Math 181 Handout 14

Rich Schwartz

December 6, 2005

The purpose of this handout is define the Veech group of a translation surface, and connect it to some hyperbolic geometry. A lot of this material can be found in various surveys of rational billiards e.g. *Dynamics and Rational Billiards*, by Howie Masur and Sergei Tabachnikov.

1 Translation Surfaces

Recall that a Euclidean cone surface is a compact surface, such that every point has a neighborhood isometric to a Euclidean cone, and only finitely many of these cones have cone angle which is not 2π. A *translation surface* is a Euclidean cone surface such that all the cone angles are multiples of 2π. We always assume these surfaces are connected.

Exercise 1: Consider the regular Euclidean octagon X. Form a surface by gluing opposite sides of X together, in "the usual way". Let Q be the quotient. One can define the distance between two points on Q to be the inf of the Euclidean lengths of paths joining the two points. Prove that Q, with this metric, is a translation surface with one nontrivial cone point having cone angle 6π.

Exercise 2: Let Σ be any surface having genus at least 1. Prove that there is a translation surface homeomorphic to Σ. In other words, you can get any compact oriented surface (except the sphere) as a translation surface.

Theorem 1.1 *Let Σ be a translation surface. Let C denote the finite set of nontrivial cone points of Σ. Then $\Sigma - C$ admits a continuous vector field in* *which the trajectories are all locally straight paths.*

Proof: Choose some basepoint $x \in \Sigma - C$. Let $v(x)$ be some unit vector tangent to x. Our goal is to define a unit vector $v(y)$ for each point $y \in \Sigma - C$. Here is the construction. Let γ be any smooth curve which joins x to y and stays in $\Sigma - C$. Say that a vectorfield along γ is *parallel* if, in the local coordinates, the vectors are all translates of each other. Since every point of γ has a neighborhood which is isometric to a disk in $\boldsymbol{R}^2$, there is a unique parallel vectorfield along γ which agrees with $v(x)$ at x. We define $v(y)$ to be the vector of this parallel vector field at y. If this is really well defined, then in small neighborhoods, our vectorfield consists entirely of parallel vectors. Hence, the trajectories are all locally straight lines.

To finish our proof, we need to see that this definition is independent of the path γ. If γ_1 and γ_2 are paths connecting x to y, and homotopic relative their endpoints, then we can produce a finite sequence of paths $\gamma_1 = \beta_1, ..., \beta_n = \gamma_2$ such that β_i and β_{i+1} agree except in a region which is contained in a single Euclidean disk. (You get the β curves just by doing the homotopy a little bit at a time.) Within the Euclidean disk, you can see that the vectorfield along β_i must be parallel to the vectorfield along β_{i+1}, because both vector fields just consist of a bunch of parallel vectors, and the two vector fields agree at some point in the disk. Since this is true for all i, the two methods for defining $v(y)$ agree.

The fundamental group $\pi_1(\Sigma - C)$ is generated by loops which travel from x into a small neighborhood of one of the cone points, wind around the cone point, and then come back. If γ_1 and γ_2 are arbitrary paths joining x to y then γ_1 is homotopic relative to the endpoints to $\delta_1 * ... * \delta_k * \gamma_2$, where each δ_i is one of the special loops just mentioned. Each loop δ_i starts and ends at x. We just have to see that the parallel vectorfield along δ_i agrees with $v(x)$ at both ends. Everything boils down to what happens in a neighborhood of the cone point. You can take the portion of δ_i which loops around the cone point to be a perfect circle; and in this case you can see that the parallel vectorfield along the circle comes back exactly to itself, because the cone angle is a multiple of 2π. ♠

Exercise 3: Prove that there is no vectorfield, like the one constructed above, on the surface of the cube.

2 Affine Automorphisms

Recall that an *affine map* of $\boldsymbol{R}^2$ is a map of the form $x \to Ax + B$, where A is a 2×2 invertible and orientation preserving matrix and B is another vector. If $B = 0$ then the map is linear. Note that the set of affine maps of $\boldsymbol{R}^2$ forms a group under composition.

