TOPIC PROPOSAL: TRANSLATION SURFACES AND TEICHMÜLLER THEORY

ALEX WRIGHT
DISCUSSED WITH ALEX ESKIN
JANUARY 2010.

1. Introduction

Pick a unit vector $v \in \mathbb{R}^2$, and consider the dynamical system on the flat torus $\mathbb{R}^2/\mathbb{Z}^2$ defined by $f_t(z) = z + tv$. This dynamical system is very well understood. For example, if $v \in \mathbb{Q}^2$, then all trajectories are closed, and if not, then all trajectories are dense. In fact much more is known.

In an attempt to generalize this example, we see immediately the problem that other closed surfaces do not possess flat structures. However, we may consider surfaces which are flat outside of a finite number of "cone type" singularities, around which the angle is an integer multiple of 2π. Such surfaces are called *translation surfaces*. For example, the regular octagon with opposite sides identified gives a translation surface of genus two with one singularity of order $8 \cdot \frac{3\pi}{4} = 6\pi$. On such a surface we can again ask about the dynamics of flow in a given direction.

Translation surfaces arise naturally in the theory of billiards. Other motivations come from physics, dynamics and Teichmüller theory.

Teichmüller theory considers the space of all complex (or hyperbolic, or conformal) structures on a closed surface, and quasi-conformal maps between such surfaces. It turns out that the moduli space of translation surfaces is a sphere bundle over the moduli space of complex structures on the underlying surface. Teichmüller space has a natural metric called the Teichmüller metric. Geodesic flow in this metric corresponds to a natural operation on translation surfaces which contracts the vertical direction and expands the horizontal direction. (Applying this operation to the regular polygon gives a very short and wide squashed octagon.) Certain very special complex geodesics in Teichmüller space correspond to translation surfaces with the sort of exceptionally nice dynamical properties we see on the flat torus.

The purpose of this topic is to define Teichmüller space and translation surfaces, give the basic theorems on these objects and some additional theorems on the dynamics of translation surfaces, and explore some of the connections between translation surfaces and Teichmüller theory.

2. Teichmüller theory

Definition. The *Teichmüller space* $\text{Teich}(S_g)$ of a closed surface S_g of genus g is the space of all marked hyperbolic structures on S_g, or equivalently (by uniformization), the space of all marked complex structures on S_g. More formally, $\text{Teich}(S_g)$ is the set of equivalence classes of pairs (X, ϕ), where $\phi : S \to X$ is a homeomorphism and X is a hyperbolic surface. Two pairs (X, ϕ) and (X', ϕ') are considered to be equivalent if there is an isometry $I : X \to X'$ such that $I \circ \phi$ is isotopic to ϕ'.

Abelian and quadratic differentials. An abelian differential a differential one form. Locally it looks like $f(z)dz$. A quadratic differential locally looks like $f(z)(dz)^2$. In a different coordinate w it may be expressed as $f(z) \cdot \left(\frac{dz}{dw}\right)^2 (dw)^2$. All our differentials will be holomorphic, which means that z must be a complex coordinate and f must be holomorphic.

Pants and the topology on Teichmüller space. By cutting a hyperbolic surface along a fixed system of $3g - 3$ non-intersecting geodesics, we obtain a decomposition into $2g - 2$ pairs of pants, or spheres with three boundary components. To get a new hyperbolic structure we could change the lengths of the boundary components, or twist one of a pair of boundary circles before gluing them back together. In fact, given a fixed pants decomposition every marked hyperbolic structure can be obtained in exactly one such way. So, to know a hyperbolic structure, we just need to keep track of the lengths of the boundary components (the *length parameters*), and the angle we twist by before gluing (the *twist parameters*). Teichmüller space is given the weak topology defined by these $6g-6$ *Fenchel-Nielson coordinates*, which identify TeichS_g up to homeomorphism as $\mathbb{R}^{6g-6}$.

