274 Curves on Surfaces, Lecture 1

Dylan Thurston
Notes by Qiaochu Yuan

Fall 2012

1 Introduction

We consider simple closed curves on surfaces.

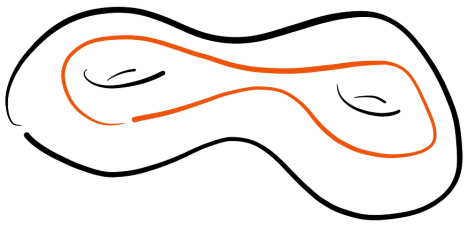


Figure 1: A simple closed curve on a 2-holed torus.

We want to understand such curves. There is a cluster algebra structure related to this, as well as other structures which look like cluster algebras but are not quite formalized. We can study curves on four levels:

1. Tropical - the curves themselves.
2. Algebraic - the cluster algebra. This involves studying Teichmüller space - the space of (uniform) hyperbolic metrics on the surface. Equivalently, the discrete faithful representations of the fundamental group into $\mathrm{PSL}_2(\mathbb{R})$.
3. Quantum - a noncommutative deformation of (the algebra of functions on) Teichmüller space. Related to the Jones polynomial and quantum 3-manifold invariants.
4. Categorical - this does not quite exist yet. Various integers should become objects in a category (and we get the numbers back by taking dimensions, for example). We should get 4-manifold invariants (the last frontier of low-dimensional topology).

The first two levels have been studied for a long time, although there are still open questions. For now we will talk about the first level.

Imagine stirring around the foam in a cup of coffee. The foam may look simple in the beginning but it gradually becomes more complicated and then becomes roughly uniformly distributed in the coffee. More mathematically, consider the disc with three punctures. Begin with a loop around, say, the leftmost two punctures. Permute the punctures and observe what happens to the curve.

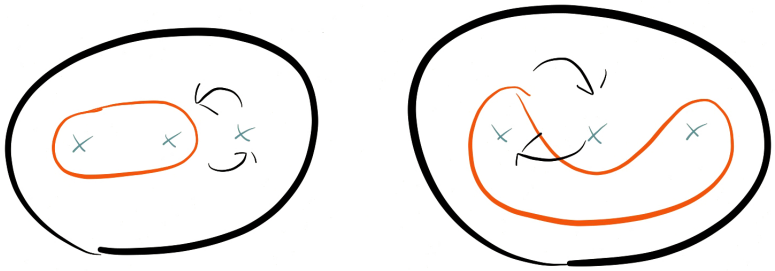


Figure 2: The curve after zero and one permutations.

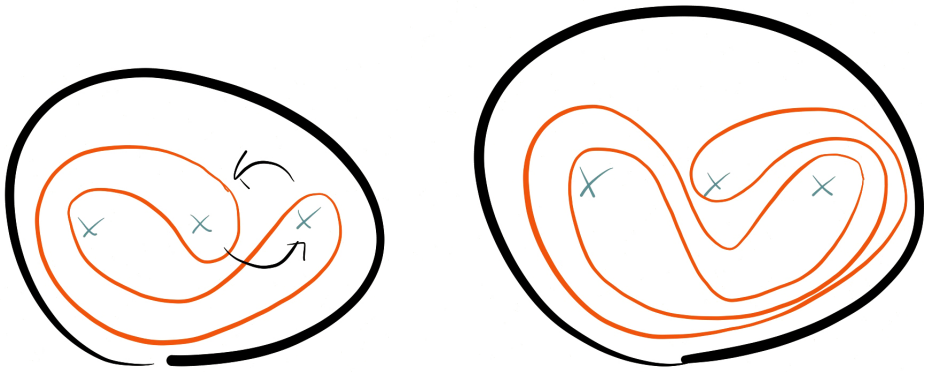


Figure 3: The curve after two and three permutations.

The curve gets more complicated; what can we say about this process?

One way to describe this process is to imagine a movie describing the permutations. The punctures trace out 3-stranded braids, and moving the curve is like trying to pull a rubber band down through the braid. How do we keep track of the curve?

One idea is to triangulate. To do this it will be convenient to add two additional punctures on the boundary. Then the disc can be triangulated by drawing lines between the punctures, and we can keep track of the curve by counting the number of intersections of the curve with the edges of the triangulation.

When we do this we get Fibonacci numbers!

These intersection numbers gives coordinates for simple closed curves up to isotopy which may have multiple components provided that we minimize the number of intersections in a given isotopy class. (These might be called normal curves by analogy with normal surfaces; the condition that the number of intersections is min-

Figure 4: The braid traced out by the punctures over time, along with snapshots of the curve. Down is forwards in time.

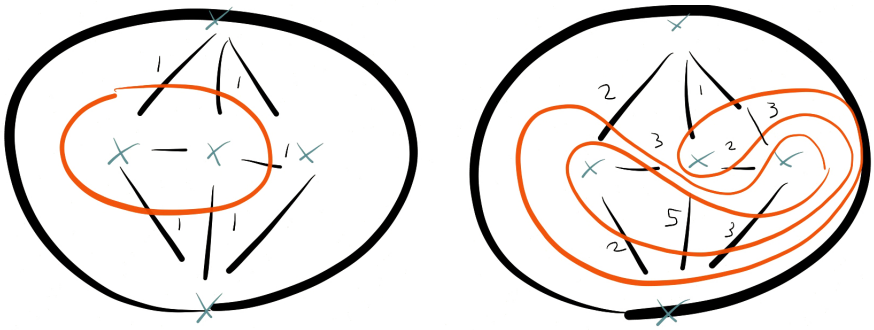


Figure 5: Counts of the number of intersections for zero and three permutations of the curve.

imal is a combinatorial analogue of being a geodesic, and we can find such curves by choosing a Riemannian metric with respect to which the edges of the triangulation are geodesics and choosing geodesic representatives.) This is because, given the number of intersections on the three edges of a triangle, there is a unique way to join them up in such a way that they could form part of a curve if it exists. This is obtained by pairing up intersections near corners.

Exercise 1.1. *Which triples of intersection numbers can be filled in to obtain a curve?*

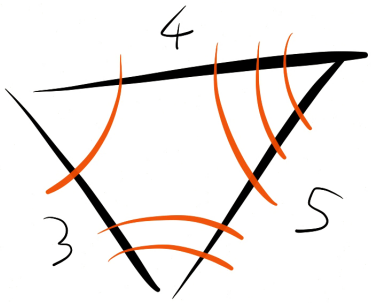


Figure 6: An example of intersection numbers and the corresponding unique curve.

Theorem 1.2. *Coordinates as specified above are unique provided that we restrict to curves with no null-homotopic components.*

One might wonder how general the process we used above to find a triangulation is.

Theorem 1.3. *Every surface with at least one puncture and at least one puncture on each boundary component admits a triangulation in which the punctures are the vertices.*

Triangulation here must be understood in a fairly general sense; for example, the edges of the triangle are allowed to be glued to each other.

Question from the audience: what kind of mathematics is this? How did we know to count the intersections without sign rather than doing the homological thing and counting, say, intersections mod 2?

Answer: there is a braid group B_3 involved which may be thought of as the mapping class group $\mathrm{MCG}(X) = \pi_0(\mathrm{Diff}(X))$ of the disc minus three punctures. (In general B_n may be thought of as either the mapping class group of the n-punctured disc or as $\pi_1(\mathrm{Conf}_n(\mathbb{R}^2))$.) In this particular case B_3 acts on certain conjugacy classes of π_1 of the disc minus 3 points (the ones which are represented by simple curves). Most conjugacy classes do not have this property. There is a nontrivial theorem we need here:

Theorem 1.4. *(Baer, Epstein, Freedman-Hass-Scott) Two simple curves are isotopic if and only if they are homotopic.*

Does this answer the question? There are a number of motivations for counting intersections without sign (perhaps Dehn was the first one to consider this?). As an application, we will get an algorithm for determining whether an element of the braid

group is the identity. Such an element must fix curves around the punctures, and in fact any element which fixes enough curves must be the identity.

We can now examine how the coordinates change when acted on by an element of the mapping class group. If we also act on the triangulation, we will get the same coordinates, but we want to see what the coordinates look like in the old triangulation.

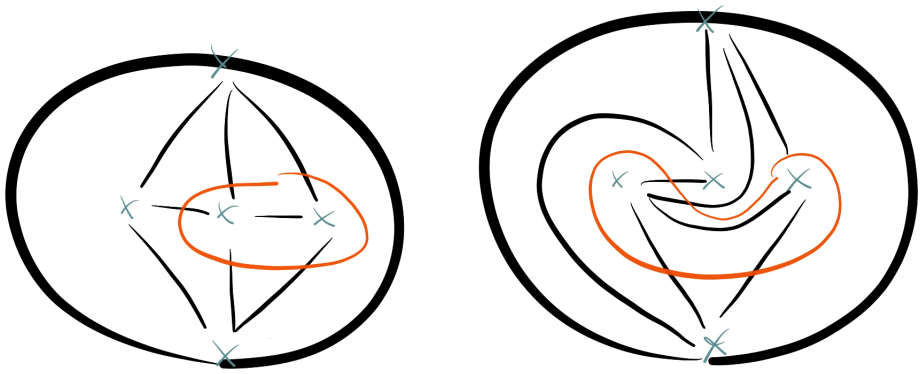


Figure 7: Dragging the triangulation along with the curve.

We can do this by changing the triangulation back. We will do this by flipping the middle edge in a quadrilateral composed of a pair of triangles (mutation?). This operation has many names; it can be interpreted in terms of rotation or the Whitehead move on binary trees.

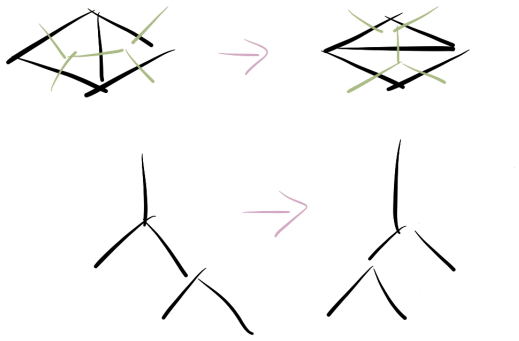


Figure 8: A quadrilateral flip / rotation / Whitehead move.

Theorem 1.5. *Any two triangulations of a surface (with at least two triangles) are related by a sequence of quadrilateral flips.*

It therefore suffices to study how coordinates change under a single flip. The only edge whose intersection number changes is the middle. If the intersection numbers of the outside edges are a, b, c, d and the intersection numbers of the old and new edges are e, f, then in fact

$$e + f = \max(a + c, b + d). \tag{1}$$

(There is a tropical cluster algebra here. This is a version of the Ptolemy relation.)

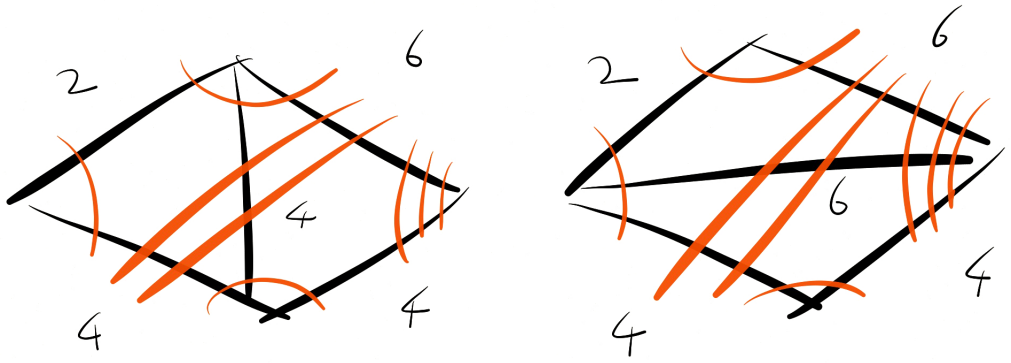


Figure 9: Intersection numbers before and after a quadrilateral flip.

To see this, note that there are six different types of curves: curves between two adjacent edges (four types) and curves crossing opposite sides. The adjacent curves add 1 to $\max(a + c, b + d)$ and to $e + f$ while the non-adjacent curves add 2 to $\max(a+c, b+d)$ and to $e+f$. Which of $a+c$ and $b+d$ is the maximum is determined by which pair of opposite sides is connected by some curve, since this can only occur for one pair.

Exercise 1.6. *Verify that we obtain the Fibonacci numbers in the example. Is there a reason why they appear?*

Exercise 1.7. *What is the asymptotic running time of the braid group algorithm?*

Exercise 1.8. *Choose coordinates for curves in the 3-punctured disc at random in some reasonable sense with some bound. What is the probability that you obtain a single curve? Two curves?*

Possible future topic: the relationship between $e + f = \max(a + c, b + d)$ and normal surfaces in 3-manifolds.

274 Curves on Surfaces, Lecture 2

Dylan Thurston

Fall 2012

2 Mapping class groups

Regarding the exercises from the previous lecture: it is cleaner to compactify and identify the point at infinity with the top and bottom punctures, so we work instead with a four-punctured sphere. Then there can be components of a curve which loop around one puncture or components which loop around two punctures (circumferences). This gives five different kinds of curves, and up to the action of the mapping class group every simple (multi)curve consists of copies of these five kinds of curves.