Suppose that Σ is a translation surface. An affine automorphism of Σ is a homeomorphism $\phi : \Sigma \to \Sigma$ such that

- ϕ permutes the nontrivial cone points of Σ.
- Every ordinary point of Σ has a neighborhood in which ϕ is an affine map.

The second condition needs a bit more explanation. Let $p \in \Sigma$ be an ordinary point. This is to say that there is a small disk Δ_p about p and an isometry I_p from Δ_p to a small disk in $\boldsymbol{R}^2$. The same goes for the point $q = \phi(p)$. The map $I_q \circ \phi \circ I_p^{-1}$ is defined on the open set $U = I_p(\Delta_p) \subset \boldsymbol{R}^2$ and the second condition says that this map is the restriction of an affine map to U.

We denote the set of all affine automorphisms of Σ as $A(\Sigma)$. It is easy to see that the composition of two affine automorphisms of Σ is again an affine automorphism. Likewise, the inverse of an affine automorphism of Σ is an affine automorphism of Σ. In short, $A(\Sigma)$ is a group.

Exercise 4: Let A be a 2×2 matrix with integer entries and determinant 1. Let B any vector. Let Σ be the square torus. You can think of Σ as $(\boldsymbol{R}/\boldsymbol{Z})^2$. Let ϕ be the map $\phi([x]) = [Ax + B]$. Prove that ϕ is an affine automorphism of Σ. Thus, the square torus has a huge affine automorphism group.

Exercise 5: Give an example of a translation surface which has no nontrivial affine automorphisms.

Exercise 6: The affine automorphisms group of the square torus is uncountable since it contains any translation. However, prove that the affine automorphism group of a surface with at least one cone point is countable. (Hin: One way to see this is that the linear differential of the affine automorphism determines the map, once its behavior on a single point is specified.)

3 The Diffential Representation

Let $SL_2(\boldsymbol{R})$ denote the group of determinant-one 2×2 matrices with real entries. Given a group A, a *representation* of A into $SL_2(\boldsymbol{R})$ is a homomorphism $\rho : A \to SL_2(\boldsymbol{R})$. Here is one explanation for this terminology: The elements of A might be somehow abstract, but a representation is a way of, well, representing these elements concretely as matrices. A representation doesn't have to be one-to-one or onto, but of course representations with these additional properties are especially nice.

Here we explain a canonical representation $\rho : A(\Sigma) \to SL_2(\boldsymbol{R})$. The first thing to recall about Σ is that it has this decomposition into straight lines. Let p and q be two regular points of Σ. There is a canonical map ϕ_{qp} from the tangent plane $T_q(\Sigma)$ to the tangent plane $T_p(\Sigma)$. This map can be specified by the following properties:

- ϕ_{qp} maps the unit vector tangent to the special line through q to the unit vector tangent to the special line through p.
- ϕ_{qp} is linear.
- ϕ_{qp} is an isometry, measured in the natural Riemannian metric on Σ.

The third point needs some explanation. Neighborhoods of p and q are isometric to little disks in the plane, and we simply use the Riemannian metric−i.e. the standard inner product−to put a Riemannian metric on the tangent spaces at p and q. Put another way, a curve through p has unit speed if and only if the image of the curve in $\boldsymbol{R}^2$ has unit speed. Likewise, the angle between two curves through p is the angle between the images of these curves in $\boldsymbol{R}^2$. The same remarks go for q.

Now, given an element $f \in A(\Sigma)$ we choose and ordinary point $p \in \Sigma$ and let $q = f(p)$. Let df_p be the differential of f at p. This means that df_p is a linear map from $T_p(\Sigma)$ to $T_q(\Sigma)$. Note that the composition

$$M(f,p) = \phi_{qp} \circ df_p$$

is a linear isomorphism from $T_p(\Sigma)$ to itself. Using the isometry I_p we can identify $T_p(\Sigma)$ with, say the tangent plane to $\boldsymbol{R}^2$ at the origin. We let $\rho(f)$ be the linear transformation of $\boldsymbol{R}^2$ which corresponds to $M(f,p)$ under the identification.