Teichmüller's existence and uniqueness theorems

Let f be a map between complex surfaces, smooth outside of a finite number of points. We define the *dilatation* of f as

$$K(f) = \sup_z \frac{|f_z| + |f_{\overline{z}}|}{|f_z| - |f_{\overline{z}}|}.$$

At each point, df takes a circle to an ellipse. The dilatation $K(f)$ is the supremum of the ratios of the major axis to the minor axis. If the dilatation is finite we say f is *quasi-conformal*; if it is 1, then f is *conformal.*

If there are quadratic differentials q_X and q_Y on X and Y so that $f : X \to Y$ takes the zeros of q_X to those of q_Y, and if furthermore f sends the horizontal (resp. vertical) foliation of q_X to the corresponding foliation of q_Y with stretch factor $\sqrt{K}$ (resp. $1/\sqrt{K}$), we say that f is a *Teichmüller mapping* with horizontal stretch factor K.

Teichmüller's theorems give that every homotopy class of quasi-conformal maps between between complex surfaces contains a unique map minimizing dilatation, and furthermore this map is a Teichmüller mapping.

The uniqueness is approached first through the analogous problem for rectangles, Grötzch's problem, which admits a beautiful elementary solution. To get the general statement, time averages of the relevant derivative are considered.

We will outline the set up for the more difficult proof of existence. Fix a point $X \in \text{Teich}(S)$. Given a quadratic differentials q, we define its norm as $\|q\| = \int_X |q|$. In order to get a possible value for a dilatation, if q is in the unit ball of quadratic differentials $QD_1(X)$ we define

$$K(q) = \frac{1 + \|q\|}{1 - \|q\|}.$$

We define a map

$$\Psi : QD_1 \to \text{Teich}(S)$$

by sending q to the terminal surface of the Teichmüller mapping with initial differential q and horizontal stretch factor $K(q)$.

It is not too hard to see that Ψ is proper. The Measurable Riemann mapping theorem gives that it is continuous and by the uniqueness theorem we already know that it is injective. Using invariance of domain we get that the image of Ψ must therefore be closed and open, hence its range is all of Teich(S).

Suppose that $(X, \phi), (Y, \psi) \in \text{Teich}(S)$. Let f be the unique Teichmüller mapping homotopic to $\psi^{-1} \circ \phi$. The *Teichmüller distance* from (X, ϕ) to (Y, ψ) is defined as $\frac{1}{2} \log(K(f))$.

The mapping class group and moduli space. The *mapping class group* of a surface S (of genus g) is defined as $\text{Mod}(S) = \text{Homeo}^+(S)/\text{Homeo}_0^+(S)$. Here $\text{Homeo}_0^+(S)$ is the connected component of the identity in the group of self homeomorphisms which preserve orientation. Mod(S) acts naturally on Teich(S) by changing the marking: $f(X, \phi) = (X, \phi \circ f)$. The quotient $\mathcal{M}_g$ is *moduli space*, and can be shown to be an orbifold finitely covered by a manifold.

If S is a torus, then Teich(S) is isometric to hyperbolic space, $\text{Mod}(S) = SL_2(\mathbb{Z})$, and moduli space is the modular surface.

3. Translation surfaces

Definition. A *translation surface* is a flat surface with trivial linear holonomy. Away from a finite number of singularities, it has an atlas of charts whose transition functions are translations of $\mathbb{R}^2$, and each singularity has a cone angle that is a multiple of 2π.

Since $\mathbb{R}^2$ can be identified with $\mathbb{C}$, and the abelian differential dz is invariant under translations, every translation surface has an underlying complex structure with an abelian differential. Away from the singularities, the complex charts are simply the flat charts. At a singularity with cone angle $(k+1)2\pi$, the charts are given using the map $z \to z^k$, and the abelian differential has a zero of order k.

The fact that every translation surface has a complex structure is our first hint at connections to Teichmüller theory.

Conversely, given a complex surface X with a non-zero abelian differential ω, the transverse foliations defined as the kernels of the real and imaginary parts of ω give horizontal and vertical foliations of X, and hence define a flat structure. A zero of order k of ω has a cone angle of $2(k+1)\pi$.

Consider a set of polygons in the complex plane, with instructions to glue each edge to some parallel edge. After gluing the appropriate edges, we obtain a translation surface, whose abelian differential on each polygon is just dz. Conversely, given a translation surface, it is possible to cut into finitely many polygons.