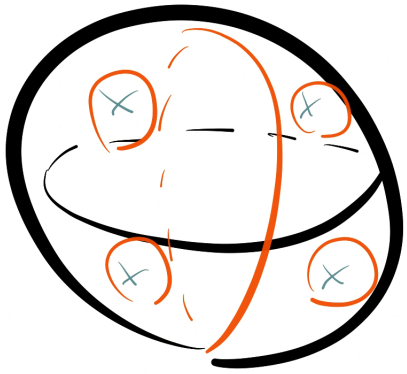
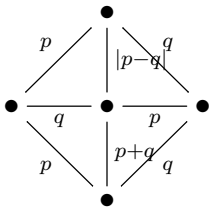


Figure 1: The five kinds of components of a simple multicurve on the four-punctured sphere.

In the first exercise, the three intersection numbers on the edges of a triangle must have the form $a+b, b+c, c+a$ where a, b, c are non-negative integers. This is equivalent to the three numbers summing to an even integer and satisfying the triangle inequality. Now, if all of the triangle inequalities among the coordinates are strict, then the curve necessarily has a component which loops around one puncture. Removing all such loops, we can work out that the coordinates must have the form



(1)

and the mapping class group acts essentially by Euclid's algorithm on p, q. This is not a coincidence; the mapping class group B_3 is related to $\mathrm{SL}_2(\mathbb{Z})$, which is in turn related to the Euclidean algorithm.

An easier case to handle first is the (orientation-preserving) mapping class group $\mathrm{MCG}^+(T^2) = \mathrm{Diff}^+(X)/\mathrm{Diff}_0(X)$ of the torus (where $\mathrm{Diff}_0(X)$ is the connected component of the group of diffeomorphisms containing the identity). This is the same as the mapping class group of the torus minus a point x, which acts on loops based at x.

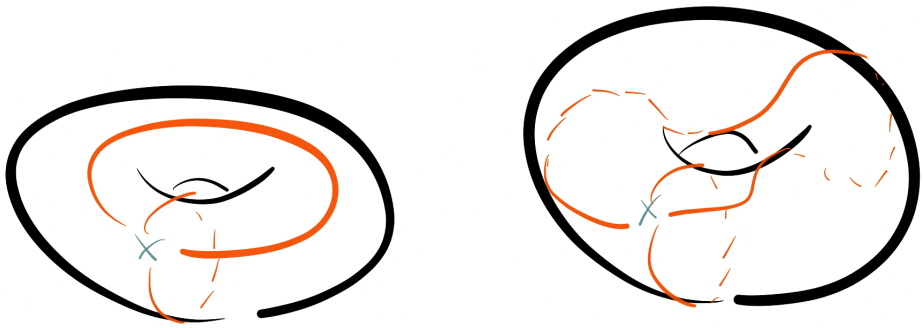


Figure 2: Two curves on a torus being acted on by the mapping class group.

This gives an action on the fundamental group, hence a homomorphism

$$\mathrm{MCG}(T^2, x) \to \mathrm{Aut}(\pi_1(T^2, x)) \cong \mathrm{Aut}(\mathbb{Z}^2) \cong \mathrm{GL}_2(\mathbb{Z}) \tag{2}$$

whose restriction to $\mathrm{MCG}^+(T^2, x)$ lands in $\mathrm{SL}_2(\mathbb{Z})$. This relies on the fact that $\pi_1(T^2, x)$ is abelian. In general we only get an action by outer automorphisms

$$\mathrm{MCG}(X) \to \mathrm{Out}(\pi_1(X)) \tag{3}$$

because the mapping class group acts on unbased loops, and changing the basepoint changes the corresponding automorphism of fundamental groups (based at different points) by conjugation by a path connecting the basepoints.

In the special case of surfaces Σ, the map $\mathrm{MCG}(\Sigma) \to \mathrm{Out}(\pi_1(\Sigma))$ is an isomorphism. This is false in general. It is also false that the mapping class group of a space is the mapping class group of a space minus a point x. There is only a commutative diagram

$$\begin{array}{ccc} \mathrm{MCG}(X) & \longrightarrow & \mathrm{Out}(\pi_1(X)) \\ \uparrow & & \uparrow \\ \mathrm{MCG}(X, x) & \longrightarrow & \mathrm{Aut}(\pi_1(X)) \end{array} \tag{4}$$

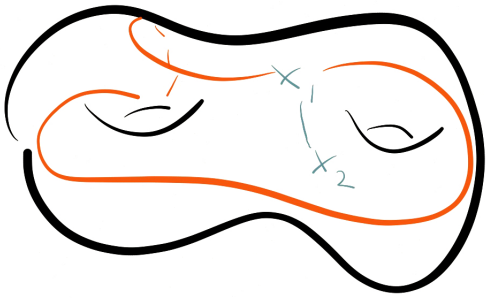


Figure 3: Changing basepoints on a surface with nonabelian fundamental group.

and the various maps in it are not isomorphisms in general, for example when X is a wedge of two circles $S^1 \vee S^1$. In this case $\mathrm{MCG}(X)$ has 8 elements (we can flip either of the two circles or switch the circles), $\mathrm{MCG}(X, x_0) = \mathrm{MCG}(X)$ if x_0 is the wedge point, and $\mathrm{MCG}(X, x_1)$ has two elements if x_1 lies on only one of the circles. Here $\pi_1 = \mathbb{Z} * \mathbb{Z}$ is not abelian, so the map $\mathrm{Aut}(\pi_1) \to \mathrm{Out}(\pi_1)$ is also not an isomorphism.

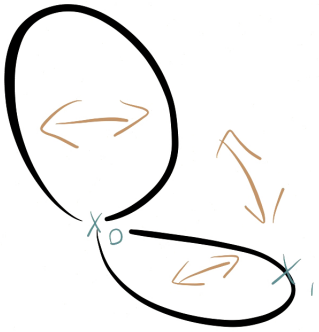


Figure 4: The wedge of two circles and the generators of its mapping class group.

Theorem 2.1. *$MCG(T^2) \cong MCG(T^2, x) \cong GL_2(\mathbb{Z})$ and $MCG^+(T^2) \cong MCG^+(T^2, x) \cong SL_2(\mathbb{Z})$.*

Regarding the distinction between $\mathrm{Aut}(\pi_1(X))$ and $\mathrm{Out}(\pi_1(X))$, we have the following result.

Theorem 2.2. *If G is a topological group, $\pi_1(G, e)$ is abelian.*

In particular T^2 is a topological group.

Exercise 2.3. *Find out what the Birman exact sequence is and report back next week.*

Back to curves. How can we describe curves on a punctured torus? We can triangulate and then count intersections as before. There are only three coordinates a, b, c, and there is a component around the puncture if various triangle inequalities are strict.

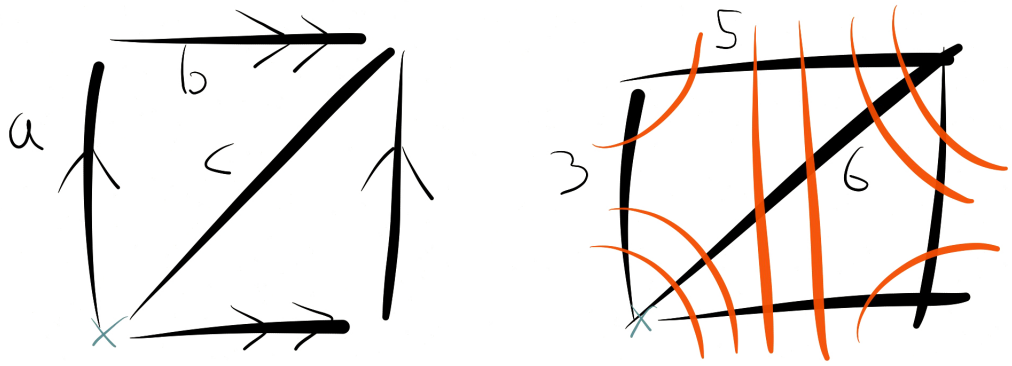


Figure 5: General coordinates for a curve on a punctured torus, and a specific curve with components around the puncture.

Removing all such components, assume WLOG that c is the greatest coordinate; then $c = a + b$. In this case curves are parameterized by two numbers (a, b), which in the universal cover $\mathbb{R}^2$ of the torus we can think of as given by a line with slope $\frac{b}{a}$ (when a, b are relatively prime).

The number of components of the corresponding curve is $\gcd(a, b)$ (the mapping class group acts by the Euclidean algorithm); consequently, we get a curve with a single component if and only if $\gcd(a, b) = 1$, so we can identify the simple closed curves on T^2 (not trivial, and not around the puncture) with the projective line $\mathbb{P}^1(\mathbb{Q})$ over $\mathbb{Q}$. The probability that $\gcd(a, b) = 1$ occurs asymptotically is

$$\frac{6}{\pi^2} = \left(\sum_{n=1}^{\infty} \frac{1}{n^2} \right)^{-1}. \tag{5}$$

(Heuristically this is because the probability that $\gcd(a, b) = n$ should be proportional to $\frac{1}{n^2}$.)

Exercise 2.4. *Explain how running Euclid's algorithm corresponds to changing the triangulation.*

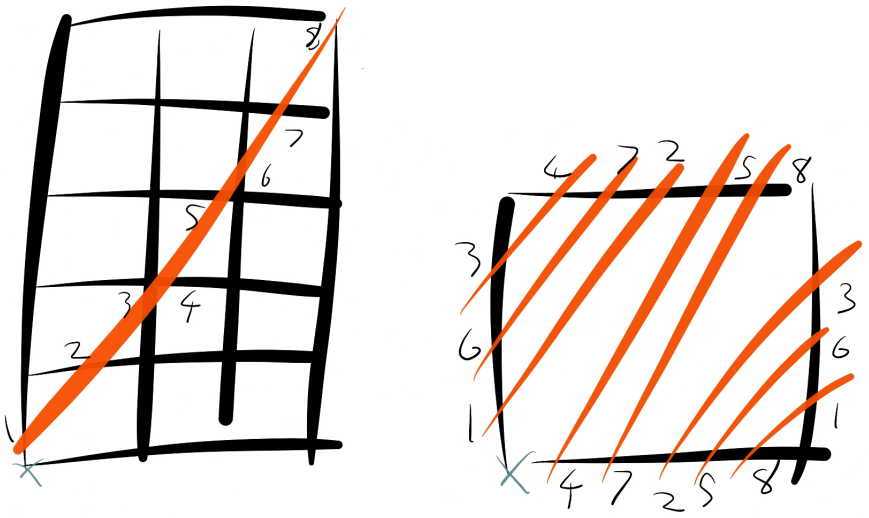


Figure 6: A line of slope $\frac{5}{3}$ in $\mathbb{R}^2$ and the corresponding curve on the torus.

Back to the braid group. What does $\mathrm{MCG}^+(T^2)$ have to do with the thrice-punctured disc or the four-punctured sphere? If Y denotes the four-punctured sphere, the double cover branched at the punctures $\tilde{Y}$ is a torus. Branched means that near the punctures the map looks locally like $z^2 \mapsto z$ in $\mathbb{C}$.

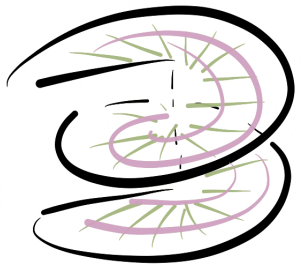


Figure 7: A local picture of the map $z^2 \mapsto z$ (given by projection down).

To obtain this double cover, skewer the torus by a line and quotient by rotation by 180° about the line. The corresponding cover is branched at the four points where we skewered the torus.

More algebraically, we quotiented by the element $-I$ in the mapping class group

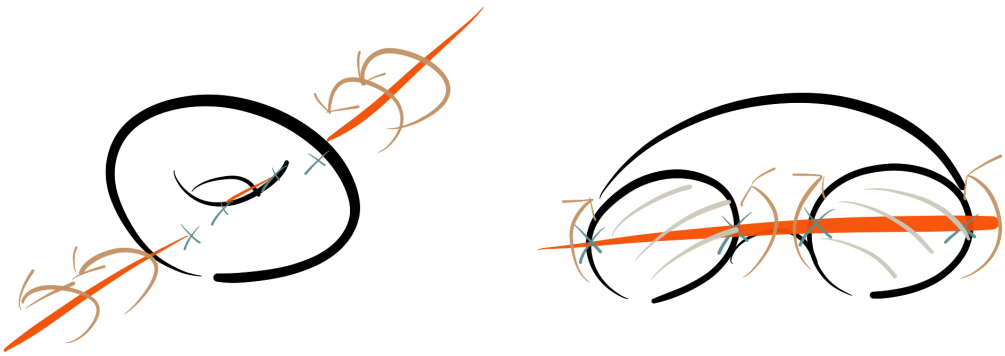


Figure 8: A skewered torus and a picture of the quotient, except that we still need to identify parts of the circles.