We claim that $\rho(f)$ is independent of the choice of point p. To see this, we note that the map $\rho(f)$ has the following alternate description. Using the coordinate charts I_p and I_q discussed above, the map $\rho(f)$ is just the linear part of

$$dI_q \circ df_p \circ dI_p^{-1}.$$

The linear part of an affine map does not depend on the point. Hence $\rho(f)$ has the same definition if independent of which point we use inside our local coordinate chart. But the surface is connected, so $\rho(f)$ does not dependent on the choice of point at all.

The determinant of $\rho(f)$ measures the factor by which f increases area in a neighborhood of any point. Since the whole surface has finite area and $\rho(f)$ is an automorphism, $\rho(f)$ must have determinant 1. Hence we can interpret $\rho(f)$ as an element of $SL_2(\boldsymbol{R})$. The map $f \to \rho(f)$ is a homomorphism because of the chain rule: The linear differential of a composition of maps is just the composition of the linear differential of the invididual maps; and composition of linear maps is the same thing as matrix multiplication in $SL_2(\boldsymbol{R})$.

We have now constructed the representation $\rho : A(\Sigma) \to SL_2(\boldsymbol{R})$. We let $V(\Sigma) = \rho(A(\Sigma))$. The matrix group $V(\Sigma)$ is sometimes called the *Veech group*.

4 Hyperbolic Group Actions

Recall that $\boldsymbol{H}^2$ is the hyperbolic plane. Every element of $SL_2(\boldsymbol{R})$ acts on $\boldsymbol{H}^2$ as an isometry. If we identify $\boldsymbol{H}^2$ with the upper half plane in $\boldsymbol{C}$, then the action is given by

$$z \to \frac{az+b}{cz+d}; \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\boldsymbol{R}).$$

Let $V = V(\Sigma)$ be as above. The *orbit* of a point $x \in \boldsymbol{H}^2$ is defined to be the set

$$\{g(x) |\ g \in V\}.$$

We define an equivalence relation on points in $\boldsymbol{H}^2$ by saying that two points are equivalent iff they lie in the same orbit.

V is said to act *properly discontinuously* on $\boldsymbol{H}^2$ for every metric ball $B \subset \boldsymbol{H}^2$ the set

$$\{g \in V |\ g(B) \cap B \neq \emptyset\}$$

is a finite set. In other words, all but finitely elements of V have such a drastic action on $\boldsymbol{H}^2$ that they move the ball B completely off itself.

Exercise 7: Prove that $SL_2(\boldsymbol{Z})$, the group of 2×2 integer determinant-one matrices, acts properly discontinuously on $\boldsymbol{H}^2$.

Before we establish the main result in this section we give one more definition. Two groups $G_1, G_2 \in SL_2(\boldsymbol{R})$ are *conjugate* if there is some $g \in SL_2(\boldsymbol{R})$ such that $G_2 = gG_1g^{-1}$.

Exercise 8: Suppose that G_1 and G_2 are conjugate. Prove that G_1 acts properly discontinuously on $\boldsymbol{H}^2$ if and only if G_2 does.

Theorem 4.1 *If V is the Veech group of a surface then V acts properly discontinuously on $\boldsymbol{H}^2$.*

We will sketch the proof of Theorem 4.1 in the next section.

Whether or not V acts properly discontinuously, we can form the quotient $\boldsymbol{H}^2/V$ as follows. We define two points $x, y \in \boldsymbol{H}^2$ to be equivalent if there is some $g \in V$ such that $g(x) = y$. Then $\boldsymbol{H}^2/V$ is defined to be the set of equivalence classes of points. In case V acts properly discontinuously the quotient is particularly nice:

Theorem 4.2 *If V acts properly discontinuously on $\boldsymbol{H}^2$ then we can remove a countable discrete set of points T from $\boldsymbol{H}^2$ such that the quotient $(\boldsymbol{H}^2 - T)/V$ is a hyperbolic surface.*

Proof: Before we start we note that all the elements of V act so as to preserve orientation, so that there are no reflections in V. (For the orientation reversing case, the statement of the result is slightly different.)