A translation surface has a natural flow in the vertical direction. Inside of a polygon in the complex plane, this flow at time t is just translation by it. We are concerned with the long term behavior of this flow. For example, we wish to know if it has invariant subsets of non-trivial measure, that is, if it is *ergodic*.

Rational billiards. One motivation for studying translation surfaces comes from rational billiards, the study of the straight line flow in a polygon whose angles are rational multiples of π. When the straight line hits an edge, it bounces back in such a way that angle of incidence equals angle of reflection; we do not consider paths that hit vertices.

Translation surfaces arise from the following insight. Instead of bouncing off a wall, we may reproduce the polygon in mirror image on the other side of edge, and let the flow continue straight through the wall into another copy of the polygon. Formally, fix the polygon P and let G be the linear part of the group of affine isometries generated by reflections in the edges of P. Since the billiard is rational, G is finite. We associate to P the translation surface given by the polygons $gP, g \in G$ with expected edge identifications, and we say that the translation surface *comes from a billiard*.

$SL_2(\mathbb{R})$ **action.** This action is easiest to see with the definition of a translation surface (X, ω) as a set of polygons with pairs of parallel sides identified. In this case $A \in SL_2(\mathbb{R})$ acts on (X, ω) simply by acting on each polygon linearly and preserving the edge identifications. In either of the two definitions, A acts by post-composition with coordinate charts. The stabilizer of this action is $SL_2(X, \omega)$, and its image in $PSL_2(\mathbb{R})$ is defined as the *Veech group*. If $\mathrm{Aff}^+(X, \omega)$ is the set of affine orientation preserving automorphisms of (X, ω), then $SL_2(X, \omega)$ is the group of derivatives of elements of $\mathrm{Aff}^+(X, \omega)$.

Strata. The moduli space of abelian differentials forms a $\mathbb{C}^g$ vector bundle over the moduli space $\mathcal{M}_g$ of Riemann surfaces. Any holomorphic one form ω in a fiber has $2g-2$ zeros, counted with multiplicity. If α is a unordered partition of $2g-2$, define the stratum $\mathcal{H}(\alpha)$ to be set of ω whose zeros have multiplicities given by α. Fix α. The size of a set of singular points Σ of $(X, \omega) \in \mathcal{H}(\alpha)$ is constant, say $|\Sigma| = n$. We can pick a basis for $H_1(X, \Sigma; \mathbb{Z}) \cong \mathbb{Z}^{2g+n-1}$. Integrating ω over this basis gives a map $\mathcal{H}(\alpha) \to \mathbb{C}^{2g+n-1}$. This map gives *local holonomy coordinates* for the stratum, proving that it is a manifold. In these coordinates, $SL_2(\mathbb{R})$ acts naturally on each copy of $\mathbb{C}$, so the pullback of the Lebesgue measure gives a $SL_2(\mathbb{R})$–invariant measure on the stratum. Similarly, an invariant measure can be obtained on the unit norm part $\mathcal{H}_1(\alpha)$ of $\mathcal{H}(\alpha)$. This measure can be shown to be finite.

In summary, $\mathcal{H}_1(\alpha)$ is a non-compact manifold with a finite invariant measure. It is sometimes referred to as the *unit hyperboloid* in $\mathcal{H}(\alpha)$.

4. Ergodic theory

Definitions. For the moment, let (X, μ) be a probability space and let $g_t : X \to X$ be a one parameter group of measure preserving transformations. Take $f \in L_1(X)$. The Birkhoff ergodic theorem gives that the time averages

$$\hat{f}(x) = \lim_{T \to \infty} \int_{-T}^{T} f(g_t(x))dt$$

exist for almost all x and that $\|f\|_1 = \|\hat{f}\|_1$.

We say that g_t is *ergodic* if the only measurable g_t invariant subset have either full or null measure. In this case, we recall that the Birkhoff ergodic theorem gives that the time averages of an L^1 function are equal to its space average.

We say that g_t is *uniquely ergodic* if μ is the only Borel measure with respect to which g_t is ergodic.