$\mathrm{SL}_2(\mathbb{Z})$. This branched cover descends to a map of mapping class groups

$$\mathrm{MCG}^+(T^2) \cong \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{PSL}_2(\mathbb{Z}) \cong \mathrm{MCG}^+(S^2, 1 \text{ pt}, 3 \text{ pts}). \tag{6}$$

The last mapping class group fixes one puncture pointwise and fixes the other three setwise. One way to see this action on three points is to look at the quotient $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{F}_2)$.

On the other hand, there is a map

$$B_3 \cong \mathrm{MCG}(D^2, 3 \text{ pts}) \to \mathrm{MCG}^+(S^2, 1 \text{ pt}, 3 \text{ pts}) \tag{7}$$

(where the first mapping class group fixes the boundary pointwise) and the claim is that this exhibits B_3 as a central extension of $\mathrm{PSL}_2(\mathbb{Z})$.

274 Curves on Surfaces, Lecture 3

Dylan Thurston
Notes by Qiaochu Yuan

Fall 2012

3 More about mapping class groups

Some background reading:

1. *Primer on Mapping Class Groups*, Farb and Margalit. Available online.
2. *Papers on Group Theory and Topology*, Dehn (introduction of Dehn-Thurston coordinates). Alex will be talking about this paper.
3. *Three-Dimensional Geometry and Topology*, Thurston Sr. Begins with a nice introduction to hyperbolic geometry. Available online.

Let S be a surface with $\chi(S) < 0$ and x a marked point. The Birman exact sequence is a short exact sequence

$$1 \to \pi_1(S, x) \to \mathrm{MCG}(S, x) \to \mathrm{MCG}(S) \to 1. \tag{1}$$

It can be iterated; for example, we can write down a short exact sequence

$$1 \to \pi_1(S \setminus 5 \text{ pts}) \to \mathrm{MCG}(S, 6 \text{ pts}) \to \mathrm{MCG}(S, 5 \text{ pts}) \to 1. \tag{2}$$

The map from $\pi_1(S, x)$ is the *point-dragging map* or *push map*. Given a curve $\gamma \in \pi_1(S, x)$, we want to send it to an element of $\mathrm{MCG}(S, x)$ which is trivial in $\mathrm{MCG}(S)$, hence it needs to be isotopic to the identity. It suffices to describe this isotopy. This isotopy will drag a neighborhood of the marked point x along γ and will be trivial outside a neighborhood of γ.

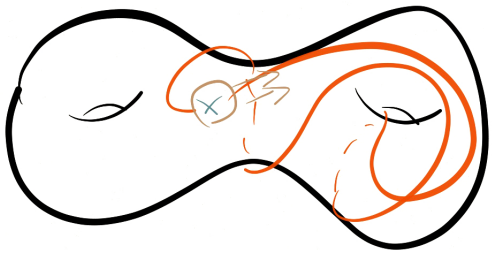


Figure 1: A marked point being pushed along a closed curve.

Why does this describe the entire kernel of the map $\mathrm{MCG}(S,x) \to \mathrm{MCG}(S)$? The general picture is as follows. For X any smooth manifold and $x \in X$ a marked point, there is a fibration

$$\mathrm{Diff}^+(X,x) \hookrightarrow \mathrm{Diff}^+(X) \to X \tag{3}$$

where the map $\mathrm{Diff}^+(X) \to X$ sends a diffeomorphism to the image of x. (A fibration behaves like a fiber bundle. The crucial property is a lifting property: in particular, any path in X lifts to a path in $\mathrm{Diff}^+(X)$.) This fibration induces a long exact sequence in homotopy

$$...\pi_1(\mathrm{Diff}^+(X)) \to \pi_1(X) \to \pi_0(\mathrm{Diff}^+(X,x)) \to \pi_0(\mathrm{Diff}^+(X)) \to \pi_0(X). \tag{4}$$

But $\pi_0(\mathrm{Diff}^+(X)) = \mathrm{MCG}(X)$ and $\pi_0(\mathrm{Diff}^+(X,x)) = \mathrm{MCG}^+(X,x)$, and $\pi_0(X)$ is a point when X is connected. The next term in the long exact sequence is a map $\pi_1(X) \to \pi_0(\mathrm{Diff}^+(X,x))$.

Theorem 3.1. *(Hamstrom) Let S be a surface with $\chi(S) < 0$. Then $\pi_1(\mathit{Diff}^+(X))$ is trivial. In fact, the connected component of the identity in $\mathit{Diff}^+(X)$ is contractible.*

This is an aspect of hyperbolic geometry. The same is true for higher-dimensional hyperbolic manifolds; this is an aspect of Mostow rigidity. (But Mostow rigidity is false for hyperbolic surfaces.)

What happens when $S = T^2$? We claimed that the map $\mathrm{MCG}(T^2,x) \to \mathrm{MCG}(T^2)$ is an isomorphism. The long exact sequence ends

$$...\pi_1(\mathrm{Diff}^+(T^2)) \to \pi_1(T^2) \to \mathrm{MCG}(T^2,x) \to \mathrm{MCG}(T^2) \to 1 \tag{5}$$

so the map $\pi_1(T^2) \to \mathrm{MCG}(T^2,x)$ needs to be trivial. There is a map $T^2 \to \mathrm{Diff}_0(T^2)$ given by T^2 acting on itself by translation, and it is a difficult theorem that this is a homotopy equivalence. (This can be proven by removing a point, which makes the Euler characteristic -1, and applying the big theorem above.) Consequently

$$\pi_1(\mathrm{Diff}^+(T^2)) = \pi_1(\mathrm{Diff}^0(T^2)) \cong \pi_1(T^2). \tag{6}$$

Similarly, T^2 admits an action by affine linear maps, and this is a homotopy equivalence to $\mathrm{Diff}(T^2)$.

In summary, the end of the long exact sequence looks like

$$\begin{array}{ccccccccccc}
\pi_1(\mathrm{Diff}^+(T^2,x)) & \longrightarrow & \pi_1(\mathrm{Diff}^+(T^2)) & \longrightarrow & \pi_1(T^2) & \longrightarrow & \mathrm{MCG}(T^2,x) & \longrightarrow & \mathrm{MCG}(T^2) & \longrightarrow & 1 \\
\downarrow\cong & & \downarrow\cong & & \downarrow\cong & & \downarrow\cong & & \downarrow\cong & & \downarrow\cong \\
1 & \longrightarrow & \mathbb{Z}^2 & \xrightarrow{\cong} & \mathbb{Z}^2 & \xrightarrow{0} & \mathrm{SL}_2(\mathbb{Z}) & \xrightarrow{\cong} & \mathrm{SL}_2(\mathbb{Z}) & \longrightarrow & 1
\end{array} \tag{7}$$

where $\cong$ denotes an isomorphism.

More generally, if G is a connected Lie group, we get a map $G \to \mathrm{Diff}_0(G)$ coming from the action of G on itself by translation, and we also get a map in the other direction coming from evaluation. This is not a homotopy equivalence in general. When $G = \mathrm{SU}(2)$ we know that $\mathrm{SU}(2) \cong S^3$, and $\mathrm{Diff}^+(S^3)$ is homotopy equivalent to $\mathrm{SO}(4)$ (the Smale conjecture, proved by Hatcher).

Recall that last time we skewered a torus (quotiented it by the central element $-I$ in $\mathrm{MCG}(T^2) \cong \mathrm{SL}_2(\mathbb{Z})$) to obtain a double cover $T^2 \to S^2$ branched at 4 points. The claim was that this showed

$$\mathrm{MCG}(S^2, 1 \text{ pt}, 3 \text{ pts}) \cong \mathrm{PSL}_2(\mathbb{Z}). \tag{8}$$

(The 1 point is the identity in T^2 regarded as a group and the 3 points are the non-identity points of order 2.)

What is the mapping class group of S^2 fixing four points pointwise? This is the congruence subgroup $\Gamma(2)$, which consists of the image of the kernel of the map $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/2\mathbb{Z})$ in $\mathrm{PSL}_2(\mathbb{Z})$. It is in fact the free group $\mathbb{Z} * \mathbb{Z}$ on two generators.

The relationship to the braid group B_3 comes from the map

$$(D^2, 3 \text{ pts}) \to (S^2, 3 \text{ pts}, 1 \text{ pt}) \tag{9}$$

given by identifying the boundary to a point (which becomes the fourth marked point).

Figure 2: A 3-punctured disc getting its boundary identified to form a 4-punctured sphere.

The mapping class group B_3 of $(D^2, 3 \text{ pts})$ (fixing the boundary pointwise) has a center generated by *Dehn twist* along a boundary curve. As a braid it is given by the *full twist*. The image of Dehn twist in $\mathrm{MCG}(S^2, 3 \text{ pts}, 1 \text{ pt})$ is trivial (we can untwist). Thus we obtain an exact sequence

$$1 \to \mathbb{Z} \to \mathrm{MCG}(D^2, 3 \text{ pts}, \partial D^2) \to \mathrm{MCG}(S^2, 3 \text{ pts}, 1 \text{ pt}) \to 1 \tag{10}$$

showing that B_3 is a central extension of $\mathrm{PSL}_2(\mathbb{Z})$.

Figure 3: A full twist and a half twist.

Recall that before we were permuting curves on the thrice-punctured disc and, looking at Dehn-Thurston coordinates, we saw the Fibonacci numbers appear. This can now be explained as follows. The element of the mapping class group we were applying was a braid in B^3 whose image in $\mathrm{PSL}_2(\mathbb{Z})$ is given by the matrix $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$.

Exercise 3.2. *Verify this.*

Hint: look at how the braid group generators lift to the torus. They can be thought of as Dehn twists.

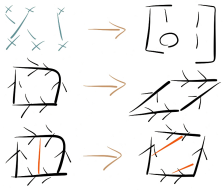


Figure 4: Some hints.

Dehn twists in general look like the following: if C is a simple closed curve on S, the Dehn twist $T_C \in \mathrm{MCG}(S)$ rotates an annular neighborhood $[0, 1] \times C$ of C as follows: $\{t\} \times C$ is rotated by $2\pi t$.

Figure 5: Dehn twist around a curve C.

Figure 6: Another picture of a Dehn twist.

Question from the audience: is this the same as the push map?

Answer: no. The push map gives a trivial element of the mapping class group. However, there is a relationship. Let γ is a simple closed curve and C_1, C_2 curves which bound an annular neighborhood of γ.

Exercise 3.3. *Push*$(\gamma) = T_{C_1} \circ T_{C_2}^{-1}$.

Theorem 3.4. *(Lickorish, ...) Let S be a closed surface. Then $MCG^+(S)$ is generated by Dehn twists.*

Dehn twists cannot generate the mapping class group of a surface with marked points because they cannot permute the marked points. With marked points, the Dehn twists instead generate the *pure* mapping class group (the subgroup fixing the marked points pointwise).

The basic invariant of an element $M \in \mathrm{SL}_2(\mathbb{Z})$ up to conjugacy is its trace (this determines its characteristic polynomial). If $\mathrm{tr}(M) = 2$ then

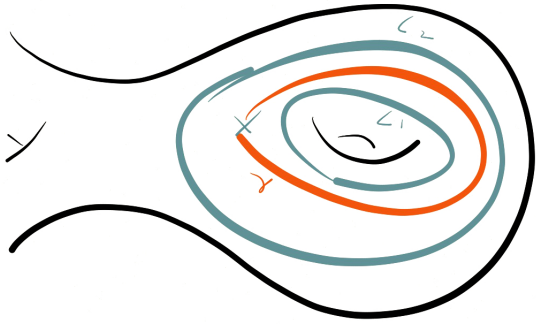


Figure 7: Dehn twists and the push map.

$$M = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \tag{11}$$

for some x, and similarly if $\mathrm{tr}(M) = -2$ then

$$M = \begin{bmatrix} -1 & x \\ 0 & -1 \end{bmatrix}. \tag{12}$$

These are the *parabolic elements*, and they look like Dehn twists when acting on the torus.

If $|\mathrm{tr}(M)| > 2$ then M has 2 distinct real eigenvalues, and iterating M we obtain exponential growth. (In the particular case above, $M = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ and the eigenvalues are ϕ^2, φ^2 where ϕ, φ are the golden ratios.) These are the *hyperbolic elements*.

If $|\mathrm{tr}(M)| < 2$ then M is in fact torsion. These are the *elliptic* or *periodic* elements. The two basic possibilities are

$$M = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \tag{13}$$

and variants.

Exercise 3.5. *Which braids do these correspond to?*

274 Curves on Surfaces, Lecture 4

Dylan Thurston

Fall 2012

4 Hyperbolic geometry

Last time there was an exercise asking for braids giving the torsion elements in $\mathrm{PSL}_2(\mathbb{Z})$. A 3-torsion element can be obtained by cyclically permuting punctures (a one-third-twist?), and a 2-torsion element can be obtained by swapping two punctures (a half-twist).