Let T be the set of points $x \in \boldsymbol{H}^2$ such that $g(x) = x$ for some nontrivial $g \in V$. The set T must be discrete in the sense that there is some $\epsilon > 0$ such that any ball of radius ϵ contains at most one point of T. Otherwise we could find some ball B which contained infinitely many points of T and we would contradict the proper discontinuity. Note that T is invariant under V: If $x \in T$ is fixed by g then $y = h(x)$ is fixed by hgh^{-1}. Thus, the quotient $(\boldsymbol{H}^2 - T)/V$ makes sense. Every $x \in \boldsymbol{H}^2 - T$ has a neighborhood Δ_x such that $g(\Delta_x) \cap \Delta_x = \emptyset$ for any nontrivial g. To see this, let d_g denote the hyperbolic distance between $g(x)$ and x. Since $x \not\in T$, the number d_g is positive. The proper discontinuity prevents there being a sequence $\{g_i\}$ with $\{d_{g_i}\}$ converging to 0. Hence there is some positive lower bound to d_g, which is what we need.

Now we know that each $x \in \boldsymbol{H}^2 - T$ has a little neighborhood which is moved completely off itself by all of G (except the identity.) This little neighborhood therefore maps injectively into the quotient $(\boldsymbol{H}^2 - T)$ and serves as a coordinate chart about x. ♠

Note that the quotient $\boldsymbol{H}^2/V$ still makes sense, and actually it is obtained from $(\boldsymbol{H}^2 - T)/V$ just by adding finitely many points. We define the *co-volume* of V to be the volume of $(\boldsymbol{H}^2 - T)/V$. The group V is said to be a *lattice* if V has finite co-volume. Σ is said to be a *Veech surface* if V is a lattice. For instance, $SL_2(\boldsymbol{Z})$ is a lattice.

5 Proof of Theorem 4.1

Before we begin, we need one piece of terminology. Let $p \in \Sigma$ be a cone point and let γ be a path on Σ which does not contain any other cone points. Then γ is *essentially straight* if $\gamma - p$ is a locally straight path in $\Sigma - p$.

Lemma 5.1 *Let f be an affine automorphism of Σ and suppose that $f(\gamma) = \gamma$ for some essentially straight path. Then f is the identity on γ. If γ has a self-intersection, then f is in the kernel of the differential representation ρ.*

Proof: The restriction of an affine map to a straight line is just a dilation. Hence, the restriction of f to γ is just a dilation. Since $f(\gamma) = \gamma$, the dilation factor must be one: The total length is preserved. So f is an isometry on γ. Since $f(p) = p$ we see that actually $\rho(f)$ is the identity on γ.

For the second half, suppose x is a self-intersection point of γ. Then $f(x) = x$ and f is the identity on two distinct paths emanating from x. Hence df is the identity on a basis of $T_x(\Sigma)$. Hence df is the identity at x. But this means that $\rho(f)$ is the identity. ♠

Now we turn to the proof of Theorem 4.1. There are three cases.

- Σ has no cone points.

- Σ has one cone point.
- Σ has more than one cone point.

5.1 No Cone Points

Σ must be a flat torus, built from some parallelogram P. There is an affine map g such that $g(P)$ is a square and hence $g(\Sigma)$ is the square torus. g conjugates $A(\Sigma)$ to the affine automorphism group of the square torus, which (after factoring out translations) we know to be $SL_2(\boldsymbol{Z})$. Hence $V = gSL_2(\boldsymbol{Z})g^{-1}$. Exercises 7 and 8 finish the proof.

5.2 One Cone Point

We suppose that there is some ball B and an infinite collection $\{g_i\} \in V$ such that $g_i(B) \cap B \neq \emptyset$. It is a general principle of compactness that there must be elements of our set which are arbitrarily close to each other. Hence, we can find an infinite list of distinct elements of V whose action on $\boldsymbol{H}^2$ converges to the action of the identity element.

What this means in terms of Σ is that we can find an infinite sequence $\{f_j\}$ of affine automorphisms such that $\rho(f_i)$ is not the identity but $\rho(f_i)$ converges to the identity as $i \to \infty$.

Consider the fundamental group $\pi_1 := \pi_1(\Sigma, p)$. Certainly we can find elements of π_1 which cannot be represented by any simple loops. Let $[\gamma]$ be such an element. Two essentially straight representatives of $[\gamma]$ must coincide if they are sufficiently close. To see this, just roll these two paths out into the plane. They would have to be the two sides of a very narrow strip, and hence would be parallel forever. But then they don't both converge to p.