Ergodicity of Teichmüller flow. The *Teichmüller flow* on the moduli space of abelian differentials is the portion of the $SL_2(\mathbb{R})$ action given by $g_t = \mathrm{diag}(e^t, e^{-t})$. This action is ergodic on each stratum.

An *interval exchange map* is a map $f : I \to I$ of an interval I that admits a partition into smaller intervals on which f is simply a translation. Let I be a horizontal interval on a translation surface (X, ω), whose interior does not contain any singularities. It is common to pick I so that the left endpoint is a singularity. The Poincaré recurrence theorem gives that every vertical trajectory leaving I comes back to I or hits a singularity. For $x \in I$, let $f(x)$ be the first point where the vertical trajectory leaving I returns to I.

A finite number of vertical trajectories from I will hit singularities before returning to I. By removing the corresponding points of I we partition I into intervals, and get that f is an interval exchange map. Assume the vertical flow is minimal. By extending rectangles up from these intervals, we get a *zippered rectangle* decomposition of (X, ω). The interval exchange map f and the heights of the rectangles determine the surface, and surface depends continuously on this data.

We can now summarize a proof of ergodicity of Teichmüller flow.

(1) First assume that (X, ω) and (X', ω') have the same vertical foliations, and have uniquely ergodic vertical foliations. Pick an interval as above. The first return maps f of (X, ω) and (X', ω') will be the same, but the heights of the rectangles might be different.
(2) Consider a very long vertical path, starting at I. Because of unique ergodicity, the number of times this path hits each subinterval of I is in proportion to its length. Hence the length of the path is approximately proportional to the number of times that it crosses I. Pick a subinterval I' of I that is small enough that every trajectory leaving I' is very long. We get a new interval exchange map f'. The length of each path leaving I' is approximately proportional to the number of times that it crosses I before returning to I'. This depends on f and not (X, ω).
(3) Consider the zippered rectangle decompositions of (X, ω) and (X', ω'). The ratio of the lengths of corresponding rectangles in the two translation surfaces will be very close to 1. Apply g_t for large t. The result is zippered rectangles with the same interval exchanges and very close heights of rectangles. Hence $g_t(X, \omega)$ and $g_t(X', \omega')$ are close.
(4) Show that $g_t(X, \omega)$ and $g_t(X', \omega')$ stay close for larger t. If f is a uniformly continuous function, conclude that the time averages starting from (X, ω) and (X', ω') are the same.
(5) Mimic the proof that ergodicity of geodesic flow on a finite volume hyperbolic surface is ergodic. Use the fact that almost every translation surface has uniquely ergodic vertical foliation. To connect a generic pair of translation surfaces, an intermediate differential is defined whose transverse foliations are the vertical foliations of the original two differentials.

Masur's criterion for unique ergodicity. Masur's criterion, proven using the technique of renormalization in dynamics, gives that if the vertical flow on (X, ω) is not uniquely ergodic then $g_t(X, \omega)$ is *divergent*, that is, it eventual leaves every compact set (and does not come back).

For example, if (X, ω) has a saddle connection in the vertical direction, we would not expect unique ergodicity. Indeed since the length of this saddle connection goes to zero, $g_t(X, \omega)$ eventually leaves every compact set.

The ergodic theory of translation surfaces. For a general translation surface, Masur has shown that almost every direction is uniquely ergodic, and a dense but countable set of directions are periodic. The first result uses Masur's criterion above. The second result uses the compactification of moduli space.

A much easier result is that for all but a countable set of directions (the directions with saddle connections), directional flow is *minimal*, that is, has dense orbits. The closure of an orbit can be seen to be a subsurface whose boundary, if it is non-empty, is the union of saddle connections. These saddle connections must be in the direction of the flow. Thus by flowing in a direction with no saddle connections, we ensure that all orbits are dense.

The Veech alternative. Let (X, ω) be a translation surface and Γ be its Veech group. The Veech alternative states that if Γ is a lattice (we say (X, ω) is a *lattice surface*) then for every direction θ either:

- $r_\theta(X, \omega)$ admits a *complete cylinder decomposition*, a decomposition into finitely many cylinders. We say $r_\theta(X, \omega)$ is *completely periodic*. Furthermore in this case the moduli of the cylinders are rationally related.
- The vertical flow on $r_\theta(X, \omega)$ is uniquely ergodic.