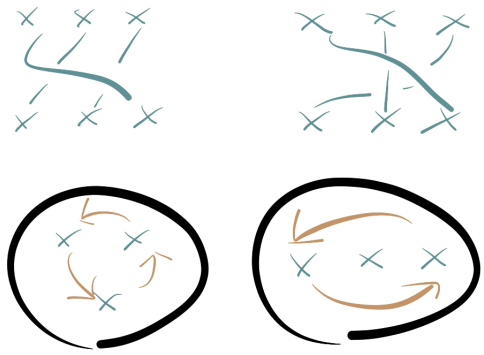


Figure 1: Braids which are torsion in $\mathrm{PSL}_2(\mathbb{Z})$.

Last time we also classified elements of $\mathrm{MCG}(T^2)$ as either periodic, parabolic, or hyperbolic. This classification generalizes to other surfaces; it is called the Nielsen-Thurston classification. We will get back to this later.

First, the hyperbolic plane. It is the unique complete simply-connected Riemannian surface with constant curvature -1, but this is not useful for computations.

The *Poincaré disk model* of the hyperbolic plane is the open disc $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$ with the metric

$$ds^2 = 4\frac{dx^2 + dy^2}{(1 - r^2)^2} \tag{1}$$

where $r^2 = x^2 + y^2$.

Since the metric is always a scalar multiple of the standard Euclidean metric, this angle is conformal with the Euclidean metric, so angles agree with Euclidean angles (even if lengths do not agree with Euclidean lengths). Geodesics are circles perpendicular to the boundary (including circles of infinite radius, or lines); in particular, there is a unique geodesic through any two points.

A geodesic triangle has the property that the sum of its angles is less than π. In fact,

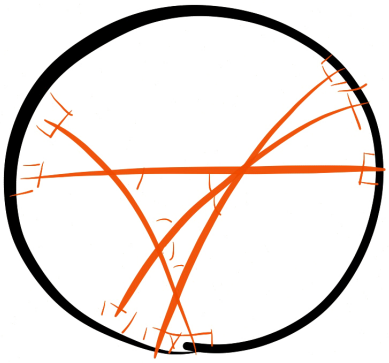


Figure 2: Geodesics in the disk model.

$$\text{Area(triangle)} = \pi - \text{sum of angles.} \tag{2}$$

Exercise 4.1. *The angle sum defect π − sum of angles is additive on triangles.*

The *Poincaré half-plane model* is the upper half-plane $\{(x, y) \in \mathbb{R}^2 : y > 0\}$ with the metric

$$ds^2 = \frac{dx^2 + dy^2}{y^2}. \tag{3}$$

Geodesics are circles perpendicular to the boundary (including circles of infinite radius). This metric is also conformal with the Euclidean metric.

Instead of thinking directly about these metrics, it is better to think about automorphisms (that is, about isometries). The automorphisms we want should be conformal; that is, they should preserve (Euclidean) angles. This is equivalent to complex analytic with derivative not equal to zero anywhere. We will look for such automorphisms within the group of Möbius transformations

$$z \mapsto \frac{az + b}{cz + d}, a, b, c, d \in \mathbb{C}, ad - bc \neq 0. \tag{4}$$

This is precisely the group of conformal automorphisms of the Riemann sphere. Abstractly, this is the group $\mathrm{PGL}_2(\mathbb{C})$ (which naturally acts on $\mathbb{CP}^1$), and consequently it admits a morphism from $\mathrm{GL}_2(\mathbb{C})$ (so Möbius transformations compose like matrices) sending $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ to the above.

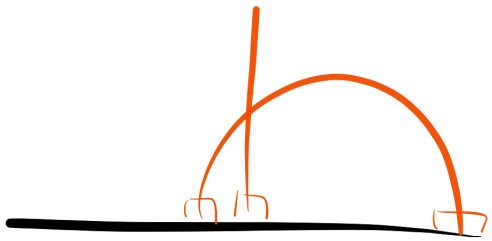


Figure 3: Geodesics in the half-plane model.

Möbius transformations have many nice properties. For example, they preserve circles (including of infinite radius).

Some Möbius transformations are relatively easy to understand. Those of the form $z \mapsto az + b$ describe translations, scalings, and rotations. The only additional Möbius transformation needed to generate the entire group is $z \mapsto \frac{1}{z}$, which is closely related to circle inversion $z \mapsto \frac{1}{\bar{z}}$. Inversion has many nice properties: it sends circles inside the unit circle to circles outside the unit circle and sends circles intersecting the unit circle to circles intersecting the unit circle.

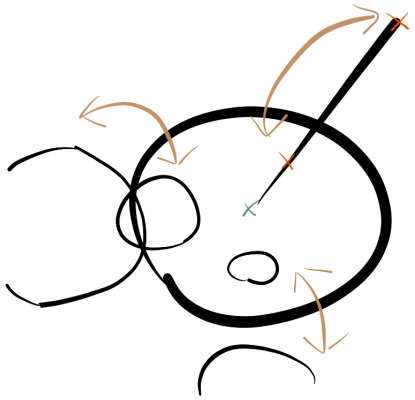


Figure 4: Circle inversion.

We want to consider Möbius transformations which in addition preserve the open disk or the upper half plane.

Theorem 4.2. *The orientation-preserving isometries of the upper half-plane are precisely $PSL_2(\mathbb{R}) \subset PGL_2(\mathbb{C})$.*

Proof. Since $\mathbb{H}^2$ has constant curvature, it has an isometry taking every pair of a point and a tangent vector to another point and a tangent vector. Every orientation-preserving isometry is conformal, hence complex-analytic, and furthermore extends to a conformal automorphism $\mathbb{CP}^1 \to \mathbb{CP}^1$ by the Schwarz reflection principle, hence must be a Möbius transformation. Moreover, it must preserve $\mathbb{RP}^1 \subset \mathbb{CP}^1$, so lies in $\mathrm{PGL}_2(\mathbb{R})$.

There is a commutative diagram of inclusions

$$\begin{array}{ccc} \mathrm{PSL}_2(\mathbb{R}) & \longrightarrow & \mathrm{PSL}_2(\mathbb{C}) \\ \downarrow & & \downarrow \\ \mathrm{PGL}_2(\mathbb{R}) & \longrightarrow & \mathrm{PGL}_2(\mathbb{C}) \end{array} \tag{5}$$

and the inclusion $\mathrm{PSL}_2(\mathbb{C}) \to \mathrm{PGL}_2(\mathbb{C})$ is an isomorphism (we can always scale by the square root of the determinant) but the inclusion $\mathrm{PSL}_2(\mathbb{R}) \to \mathrm{PGL}_2(\mathbb{R})$ is not; the latter has two connected components, one consisting of matrices with positive determinant (which $\mathrm{PSL}_2(\mathbb{R})$ maps to isomorphically) and one consisting of matrices with negative determinant.

The elements of $\mathrm{PGL}_2(\mathbb{R})$ of negative determinant take the upper half-plane to the lower half-plane, so the elements of $\mathrm{PSL}_2(\mathbb{R})$ are the ones we want. □

We can now obtain the metric on the upper half-plane we wrote down earlier as follows: first assume that it is $ds^2 = dx^2 + dy^2$ at the point $(0, 1)$ and find an automorphism taking $(0, 1)$ to another point (x, y). It suffices to take $z \mapsto yz + x$, and we want this to be an isometry, which in fact forces $ds^2 = \frac{dx^2+dy^2}{y^2}$ by inspecting the Jacobian.

A similar idea works for the metric on the disk; alternately, there is a Möbius transformation taking the upper half-plane to the disk given by

$$z \mapsto \frac{z - i}{z + i}. \tag{6}$$

To see this, note that it takes the boundary of the upper half-plane to the boundary of the open disc and takes i to 0. (Multiplying by i gives a map which fixes ± 1 and which sends ∞ to i.) This gives us a description of the Möbius transformations fixing the disk by conjugating by the above map; explicitly, they have the form

$$z \mapsto e^{i\theta} \frac{z - a}{\bar{a}z - 1} \tag{7}$$

where a lies in the disk.

More generally, the Riemann mapping theorem asserts that for any open simply-connected subset U of $\mathbb{C}$ there exists a unique biholomorphic map from U to the open disc sending a particular point and tangent vector to a particular point and tangent vector in the open disc. This gives us a metric on U.

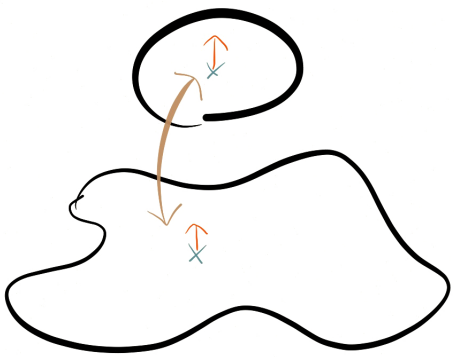


Figure 5: The Riemann mapping theorem.

Moving away from complex analysis, we can write down a different kind of model of the hyperbolic plane which generalizes better to higher dimensions as follows. There is an exceptional isomorphism

$$\mathrm{PSL}_2(\mathbb{R}) \cong \mathrm{SO}^+(2,1) \tag{8}$$

where the RHS describes the group of all linear transformations of $\mathbb{R}^3$ preserving the quadratic form $x^2+y^2-z^2$ which have determinant 1 and which maps the upper half of the cone $x^2+y^2=z^2$ to itself.

This can be seen as follows. $\mathrm{SL}_2(\mathbb{R})$ has a natural representation on $\mathbb{R}^2$. This gives a representation of $\mathrm{PSL}_2(\mathbb{R})$ on $\mathrm{Sym}^2(\mathbb{R})$, or equivalently on 2×2 real symmetric matrices as follows:

$$\begin{bmatrix} p & q \\ q & r \end{bmatrix} \mapsto \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} p & q \\ q & r \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix}. \tag{9}$$

This action preserves the determinant $pr-q^2$. Writing $x=q, p=z+y, q=z-y$ this gives the quadratic form $z^2-y^2-x^2$. Alternately, $\mathrm{PSL}_2(\mathbb{R})$ has an adjoint action on its Lie algebra $\mathfrak{sl}_2(\mathbb{R})$, which is 3-dimensional. This action preserves the Killing form on $\mathrm{sl}_2(\mathbb{R})$, which also has signature $(2,1)$ as above.

$\mathrm{SO}(2,1)$ preserves the two-sheeted hyperboloid $\{(x,y,z) \in \mathbb{R}^3 : x^2+y^2-z^2 = -1\}$, and $\mathrm{SO}^+(2,1)$ preserves the sheet $z > 0$. This has an induced metric

$$ds^2 = dx^2 + dy^2 - dz^2 \tag{10}$$

which gives another model of the hyperbolic plane, the *hyperboloid model*. This is a pleasant model for computations because of the lack of division.

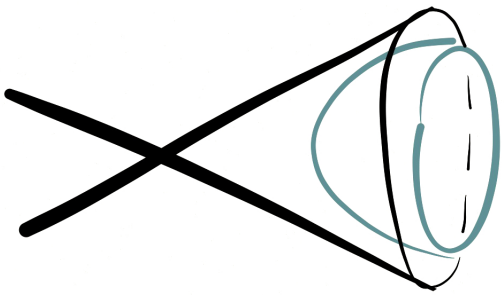


Figure 6: The hyperboloid model.

We need to check that the above metric is actually Riemannian; as a metric on $\mathbb{R}^3$ it is Lorentzian. $\mathrm{SO}^+(2,1)$ acts transitively on the hyperboloid, so to check the signature of the metric it is enough to check the signature at a point. We will use the point $(0,0,1)$. The tangent plane at this point is the xy-plane, and the induced metric is Riemannian as desired.

The hyperboloid may be thought of as a sphere of radius i. This is not as silly as it sounds; it turns many trigonometric identities into hyperbolic trigonometric identities.

We now have three models for the hyperbolic plane. To do computations in the future, we will choose whichever model makes our computations easiest.

274 Curves on Surfaces, Lecture 5

Dylan Thurston
Notes by Qiaochu Yuan

Fall 2012

5 Ideal polygons

Previously we discussed three models of the hyperbolic plane: the Poincaré disk, the upper half-plane, and the hyperboloid. We did not describe geodesics in the hyperboloid model, but they are given by intersections of planes through the origin.

Figure 1: A plane giving a geodesic on the hyperboloid.

This is precisely the same way we obtain geodesics on the sphere.

Figure 2: A plane giving a geodesic on the sphere.

To relate the models, in the hyperboloid model we can look at the hyperboloid from the tip of the other hyperboloid, and we get the disk model. (For an observer at the origin, the corresponding model is the Klein disk model, which is not conformal. Its geodesics are straight chords. This is analogous to how stereographic projection is conformal on the sphere, but projection from a random point is not.)

$\mathbb{H}^2$ has an *ideal boundary*. For example, in the disk model the boundary of the disk is this boundary. It is often convenient to complete $\mathbb{H}^2$ to include this boundary, obtaining $\overline{\mathbb{H}^2}$. The metric does not extend, but we still get a topological space. $\overline{\mathbb{H}^2}$ is obtained from $\mathbb{H}$ by adding *ideal points*, or equivalence classes of geodesic rays. Most pairs of geodesics get exponentially far away from each other, but sometimes two geodesic rays will approach each other; intuitively they are approaching the same point of the boundary. (We do not get this behavior in the Euclidean plane, where two geodesic rays can be parallel.)