So, $[\gamma]$ has only finitely many essentially straight representatives. Call them $\gamma_1, ..., \gamma_k$. If i is large then $f_i(\gamma_j)$ is an essentially straight path which is quite close to γ_j. But then these paths are homotopic by a fairly obvious homotopy. Hence f_i permutes the curves $\gamma_1, ..., \gamma_k$ for i large. There is some fairly small power $f = f_i^r$ such that $f(\gamma_1) = \gamma_1$. But then $\rho(f)$ is the identity by Lemma 5.1. Here r only depends on k. If $\rho(f_i)$ is very close to the identity and $\rho(f_i^r)$ is the identity then $\rho(f_i)$ is also the identity. This is a contradiction.

5.3 More than one Cone Point

Here we can play the same game, considering almost straight representatives of self-intersecting paths connecting possibly different cone points on Σ. We omit the details.

6 Triangle Groups

Recall that a geodesic hyperbolic triangle is a triangle in $\boldsymbol{H}^2$ whose sides are either geodesic segments, geodesic rays, or geodesics. The case of interest to us is the $(8, \infty, \infty)$ triangle. This is a geodesic triangle, two of whose sides are geodesic rays and whose remaining side is a geodesic. Two of the vertices of this triangle are on the ideal boundary of $\boldsymbol{H}^2$ and the remaining vertex $v \in \boldsymbol{H}^2$ is the common endpoint of the two rays. The two rays make an angle of $2\pi/8$ at v.

Lemma 6.1 *Let γ be any geodesic in $\boldsymbol{H}^2$. Then there is an order 2 hyperbolic isometry which fixes γ.*

Proof: Thinking of $\boldsymbol{H}^2$ as the upper halfplane, the map $z \to -\overline{z}$ fixes the imaginary axis, which is a geodesic. We've already seen that any two geodesics are isometric to each other. If g is an isometry taking the geodesic γ_1 to the geodesic γ_2, and I is an order 2 isometry fixing γ_1 then gIg^{-1} is the desired order 2 isometry fixing γ_2. Thus, we can start with the one reflection desribed above and construct all the others by conjugation. ♠

The order 2 hyperbolic isometry fixing γ is called a *hyperbolic reflection* in γ. Given any geodesic triangle Δ we can form the group $G(\Delta) \subset SL_2(\boldsymbol{R})$ as follows. We let I_1, I_2, I_3 be hyperbolic reflections fixing the 3 sides of Δ and then we let $G(\Delta)$ be the group generated by words of even length in I_1, I_2, I_3. For instance, I_1I_2 and $I_1I_2I_1I_3$ all belong to G but $I_1I_2I_3$ does not. All the elements in G are orientation preserving and it turns out that we can find matrices in $SL_2(\boldsymbol{R})$ for the three elements I_iI_j. This is enough to show that G actually comes from a subgroup of $SL_2(\boldsymbol{R})$.

7 Behold, The Octagon!

We've already seen that one can make a genus 2 hyperbolic surface by gluing the correctly sized regular hyperbolic octagon together. Here will discuss another connection between the regular octagon and hyperbolic geometry. This connection was discovered essentially by Veech. Veech's result works for general n-gons, but its easier just to discuss it for one of them.

Let Σ be the genus 3 translation surface constructed by the following procedure. Take two copies O_1 and O_2 of the regular unit octagon. Glue each side of O_j to the opposite side of O_{3-j}. This produces a surface with two cone points of cone angle 6π which double covers the usual surface made by gluing together a single octagon.

Theorem 7.1 *$V(\Sigma)$ is the same group as $G(\Delta)$ where Δ is the $(8, \infty, \infty)$ hyperbolic triangle group.*

We will sketch a proof of this result. [1] To make things work well, we define an *anti-affine automorphism* to be a homeomorphism of Σ which is locally anti-affine, meaning that the map locally has the form $x \to Ax + B$ where A is an orientation reversing linear map. We let $\widehat{A}(\Sigma)$ be the group of these maps and we let $\widehat{V} = \rho(\widehat{A})$. We will show that $\widehat{V}$ coincides with the group $\widehat{G}$ generated by the reflections in the sides of the $(8, \infty, \infty)$ triangle.