The dichotomy of completely periodic versus uniquely ergodic is called *optimal dynamics*. Here is a sketch of the proof.

(1) Assume $\theta = 0$ and the flow is not uniquely ergodic. By Masur's criterion, $g_t(X, \omega)$ diverges to infinity on the finite volume hyperbolic surface $\mathbb{H}/\Gamma$ inside of the moduli space of abelian differentials.

(2) Hence $g_t(X,\omega)$ goes to infinity along a cusp corresponding to a parabolic $g \in SL_2(X,\omega)$. There is a diffeomorphism ϕ of (X,ω) whose derivative is constantly g. Replacing ϕ with a power, we may assume that ϕ fixes the singularities of (X,ω). We may rotate the surface so that g is upper triangular.
(3) ϕ is the identity on vertical trajectories leaving singularities, hence on the closure on these singularities. Closures of trajectories are sub-manifolds with boundary, so each vertical trajectory leaving a singularity is a saddle connection. These saddle connections bound cylinders.

Smillie's theorem. In the case of regular polygons with opposite edges identified, elementary computations using the theory of Fuchsian groups show that Γ is a triangle group. However, when Γ is not generated by rotations and parabolics, for example when $\mathbb{H}/\Gamma$ has positive genus, then Γ cannot be computed by hand, even up to finite index. So a more abstract tool is required to identify lattice surfaces.

This tool, Smillie's theorem, says that if the $SL_2(\mathbb{R})$ orbit of of (X,ω) is closed, then Γ is a lattice. Here is an outline of a proof.

(1) Write $B = SL_2(\mathbb{R}) \cdot (X,\omega)$, and assume B is closed. Denote horocycle flow by u_t. The map $SL_2(\mathbb{R})/SL(X,\omega) \to B$ is a homeomorphism.
(2) Get a locally finite G invariant measure μ on $G \cdot (X,\omega)$.
(3) Pick an everywhere positive continuous function $f \in L^1(\mu)$. Define

$$S_n = \{y \in B : \liminf_{T\to\infty} f(u_t y)dt \geq \frac{1}{n}\}.$$

Note that S_n has finite measure and is u_t invariant and closed.
(4) Use quantitative recurrence of horocycle flow to show that for some n, $\mu(S_n) > 0$.
(5) A Mautner type theorem holds: Any vector in $L^2(\mu)$ that is fixed by u_t is fixed by all of $SL_2(\mathbb{R})$. Apply this theorem to the indicator function of S_n to get that S_n is all of B.
(6) We have found a $SL_2(\mathbb{R})$–invariant finite measure on $SL_2(\mathbb{R})/SL(X,\omega)$, so $SL(X,\omega)$ must be a lattice.

Teichmüller disks. Suppose (X,ω) is a lattice surface. Pick a point in Teichmüller space corresponding to X, and consider the geodesic through X given by ω^2. The complexification of this geodesic is a $SL_2(\mathbb{R})$–orbit, and its image in $\mathcal{M}_g$ space is closed.

When the projection to $\mathcal{M}_g$ of a complex geodesic in Teichmüller space is closed, we call it a *Teichmüller curve*. Smillie's theorem gives that lattice surfaces correspond to Teichmüller curves.

Since the $SL_2(\mathbb{R})$ action on strata is ergodic, typically the projection of a closed geodesic to $\mathcal{M}_g$ is dense. Teichmüller geodesics are very rare. They are totally geodesic subspaces of $\mathcal{M}_g$. This fact can be shown using the following observation and the KAK decomposition. The geodesic flow g_t on hyperbolic space moves points a distance of t. On a translation surface (X,ω), g_t acts as a Teichmüller mapping with dilatation e^{2t}. Hence the Teichmüller distance between (X,ω) and $g_t(X,\omega)$ is $\frac{1}{2}\log e^{2t} = t$.