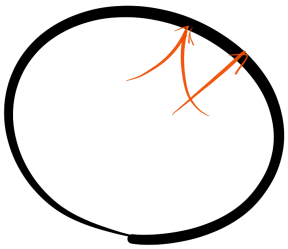


Figure 3: Some equivalent and inequivalent geodesic rays.

A *triangle* in $\mathbb{H}^2$ can be constructed from three geodesics. If α, β, γ are the corresponding angles of the triangle, the claim was that the area of the triangle is $\pi - (\alpha + \beta + \gamma)$. This is a corollary of the Gauss-Bonnet theorem, one intuitive statement of which is that the total turning of a curve bounding a disk D in a Riemannian manifold is

$$2\pi - \int_D K \, dA \tag{1}$$

where K is the curvature.

Exercise 5.1. *Measure the curvature of some surface. (A leaf? An orange peel? Can you find a leaf with positive curvature?)*

The Gauss-Bonnet theorem can be used to compute the area of a sphere. Considering a great circle on a sphere of radius 1, we compute that 2π is half the surface area of such a sphere, hence the total surface area is 4π. When applied to a

geodesic triangle, the total turning of a geodesic triangle with angles α, β, γ is given by $(\pi - \alpha) + (\pi - \beta) + (\pi - \gamma)$, which is 2π plus the area of the triangle.

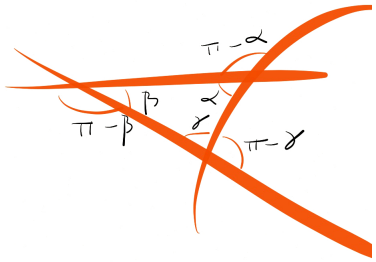


Figure 4: The turning angles of a hyperbolic triangle.

In particular, this area is positive, so $\alpha + \beta + \gamma < \pi$. In addition, we conclude that the area of a geodesic triangle is always strictly less than π. As the three vertices approach the ideal boundary, the corresponding angles become very small, so the area approaches π. In the limiting case, the three vertices are on the ideal boundary, and we get an *ideal triangle*. Since geodesics meet the boundary at right angles, the angles in an ideal triangle are all equal to 0, so an ideal triangle has area π.

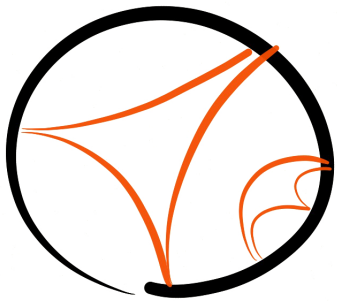


Figure 5: Some ideal triangles.

Ideal triangles turn out to be simpler and easier to understand than ordinary triangles. For example:

Proposition 5.2. *Any two ideal triangles are related by an orientation-preserving isometry.*

Proof. We will work in the upper half-plane model. In this model the ideal boundary is $\mathbb{RP}^1$ acted on by the isometry group $\mathrm{PSL}_2(\mathbb{R})$. We know that $\mathrm{PGL}_2(\mathbb{R})$ acts triply transitively. A short way to see this is to try to send any three points a, b, c to $0, 1, \infty$. This gives, uniquely,

$$x \mapsto \frac{(x-a)(b-c)}{(x-c)(b-a)}. \tag{2}$$

By choosing an appropriate ordering of a, b, c we can arrange for this isometry to lie in $\mathrm{PSL}_2(\mathbb{R})$. □

It follows that the *moduli space of ideal triangles* (the space of ideal triangles up to orientation-preserving isometry) is a point. What about *ideal quadrilaterals*? All of the interior angles are still 0, so by Gauss-Bonnet any ideal quadrilateral has area 2π; alternately, an ideal quadrilateral is the union of two ideal triangles. Now not all ideal quadrilaterals are equivalent. By the previous proof, we can send three of the vertices to $0, 1, \infty$, and the fourth point cannot be moved because the pointwise stabilizer of $0, 1, \infty$ is the identity.

This gives an invariant of ideal quadrilaterals: take three of the points a, b, c to $0, 1, \infty$ and look at where the fourth point d goes. This is

$$d \mapsto \frac{(d-a)(b-c)}{(d-c)(b-a)} \tag{3}$$

or the *cross-ratio* of a, b, c, d. This can be used to describe the moduli space of ideal quadrilaterals. In general, we expect the moduli space of ideal n-gons to have dimension $n-3$. We can think of the n as the number of parameters describing vertices and the 3 as the dimension of $\mathrm{PSL}_2(\mathbb{R})$.

There is something funny going on here. If we think of an ideal quadrilateral as being composed of two ideal triangles, we can slide the ideal triangles against each other by moving the opposite vertex, and this does not change the triangles themselves up to isometry. Ordinary Euclidean triangles do not have this property.

We would like to make ideal geometry look more like ordinary geometry. One way to do this would be to associate a number to an edge of an ideal polygon, which we want to think of as its length. It has infinite length in the hyperbolic metric, so we will have to be clever. First, what do circles (points equidistant to a given point) look like in the hyperbolic plane?

Proposition 5.3. *Circles in the disk model or half-plane model are still ordinary circles.*

Proof. In the disk model, take the center of the circle to be the origin. Isometries include rotational symmetries, so the circle must be an ordinary Euclidean circle.

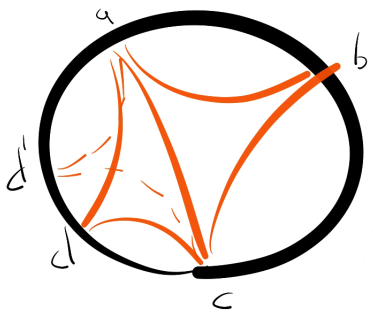


Figure 6: Sliding triangles in an ideal quadrilateral.

Conformal maps preserve Euclidean circles, so the same is true for any center, and the same is true for the half-plane. □

However, the Euclidean center is usually not the center. As a circle gets closer to the boundary, its center also gets closer to the boundary.

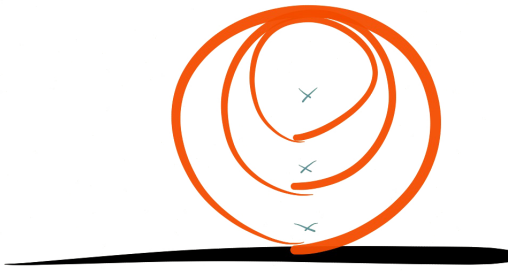


Figure 7: Centers getting closer to the boundary.

As we send the circle to the boundary and send the radius to ∞ so that the circle continues to pass through a given point, we obtain a *horocycle*. In the disk and half-plane models these look like Euclidean circles tangent to the boundary (and in the latter case this includes circles of infinite radius). The corresponding construction in the Euclidean plane gives a line, but horocycles are not geodesics in the hyperbolic plane. They can be thought of as circles centered at an ideal point.

One property of Euclidean circles is that a geodesic through the center is perpendicular to the boundary. This is still true for hyperbolic circles. It is also true that a horocycle is perpendicular to any geodesic approaching its ideal point.

We now know how to give meaning to a circle tangent to the boundary.

Exercise 5.4. *What is the meaning of a circle intersecting the boundary at an arbitrary angle?*

General hint: find the right model and put the interesting point in a convenient place.

We now want to describe *decorated* ideal polygons, which are given by an ideal polygon together with a choice of horocycle around each vertex. This gives us a way to measure length: we can take the length of an edge of a decorated ideal polygon to be the oriented distance between the horocycles around the corresponding vertices (so it is negative if the horocycles overlap).

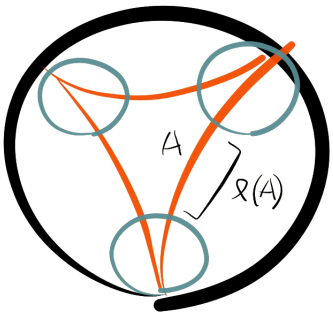


Figure 8: A decorated triangle and a length.

Lemma 5.5. *Any three choices $\ell(A), \ell(B), \ell(C) \in \mathbb{R}$ of lengths is realizable by a unique decorated ideal triangle (up to isometry).*

Proof. We first consider the case where all three lengths are equal to 0. This means that the horocycles are all tangent to each other as well as to the ideal boundary. The corresponding configuration of four circles is unique up to fractional linear transformations.

Alternately, we can work with an ideal triangle in the upper half-plane with one of the vertices at infinity. One of the horocycles is then a line parallel to the x-axis, and the other two horocycles are uniquely determined by this.

Figure 9: Four tangent circles.

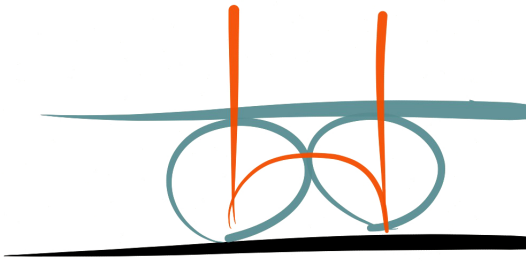


Figure 10: Three tangent horocycles in the upper half-plane.

If from here we move one of the horocycles by some distance along one geodesic, we move it by the same distance along the other geodesic. If we move all three horocycles by distances a, b, c, we get distances $a+b, b+c, c+a$, and this linear transformation is invertible. □

We can also decorate ideal quadrilaterals and assign lengths to their edges. This is equivalent to gluing decorated ideal triangles with matching lengths, and now there is no sliding provided that all lengths are fixed. This leads to a parameterization of decorated ideal polygons.

Theorem 5.6. *Given a decorated ideal polygon and a triangulation, the lengths between horocycles of edges determine the polygon (up to isometry).*

Figure 11: The general case.

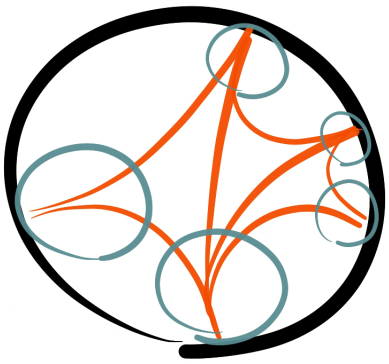


Figure 12: A decorated and triangulated pentagon.

This gives us some number of parameters describing a decorated ideal n-gon. First there are n edges to consider. A triangulation adds $n-3$ edges, so we get $2n-3$ parameters. The moduli space of ideal n-gons has dimension $n-3$, and the difference between these are the n parameters describing the horocycles.

Next time we will discuss how these parameters change under change of triangulations.

274 Curves on Surfaces, Lecture 6

Dylan Thurston

Fall 2012

6 Dehn-Thurston coordinates (Alex)

We want to study the action of the mapping class group on isotopy classes of curves on a surface. Ideally this action should be faithful. Dehn-Thurston coordinates are a way to parameterize isotopy classes of curves.

Let S_g be a compact orientable surface of genus g with negative Euler characteristic, possibly with boundary. (The Euler characteristic condition only excludes the sphere, the torus, the cylinder, and the disks.) We will consider *multicurves* on S_g, which are 1-dimensional submanifolds such that no component bounds a disk and such that no component is homotopic to an arc on the boundary. This gives a collection of non-intersecting, non-self-intersecting, non-null-homotopic curves.

We parameterize multicurves by first choosing a decomposition into pairs of pants.

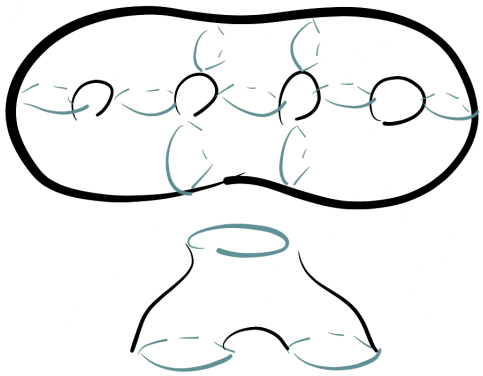


Figure 1: A pair of pants decomposition.

We want to know the intersections of a multicurve with the boundary of each pair of pants. This gives a collection of intersection numbers $m_1, ...m_n$. Additionally, we have N twisting numbers which tell us how to glue the pairs of pants together. If S_g has no boundary, then $N = 3g - 3$.

Definition The *geometric intersection number* of two curves is

$$(\gamma, \delta) = \min_{c,d} |c \cap d| \tag{1}$$

where c, d are curves isotopic to γ, δ respectively.

The claim we need for these intersections to determine a multicurve is that up to isotopy preserving the boundary componentwise, a multicurve on a pair of pants is

determined by its intersection numbers with the boundary (except for components parallel to a boundary component).

Fix intervals on the boundary components of the pants; these are *windows*. We will require that our curves only intersect the boundaries in windows. This only gives a few possibilities for the components of a multicurve: it can either connect adjacent windows, loop around to connect a window with itself, or loop around a leg (parallel to a boundary component).

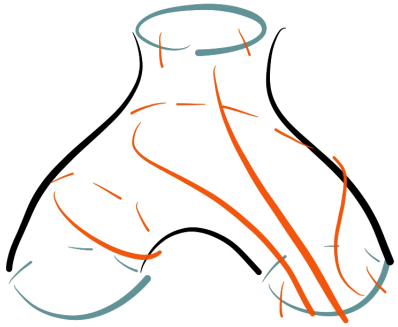


Figure 2: Curves on a pair of pants.