Think of O as being made from two octagons placed side to side, as in Figure 1 below. Simultaneous reflection in the vertical lines of symmetry of the two octagons is an anti-affine automorphism. Call it f_1. The lines through the centers of the octagons, which make an angle of $\pi/8$ with the vertical, are also lines of bilateral symmetry. Let f_2 be the anti-affine automorphism defined by simultaneous reflections in these two lines. The elements $\rho(f_1)$ and $\rho(f_2)$ are two of the three generators of $\widehat{G}$. The axes of f_1 and f_2 are drawn thickly on the right hand side of Figure 1.

The third element is the non-trivial one. Consider the decomposition of Σ into cylinders, as indicated by Figure 1.

[1] I learned this proof from Pat Hooper, and Pat has a great paper on his website which gives an argument like this for a new Veech surface he discovered.

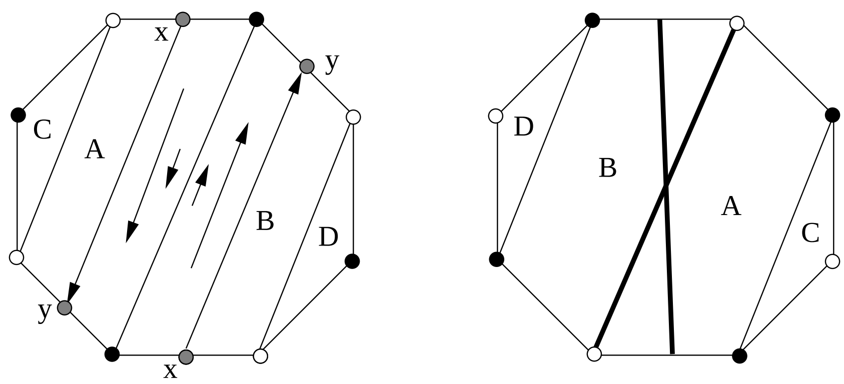

Figure 1

Let g be the affine automorphism characterized by the following properties:

- g does the same thing to both octagons.
- g maps the points labelled x to the points labelled y, in the manner suggested by the arrows. These points are at the midpoints of the edges they lie on.
- Figure 1 shows a decomposition of Σ into 4 cylinders, A, B, C, D. Every point on the boundary of any of these cylinders is fixed by g. That is, g is a Dehn twist of each cylinder.
- $f_2 g f_2 = g^{-1}$. (Actually, this is a consequence of symmetry and the other properties.)
- The differential dg is as close to the identity as possible. By this we mean, informally, that g is the "shear" of minimal "strength" which has the other properties.

The arrows on the left sort of indicate the action of g in the vicinity of the axis of symmetry. Let

$$f_3 = f_2 g.$$

Exercise 9: Show that f_3 fixes the axes of both f_1 and f_2. On the right octagon, these axes are the thick lines.

Note that f_3 has order 2 because

$$f_3^2 = f_2 g f_2 g = g^{-1} g = Id.$$

The fact that f_3 has order 2 and fixes the axes of both f_1 and f_2 forces $\rho(f_3)$ to be the reflection in the third side of our $(8, \infty, \infty)$ triangle. Thus we see that $\widehat{V}$ contains the three generators of $\widehat{G}$. Hence $\widehat{G} \subset \widehat{V}$.

Exercise 10: Suppose that Γ is a group acting properly discontinuously on $\boldsymbol{H}^2$ and $\widehat{G} \subset \Gamma$. Then either $\Gamma = \widehat{G}$ or else Γ is the group generated by the reflections in the sides of the geodesic triangle obtained by bisecting Δ in half. (Hint: Consider all the geodesics fixed by reflections of elements of Γ. These lines decompose $\boldsymbol{H}^2$ into regions which must be permuted by Γ. In particular, these regions partition Δ in some way. The only possibility that leads to a properly discontinuous group is the bisection of Δ into two equal halves.)

If $\widehat{V}$ does not equal $\widehat{G}$ then Σ has an extra isometric symmetry which fixes the centers of the octagons. (This corresponds to the extra element, reflection in the bisector of Δ.) But the octagons do not have any line of symmetry between the two drawn in figure 1. Hence $\widehat{V} = \widehat{G}$.