5. Recent work and hopes for the future.

Some central problems. A Teichmüller curve is *primitive* if it does not arise via a branched cover. One central problem is the classification of primitive Teichmüller curves.

Masur and others have shown that the growth of the number of saddle connections, and the number of cylinders, is bounded quadratically above and below, and is in many instances exactly quadratic. The calculation of associated constants is a central problem.

There are numerous results which hold for almost every translation surface. Since billiards are measure zero in translation surfaces, this is a enormous obstacle for the application of the theory of translations surfaces to billiards. A Ratner's theorem might allow precise description on the sets where these statements fail; hopefully new theorems could be proven on these sets. In fact, a Ratner's theorem has been called a hoped for "magic wand" **[10]**.

There are many other open problems which are seemingly more elementary. For example, the problem of determining when two surfaces are in the same $SL_2(\mathbb{R})$ orbit is open, even when the surfaces are square tiled. It is completely unknown which groups arise as Veech groups, although it is known that infinitely generated groups do occur.

Success in genus 2. For a lattice surface one might hope for symmetries in the underlying complex surface. It turns out that symmetries are more apparent on the much better behaved Jacobian.

If $(X, \omega) \in \mathcal{H}(2)$, then $\mathrm{Jac}(X)$ of a lattice surface (X, ω) admits real multiplication by the trace field $\Bbbk$ of the Veech group. That is, there is a linear map $\Bbbk \to \mathrm{End}(\mathrm{Jac}(X)) \otimes \mathbb{Q}$. Furthermore, ω is an eigenform for this real multiplication.

Jacobians are principally polarized abelian varieties. The moduli space of principally polarized abelian varieties admitting real multiplication by a given totally real number field is a *Hilbert modular variety*. McMullen showed that all primitive Teichmüller curves in $\mathcal{H}(2)$ lie on countably many Hilbert modular surfaces, and in doing so identified a new infinite family of primitive Teichmüller curves given by L–shaped billiard tables.

McMullen also shown that when the Veech group of a genus two translation surface contains a hyperbolic, then the limit set of the Veech group is the entire circle. This result gives an infinite family of translation surfaces with infinitely generated Veech group. This work involves defining a quantity called the Galois flux of an interval exchange map and also a translation surface, which measures the growth of the Galois conjugates of a first return map. Other tools involve Diophantine properties of square roots and surgery on foliations.

Other work of McMullen and separately Calta has given other successes in genus 2, even Ratner's theorem. Calta's approach uses the J–invariant, a generalization of the SAF invariant for interval exchange maps.

References

[1] M. B. Mayer and M. Bekka, *Ergodic Theory and Topological Dynamics of Group Actions on Homogeneous Spaces*. London Math Society, 2000.

[2] B. Farb and D. Margalit, *A primer on the mapping class group*. Working draft, version 3.1. Available online at `www.math.utah.edu/~margalit/primer/`.

[3] A. Eskin, *Counting problems in moduli space*. Handbook of dynamical systems. Vol. 1B, 581–595, 2006.

[4] Y. Katznelson and B. Weiss, *A simple proof of some ergodic theorems*. Israel J. Math. **42** (1982), no. 4, 291–296.

[5] D. Masur and S. Tabachnikov, *Rational billiards and flat structures*. Handbook of dynamical systems, Vol. 1A, 1015–1089, 2002.

[6] C. T. McMullen, *Billiards and Teichmller curves on Hilbert modular surfaces*. J. Amer. Math. Soc. **16** (2003), no. 4, 857–885.

[7] Y. Minsky and B. Weiss, *Nondivergence of horocyclic flows on moduli space*. J. Reine Angew. Math. 552 (2002), 131–177.

[8] J. Smillie and B. Weiss, *Minimal sets for flows on moduli space*. Israel J. Math. 142 (2004), 249–260.

[9] W. A. Veech, *Geometric realizations of hyperelliptic curves*. Algorithms, fractals, and dynamics (Okayama/Kyoto, 1992), 217–226, Plenum, New York, 1995.

[10] A. Zorich, *Flat surfaces*. Frontiers in number theory, physics, and geometry. I, 437–583, Springer, Berlin, 2006.