To define twisting numbers, we will now decompose S_g into pairs of pants and cylinders (and again fix windows).

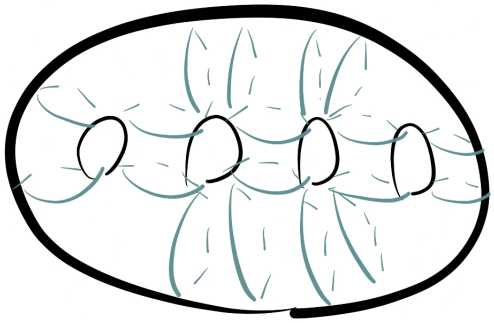


Figure 3: Pairs of pants and cylinders.

In a given cylinder, the twisting number is then the geometric intersection of a

multicurve with either of two curves connecting the boundaries of the windows, with sign determined by handedness.

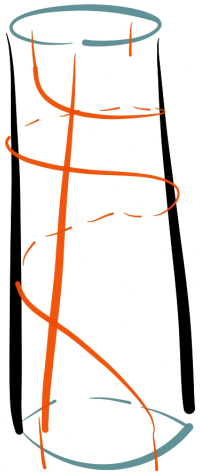


Figure 4: Twisting numbers in a cylinder.

We also need twisting numbers counting components parallel to boundary components.

In summary, we parameterize multicurves by elements of the set $\mathbb{Z}_{\geq 0}^{3g-3} \times \mathbb{Z}^{3g-3}$ quotiented by the equivalence relation $(0, x) \sim (0, -x)$.

Theorem 6.1. *The mapping class group is generated by Dehn twists.*

If an element of the mapping class group does not act faithfully on multicurves, then it fixes all such curves (up to isotopy), hence commutes with all Dehn twists, hence lies in the center. To show that the action of the mapping class group on multicurves is faithful, it suffices to show that the center is trivial. This will be true whenever $g > 2$ and S_g does not have boundary.

To see this we will draw a suitable collection of circles on S_g. Any element of the center preserves (up to isotopy) every circle, so it preserves the graph describing how circles intersect.

But when $g > 2$ we can arrange these circles so that the corresponding graph has no automorphisms. It follows that up to isotopy an element of the center of the mapping class group fixes the graph pointwise.

The complement of the graph is a collection of disks, and a homeomorphism of the disk fixing the boundary is isotopic to the identity through homeomorphisms fixing the boundary, so we conclude.

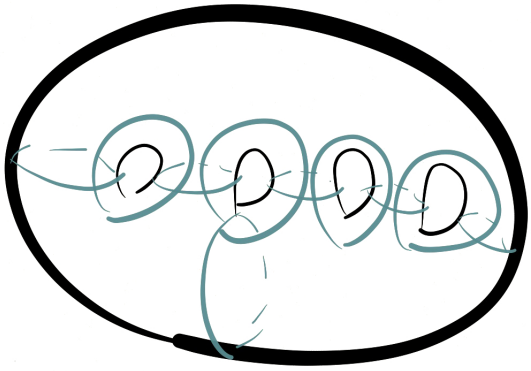


Figure 5: Circles on a 4-holed torus.

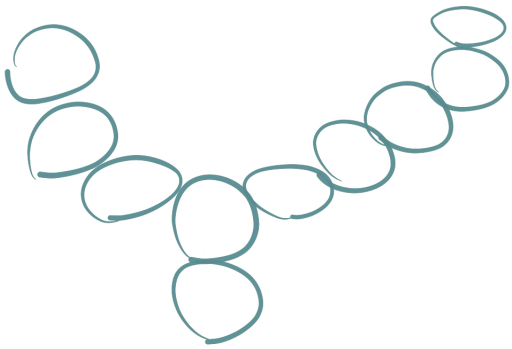


Figure 6: A graph describing the intersections of the circles which has no automorphisms.

This argument does not work when $g = 2$, when we can take a 180° rotation. This is a reflection of the fact that when $g = 2$ a smooth projective algebraic curve over $\mathbb{C}$ is hyperelliptic, so alwways has a hyperelliptic involution, but when $g > 2$ not all curves are hyperelliptic.

Dehn-Thurston coordinates were discovered by Dehn and rediscovered by Thurston.

7 More about ideal polygons

Last time we discussed two methods for understanding ideal polygons. One was to send three of their vertices to $0, 1, \infty$, and another was to choose horocycles around the vertices and count distances. We want a more natural version of this picture (which does not depend on a choice of triangulation).

Consider the hyperboloid model $x^2 + y^2 = z^2 - 1$. What do horocycles look like here? First, what do circles look like? Using the projection to a plane, they come from cones coming from the other hyperboloid.

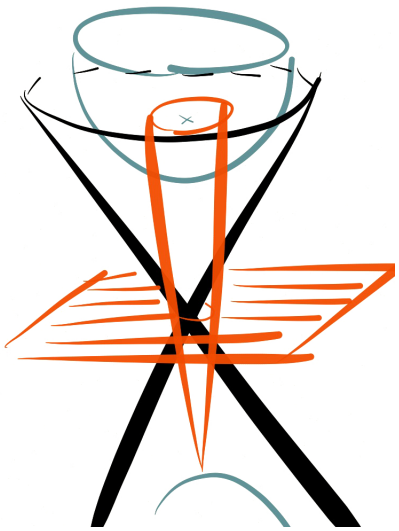


Figure 7: A cone and the corresponding circle.

Alternatively, we can intersect the hyperboloid with a plane (analogous to what happens with a sphere). On a sphere, the center of the corresponding circle is the unique point whose tangent plane is parallel to the intersecting plane, and the same is true on the hyperboloid.

This is clearest to see for the lowest point on the hyperboloid, and everything is invariant under $\mathrm{SO}^+(2,1)$, so it follows everywhere.

To get horocycles, we take tangent planes to the cone (circles centered at infinity), then translate them so that they intersect with the hyperboloid.

We want to describe this situation algebraically in terms of the inner product on $\mathbb{R}^3$ with corresponding quadratic form $x^2 + y^2 - z^2$. The hyperboloid is the set of vectors v such that $\langle v, v\rangle = -1$. The (null) cone is the set of vectors v such that $\langle v, v\rangle = 0$ (*null vectors*). Planes can be described in the form

Figure 8: A plane intersecting a sphere and the corresponding circle.

Figure 9: A plane intersecting a hyperboloid and the corresponding circle.

$$P_v = \{w : \langle v, w\rangle = k\}. \tag{2}$$

When v is a null vector, $v \in P_{v,0}$. To get horocycles, we take hyperplanes of the form $P_{v,-1}$ and intersect them with the hyperboloid. In other words, horocycles have the form

$$h_v = \{w : \langle v, w\rangle = \langle w, w\rangle = -1\}. \tag{3}$$

We therefore have a natural correspondence between horocycles and nonzero null vectors in the upper cone ($z \geq 0$). Now, given two vectors v_1, v_2 at which we have

centered two horocycles h_1, h_2, we want to describe algebraically the corresponding length. We begin by defining the λ-*length*

$$\lambda(h_1, h_2) = \sqrt{-\frac{1}{2}\langle v_1, v_2\rangle}. \tag{4}$$

This is some function $f(\ell(h_1, h_2))$ of the length. To explain the factor of $\frac{1}{2}$, first take $v_1 = (1, 0, 1)$. The corresponding horocycle contains $x = (0, 0, 1)$, and so does the horocycle corresponding to $v_2 = (-1, 0, 1)$. In fact, the horocycles are tangent at x, so $\ell(h_1, h_2) = 0$ in this case. On the other hand, $\langle v_1, v_2\rangle = -2$, so the above normalization gives $\lambda(h_1, h_2) = 1$.

Now we should talk more about the relationship between the upper half-plane model and the hyperboloid model. In the former the isometry group is $\mathrm{PSL}_2(\mathbb{R})$ while in the latter the isometry group is $\mathrm{SO}^+(2, 1)$. In the former the ideal points are $\mathbb{RP}^1$ while in the latter the ideal points are rays in the null cone. We would like a correspondence between them, hence a way to take vectors in $\mathbb{R}^2$ to null vectors in $\mathbb{R}^{2,1}$.

We can do this by thinking of $\mathbb{R}^{2,1}$ as symmetric 2×2 matrices with the negative of the determinant as the quadratic form. Given a vector in $\mathbb{R}^2$, we can now tensor it with itself to get such a symmetric matrix, giving

$$\begin{bmatrix} a \\ b \end{bmatrix} \mapsto \begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix} \tag{5}$$

Exercise 7.1. *Check that diagonalizing this quadratic form gives a map* $(a, b) \mapsto (a^2 - b^2, 2ab, a^2 + b^2)$ *from* $\mathbb{R}^2$ *to null vectors in* $\mathbb{R}^{2,1}$.

This gives a map from $\mathbb{R}^2$ to horocycles. What is the λ-length in these terms? Given (a, b) and (c, d), the dot product of the corresponding null vectors is

$$abcd - \frac{1}{2}(a^2d^2 + b^2c^2) = -\frac{1}{2}(ad - bc)^2 = -\frac{1}{2}\det\begin{bmatrix} a & c \\ b & d \end{bmatrix}^2. \tag{6}$$

So the corresponding λ-length is

$$\frac{1}{2}\left|\det\begin{bmatrix} a & c \\ b & d \end{bmatrix}\right|. \tag{7}$$

Exercise 7.2. *Was this detour necessary? Is there a natural way to write down horocycles as subsets of* $\mathbb{CP}^1$ *(which contains both the disk and the half-plane models) and do these computations there instead?*

274 Curves on Surfaces, Lecture 7

Dylan Thurston

Fall 2012

8 Horocycles and lengths

Last time we saw that in the hyperboloid model there is a nice way to write down horocycles: they can be given by sets of the form

$$h_v = \{w : \langle v, w\rangle = \langle w, w\rangle = -1\} \tag{1}$$

where v is a null vector. On the other hand, it is often more convenient to work in the upper half-plane. Here the boundary is $\mathbb{RP}^1$ and we saw previously that there is a way to write down from a vector $(a, b) \in \mathbb{R}^2$ a null vector $(2ab, a^2 - b^2, a^2 + b^2)$, and this gives a map from $\mathbb{RP}^1$ to null vectors, which give horocycles. It would be nice to be able to avoid this and directly describe the horocycle associated to (a, b) in the upper half-plane.

Last time we also described the λ-length

$$\lambda(h_1, h_2) = \sqrt{-\frac{1}{2}\langle v_1, v_2\rangle} = \frac{1}{2}\left|\det(w_1, w_2)\right| \tag{2}$$

between two horocycles in terms of the inner product of the corresponding null vectors and then in terms of the determinant of the corresponding vectors in $\mathbb{R}^2$. This is some function $f(\ell(h_1, h_2))$ of the length which we have not yet worked out. To compute this function, we will work in the upper half-plane with a horocycle at 0 and a horocycle at ∞ (which is a horizontal line). We have one more degree of freedom, so we will choose the diameter of the horocycle at 0 to be 1.

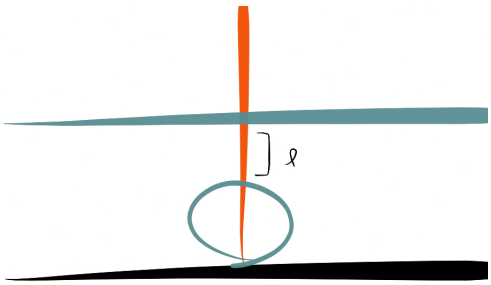


Figure 1: Two horocycles in the upper half-plane.

Then the horocycle at ∞ is the line $y = c$ for some c and the distance between them is

$$\int_1^c \frac{dy}{y} = \log c. \tag{3}$$

The nicest case is when we also have $c = 1$, in which case the distance is 0. We can do this by scaling the horocycle at ∞ by $\frac{1}{c}$ without scaling the horocycle at 0. This is not an isometry, so it will change the length and λ-length. In intrinsic hyperbolic terms, we move each point of the horocycle by $\log c$ away from ∞. To determine what this does to v_i or w_i we need to determine the corresponding element of $\mathrm{PSL}_2(\mathbb{R})$. As a fractional linear transformation, this is $x \mapsto \frac{x}{c}$, which is associated to the matrix

$$\begin{bmatrix} \frac{1}{\sqrt{c}} & 0 \\ 0 & \sqrt{c} \end{bmatrix} \in \mathrm{SL}_2(\mathbb{R}). \tag{4}$$

Now we should figure out what w_1 and w_2 are (the elements of $\mathbb{R}^2$ associated to our horocycles). We will have $w_1 = (0, y_1)$ for some y_1 and $w_2 = (x_2, 0)$ for some x_2. Scaling by $\frac{1}{c}$ multiplies w_2 by $\frac{1}{\sqrt{c}}$ and changes the λ-length to 1 by our normalization, so

$$\lambda(h_1, h_2) = \sqrt{c} = \exp\left(\frac{\ell(h_1, h_2)}{2}\right). \tag{5}$$

Exercise 8.1. *Find $\lambda(h_1, h_2)$ in terms of the Euclidean geometry of two horocycles in the upper half-plane.*

Given a decorated and triangulated ideal polygon, we can ask about how λ-lengths change when we change triangulations. Hyperbolically this is a messy computation, but algebraically it becomes nicer. For an ideal quadrilateral determined by four vectors w_1, w_2, w_3, w_4, we want to compute one of the corresponding determinants in terms of the others.

