mordell-Weil group over a Teichmüller curve. *Invent. Math.* **165**(2006), 633–649.
- [Mö2] M. Möller. Affine groups of flat surfaces. In *Handbook of Teichmüller theory. Vol. II*, volume 13 of *IRMA Lect. Math. Theor. Phys.*, pages 369–387. Eur. Math. Soc., Zürich, 2009.
- [Pu] J.-C. Puchta. On triangular billiards. *Comment. Math. Helv.* **76**(2001), 501–505.
- [Ra] M. Ratner. Interactions between ergodic theory, Lie groups, and number theory. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*, pages 157–182, Basel, 1995. Birkhäuser.
- [Sch] R. E. Schwartz. Obtuse triangular billiards. II. One hundred degrees worth of periodic trajectories. *Experiment. Math.* **18**(2009), 137–171.
- [SW] J. Smillie and B. Weiss. Veech’s dichotomy and the lattice property. *Ergodic Theory Dynam. Systems* **28**(2008), 1959–1972.
- [Ta] K. Takeuchi. Arithmetic triangle groups. *J. Math. Soc. Japan* **29**(1977), 91–106.
- [Ve] W. A. Veech. Teichmüller curves in moduli space, Eisenstein series and an application to triangular billiards. *Invent. Math.* **97**(1989), 553–583.
- [Wa] C. C. Ward. Calculation of Fuchsian groups associated to billiards in a rational triangle. *Ergodic Theory Dynam. Systems* **18**(1998), 1019–1042.

DEPARTMENT OF MATHEMATICS, STATISTICS, AND COMPUTER SCIENCE, UNIVERSITY OF ILLINOIS AT CHICAGO, DEMARCO@MATH.UIC.EDU