Writing the λ-lengths of the four sides as A, B, C, D and writing the λ-lengths of the diagonals as E, F, we get the following result.

Lemma 8.2. *(Ptolemy relation)* $\lambda(F)\lambda(E) = \lambda(A)\lambda(C) + \lambda(B)\lambda(D)$.

This relation is named after the corresponding relation for a quadrilateral inscribed in a circle in Euclidean geometry.

Proof. This is really a statement about a 2×4 matrix (assembled from the w_i)

$$\begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \end{bmatrix}, \tag{6}$$

namely a quadratic relation between the 2×2 minors. Both sides of this quadratic relation are invariant under scaling any of the columns, so assuming that the $y_i \neq 0$

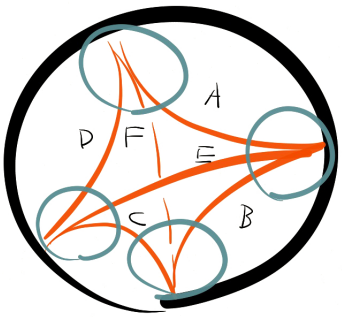


Figure 2: λ-lengths in a decorated ideal quadrilateral.

we may assume WLOG that the y_i are all equal to 1. Then the quadratic relation becomes

$$(x_2 - x_4)(x_1 - x_3) = (x_1 - x_2)(x_3 - x_4) + (x_1 - x_4)(x_2 - x_3). \tag{7}$$

Alternatively, by a suitable change of coordinates we can arrange $x_1 = 1, x_2 = 0, y_1 = 0, y_2 = 1$, which simplifies the relation considerably.

Alternatively, let $a = \sum x_i e_i$ and $b = \sum y_i e_i$ be vectors in $\mathbb{R}^4$ with e_i the standard basis. Then $a \wedge b$ has components which are the 2×2 minors above in the standard basis $e_i \wedge e_j, i < j$ of $\Lambda^2(\mathbb{R}^4)$. We have $(a \wedge b) \wedge (a \wedge b) = 0$ by standard properties of the exterior product, and expanding this out in the standard basis gives the relation above. □

This is also known as the Plücker relation.

Exercise 8.3. *Relate $\lambda(h_1, h_2)$ to Euclidean geometry in the disk model. Prove the hyperbolic Ptolemy relation using the Euclidean one.*

Exercise 8.4. *Relate the cross-ratio of four ideal points to λ-lengths of a corresponding decorated ideal quadrilateral.*

Since a decorated ideal n-gon is determined by its n horocycles, specifying such an n-gon is equivalent to specifying a collection of n points in $\mathbb{R}^2/\{\pm 1\}$ modulo the action of $\mathrm{PSL}_2(\mathbb{R})$, which is very close to specifying a point in the Grassmannian $\mathrm{Gr}_{2,n}$ except that there are some cyclic order and positivity conditions. The positivity condition is equivalent to specifying a $2 \times n$ matrix all of whose minors have positive determinant.

Let's consider more general surfaces than the hyperbolic plane, such as the punctured torus. This surface admits a complete hyperbolic metric of constant curvature -1, and so we can talk about geodesics on it, which go to infinity (the puncture). We can write down three such geodesics giving an ideal triangulation of the torus.

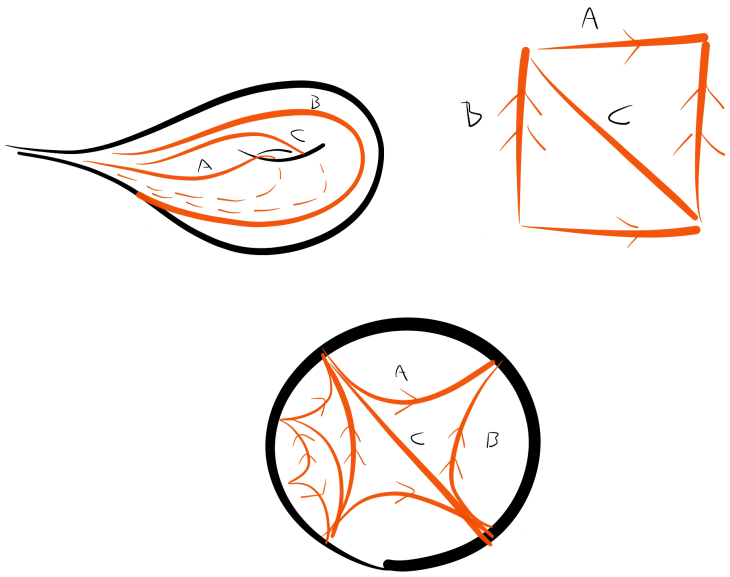


Figure 3: Three views of a punctured torus and three geodesics on it.

We want to measure lengths by decorating using horocycles as before. Here by horocycle we mean a curve which lifts to a horocycle in the universal cover (which is the hyperbolic plane).

Theorem 8.5. *Decorated hyperbolic structures on the punctured torus are parameterized by the three λ-lengths in a triangulation.*

Alternatively, let $T^{g,n}$ be the space of complete hyperbolic structures on a (compact, orientable) surface of genus g with n punctures up to isotopy (*Teichmüller space*), and let $\tilde{T}^{g,n}$ be the space of correspondingly decorated hyperbolic structures, where we also choose horocycles (*decorated Teichmüller space*). Then $\tilde{T}^{1,1}$ is homeomorphic to $\mathbb{R}^3$ with the homeomorphism given by λ-lengths.

There is an interesting relationship between these ideas and number theory. A *Markov triple* is a solution to $x^2 + y^2 + z^2 = 3xyz$. There is an obvious solution $(1, 1, 1)$, and new solutions can be generated from old solutions by permutation or by applying

$$(x, y, z) \mapsto (x, y, \frac{x^2 + y^2}{z}). \tag{8}$$

Consider the Ptolemy relation for a decorated ideal triangulation of a punctured torus. This gives $\lambda(C)\lambda(C') = \lambda(A)^2 + \lambda(B)^2$ where C' is the fourth diagonal, or

$$\lambda(C') = \frac{\lambda(A)^2 + \lambda(B)^2}{\lambda(C)}. \tag{9}$$

What is the relationship? In the special case that $\lambda(A) = \lambda(B) = \lambda(C) = 1$ (equilateral; all of the horocycles touch), the Ptolemy relation allows us to compute other λ-lengths, such as the lengths of various diagonals, and these are precisely the Markov triples we get starting from $(1, 1, 1)$.

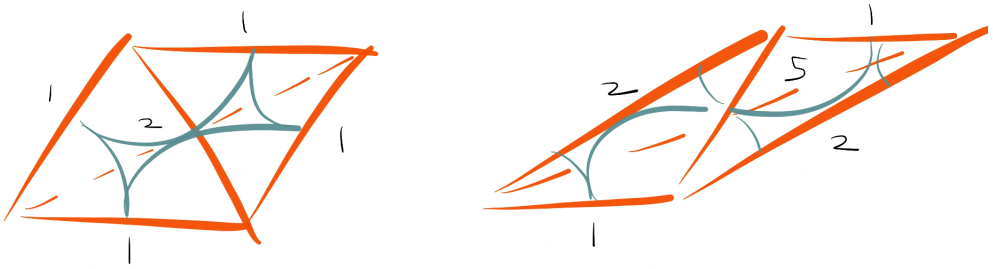


Figure 4: Changing triangulations on an equilateral torus.

A *Markov number* is a number occurring in a Markov triple.

Conjecture 8.6. *For every Markov number n, the Markov triples containing n can be connected by the transformation above without removing n.*

Equivalently, the *simple length spectrum* of the equilateral punctured torus is simple up to symmetries. The length spectrum is the multiset of lengths of closed geodesics. The simple length spectrum is the multiset of lengths of closed simple (non-intersecting) geodesics. A multiset is simple if every element occurs with multiplicity 1.

274 Curves on Surfaces, Lecture 8

Dylan Thurston
Notes by Qiaochu Yuan

Fall 2012

9 Teichmüller space and Markov triples

Last time there were some exercises. One was to compute λ-lengths between two horocycles in terms of Euclidean geometry in the upper half-plane. If x is the distance between the corresponding points on the boundary and Δ_1, Δ_2 are the diameters, then this is $\frac{x}{\sqrt{\Delta_1 \Delta_2}}$.

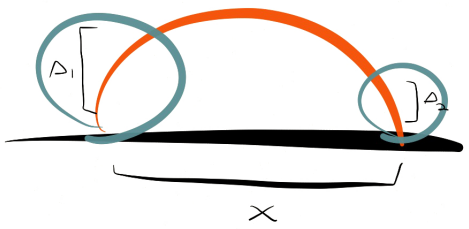


Figure 1: Euclidean distances determining the λ-length.

To check this, it suffices to check that it is invariant under hyperbolic isometries, behaves correctly with respect to scaling one of the circles, and is equal to 1 when the circles are tangent. The only nontrivial step is invariance under inversion with respect to some circle, but the circle determined by the geodesic connecting the boundary points fixes both horocycles setwise, so the conclusion follows. Alternatively, we can invert so that one of the horocycles is sent to ∞.

One was to relate the Euclidean and hyperbolic Ptolemy relations. Hint: if (x, y) is a point on the unit circle, then $x^2 + y^2 = 1$, and there is a natural way to construct a null vector from this data.

One was to relate the cross ratio of four ideal points a, b, c, d to λ-lengths. To access the cross ratio, we will work in the upper half-plane and send three of the points to $0, 1, \infty$. Let the fourth point be τ.

We want to compute τ in terms of λ-lengths, so we need to choose four horocycles (one of which is at ∞). Letting the diameters of the three finite horocycles be $\Delta_a, \Delta_b, \Delta_d$ and letting the infinite horocycle be at height y_c, we have

$$\lambda(a, b) = \frac{1}{\sqrt{\Delta_a \Delta_b}} = \lambda(B) \tag{1}$$

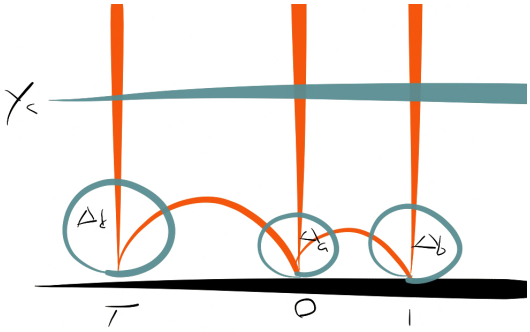


Figure 2: A decorated ideal quadrilateral in the upper half-plane.

$$\lambda(a,d) = \frac{-\tau}{\sqrt{\Delta_a \Delta_d}} = \lambda(A) \tag{2}$$

$$\lambda(c,d) = \sqrt{\frac{y_c}{\Delta_d}} = \lambda(D) \tag{3}$$

$$\lambda(b,c) = \sqrt{\frac{y_c}{\Delta_b}} = \lambda(C) \tag{4}$$

and we want some combination of these which does not depend on the choice of horocycle. This turns out to be

$$\tau = -\frac{\lambda(A)\lambda(C)}{\lambda(B)\lambda(D)}. \tag{5}$$

We should give a more careful definition of Teichmüller space. Suppose we want to classify hyperbolic metrics (complete, finite volume) on a (smooth) surface S. The first method would be to look at all hyperbolic metrics, where we declare two metrics to be equivalent if there is an isometry between them. If S is closed of genus g, this gives the moduli space $\mathcal{M}_g$. This is not a manifold but an orbifold due to metrics with additional symmetries.

A space which is easier to parameterize is to look at all hyperbolic metrics, where we declare two metrics to be equivalent if there is an isometry between them which is isotopic to the identity (in the diffeomorphism group; for surfaces, this is equivalent to being homotopic to the identity). If S is closed of genus g, this is Teichmüller space. Surprisingly, it is a finite-dimensional manifold of dimension $6g-6$ for $g > 1$ even though it is defined as the quotient of some infinite-dimensional space by some other infinite-dimensional space.

When $g = 1$ we can instead look at Euclidean metrics on the torus modulo isometries isotopic to the identity and by global scaling. By choosing an oriented basis for homology, this can be identified with oriented pairs of vectors in $\mathbb{R}^2$ up to scaling and rotation, which can be identified with the upper half-plane. If we instead quotient by all isometries, we get the upper half-plane modulo $\mathrm{PSL}_2(\mathbb{Z})$. This is the moduli space $\mathcal{M}_1$.

We can also allow n punctures with *cusps* at the punctures (we want the metric to remain complete and of finite volume). This gives moduli spaces $\mathcal{M}_{g,n}$.

To relate moduli space and Teichmüller space, we should quotient Teichmüller space by all diffeomorphisms and not just those isotopic to the identity. The additional symmetries by which we have to quotient are precisely the mapping class group $\mathrm{MCG}(S)$. For example, $\mathcal{M}_{1,1} = \mathcal{T}_{1,1}/\mathrm{MCG}(T^2, x)$, which exhibits $\mathcal{M}_{1,1}$ as the quotient of the upper half-plane by $\mathrm{SL}_2(\mathbb{Z})$.

Last time we also chose horocycles, giving us a decorated Teichmüller space $\tilde{\mathcal{T}}_{g,n}$. This has dimension $6g-6+3n$, which is the number of arcs in a triangulation. Given a triangulation, we can measure its λ-lengths to obtain a map

$$\tilde{\mathcal{T}}_{g,n} \to \mathbb{R}^{6g-6+3n} \tag{6}$$

and this is a homeomorphism; moreover, these maps coming from triangulations are related by exchange relations coming from the Ptolemy relation.

Last time we also discussed Markov triples, the integer solutions to $x^2+y^2+z^2 = 3xyz$. These arise when considering lengths of the segment of a horocycle contained in an ideal triangle that it decorates.

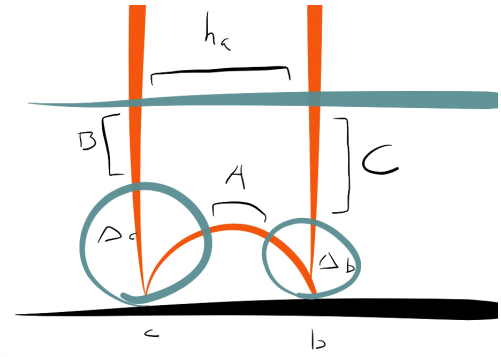


Figure 3: A segment of a horocycle at infinity in the upper half-plane.

This is necessarily some function of the λ-lengths involved, so we want to determine

what this function is. Sending the corresponding horocycle to ∞ in the upper half-plane and letting its height be y_a, we compute that this length is

$$\int_0^1 \frac{dx}{y_a} = \frac{1}{y_a} = h_a. \tag{7}$$

But we want an expression in terms of λ-lengths. The relevant λ-lengths are $\lambda(A) = \frac{1}{\sqrt{\Delta_b \Delta_c}}, \lambda(B) = \sqrt{\frac{y_a}{\Delta_c}}, \lambda(C) = \sqrt{\frac{y_a}{\Delta_b}}$. This gives

$$h_a = \frac{\lambda(A)}{\lambda(B)\lambda(C)}. \tag{8}$$

Consider now a horocycle on a punctured torus with an ideal triangulation. Let the λ-lengths be x, y, z.

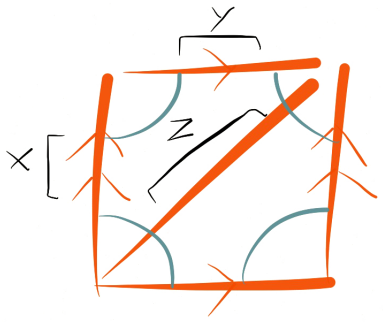


Figure 4: A horocycle on a punctured torus.

The total length of the horocycle can be computed using the above to be

$$2\left(\frac{x}{yz} + \frac{y}{zx} + \frac{z}{xy}\right) = 2\frac{x^2 + y^2 + z^2}{xyz}. \tag{9}$$

The Markov triple equation then says precisely that the total length of the corresponding horocycle is equal to 6, which occurs, for example, when $x = y = z = 1$. This total length is independent of triangulation, and changing triangulations corresponds to getting new Markov triples.

Exercise 9.1. *The above discussion was for the equilateral torus. What are λ-lengths in the square torus? What is the corresponding Diophantine equation?*

Next time we will talk about cluster algebras. These are, among other things, algebras A with sets of distinguished variables called clusters. For any cluster $C = (x_i)_{i=1}^n$, the entire algebra A is contained in the ring of Laurent polynomials $\mathbb{Q}[x_i^{\pm}]$ in the x_i. We can also move between clusters as follows: for any C and i, there is a unique cluster $C' = C \setminus \{x_i\} \cup \{x_i'\}$ satisfying an exchange relation of the form

$$x_i x_i' = P(x_1, ...\hat{x_i}, ...x_n) \tag{10}$$

where P is a polynomial with two terms. There are also rules determining which polynomials P can occur and how they are related.

The model examples to keep in mind come from surfaces. If S is a surface, possibly with boundary, and M is a set of marked points, then a choice of hyperbolic metric and horocycles allows us to associate to any triangulation a collection of λ-lengths. Changing the triangulation by replacing a diagonal in a quadrilateral changes the λ-lengths by a Ptolemy relation, which is an exchange relation. The λ-lengths associated to a given triangulation form the clusters of a cluster algebra of functions on decorated Teichmüller space.

Example Consider a pentagon. We can choose horocycles so that the λ-lengths are all equal to 1. Choosing two more diagonals gives a triangulation with two new λ-lengths x_1, x_2.

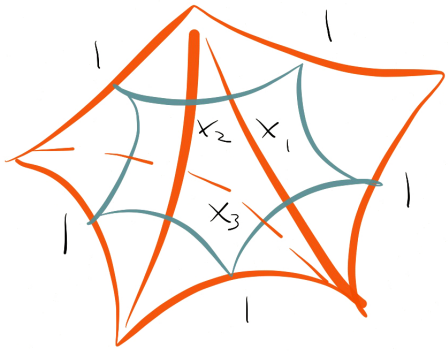


Figure 5: A pentagon and its diagonals.

Now alternately applying quadrilateral flips to get new diagonals gives a sequence of λ-lengths x_n satisfying the recurrence

$$x_{n+1}x_{n-1} = x_n + 1. \tag{11}$$

This recurrence is periodic with period 5 since there are only 5 diagonals in the pentagon. It consists of Laurent polynomials in x_1, x_2, and this is not obvious. Computing the terms gives

$$x_3 = \frac{x_2}{x_1} + \frac{1}{x_1} \tag{12}$$

$$x_4 = \frac{1}{x_1} + \frac{1}{x_1 x_2} + \frac{1}{x_2} \tag{13}$$

$$x_5 = \frac{1 + x_1}{x_2} \tag{14}$$

$$x_6 = x_1. \tag{15}$$

Exercise 9.2. *Consider the recurrence $x_{n+1}x_{n-1} = x_n^2 + 1$. Check a few terms to verify that we get Laurent polynomials. Can you find a surface giving rise to this recurrence as above?*

274 Curves on Surfaces, Lecture 9

Dylan Thurston
Notes by Qiaochu Yuan

Fall 2012

10 Cluster algebras

Yesterday Zelevinsky gave a talk about cluster algebras.

Question from the audience: in a surface cluster algebra, we can show that the denominators of all the relevant rational functions don't have positive real roots (because the λ-lengths described by these rational functions exist). But the theory of cluster algebras shows that these denominators are in fact monomials (the Laurent phenomenon). Can this be shown using some notion of complex λ-lengths?

Answer: hyperbolic structures correspond to representations of the fundamental group into $\mathrm{PSL}_2(\mathbb{R})$. To complexify this, we could instead look at representations of the fundamental group into $\mathrm{PSL}_2(\mathbb{C})$. This is the group of orientation-preserving isometries of hyperbolic space $\mathbb{H}^3$.

First, some background: the isometries of $\mathbb{H}^2$ can be organized into three classes. The *elliptic* elements fix some point, which in the disk model we can arrange to be the origin. The *hyperbolic* elements fix some geodesic, which in the disk model we can arrange to be a diameter, and also fix two points on the boundary. The *parabolic* elements fix one point on the boundary.

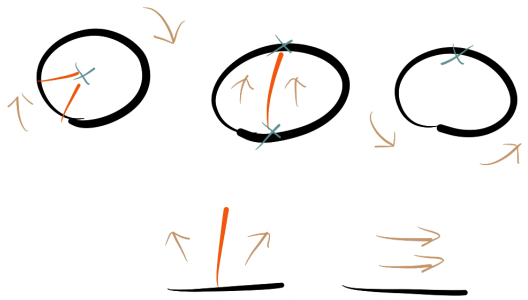


Figure 1: The three types of isometries.

This classification is obtained as follows. An element of $\mathrm{PSL}_2(\mathbb{R})$ is represented by a 2×2 real matrix with two eigenvalues λ, λ^{-1}. It could have two distinct real eigenvalues, in which case it can be conjugated to a diagonal matrix (hyperbolic). It could have two complex eigenvalues, in which case λ is on the unit circle (elliptic). Or it could have a repeated eigenvalue, so $\lambda^2 = 1$ and $\lambda = \pm 1$. We can choose a representative with $\lambda = 1$, in which case we obtain a Jordan block (parabolic).

Consider a closed surface with punctures (but no boundary components), with cusps at the boundaries. Having a cusp means that the monodromy around the

puncture is parabolic. (Hyperbolic monodromy corresponds to a geodesic boundary component.)

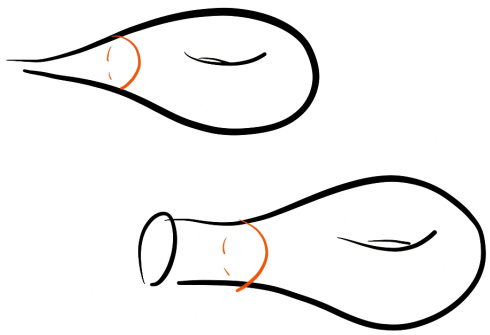


Figure 2: A cusp vs. a geodesic boundary.

Choosing a horocycle at the cusp corresponds to choosing an eigenvector of the monodromy in $\mathbb{R}^2$ up to sign, and this allows us to define λ-lengths using the determinant formula in a way which continues to work when $\mathrm{PSL}_2(\mathbb{R})$ is replaced with $\mathrm{PSL}_2(\mathbb{C})$. Unfortunately, the corresponding λ-lengths are not always defined, so this doesn't appear to prove the Laurent phenomenon.

We know from Zelevinsky's talk that (some) cluster algebras can be described using quivers (directed graphs). The quiver associated to a triangulation of a surface has vertices the edges of the triangulation and directed edges are given by clockwise adjacency. We need to check that mutation corresponds to changing the triangulation.

Mutation at a vertex occurs in three steps:

1. Add composite arrows through the vertex.
2. Reverse arrows through the vertex.
3. Delete all oriented 2-cycles that occur.

Doing this does indeed correspond to changing the triangulation (by a quadrilateral flip).

Moreover, the λ-lengths in the triangulation (relative to some choice of horocycles and hyperbolic structure) are cluster variables with exchange relation given by the Ptolemy relation

$$\lambda(E)\lambda(F) = \lambda(A)\lambda(C) + \lambda(B)\lambda(D). \tag{1}$$

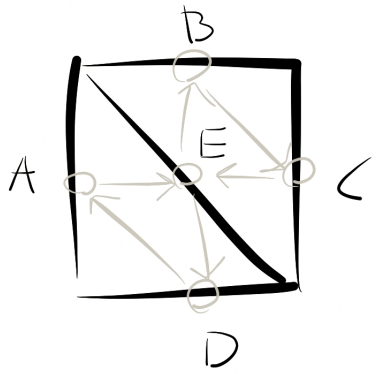


Figure 3: The quiver associated to a triangulation.

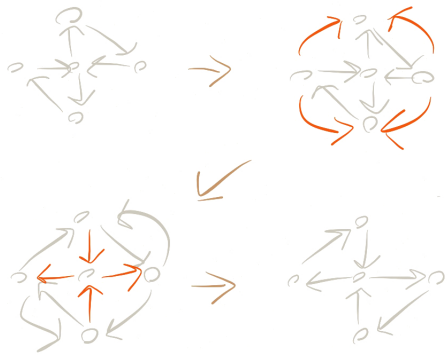


Figure 4: Mutation and changing the triangulation.

Here $\lambda(F)$ is the new cluster variable given by the length of the new diagonal.

In general, some vertices of the quiver are frozen and cannot be mutated (for a surface with boundary these are the boundary edges). We can indicate the difference by drawing frozen vertices in black and non-frozen vertices in white.

The cluster algebra associated to a pentagon is finite type of type A_2. We can get A_3 by triangulating the hexagon appropriately. If we reverse one of the arrows in the corresponding quiver, we get a different triangulation of the hexagon. More generally we can get A_n by triangulating an $(n+3)$-gon.

Similarly, we get D_n from a punctured n-gon. (To do this we needed to cancel an edge, which makes the exchange relation a little different from what one expects.)

One reason to care about cluster algebras is that they occur in many different

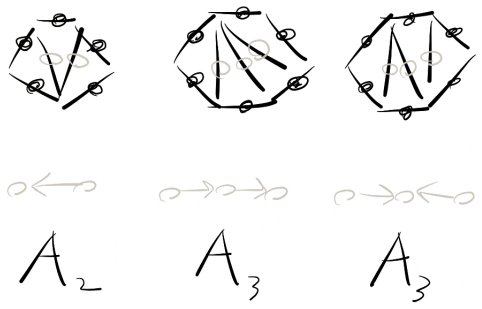


Figure 5: Cluster algebras of type A_n from polygons.

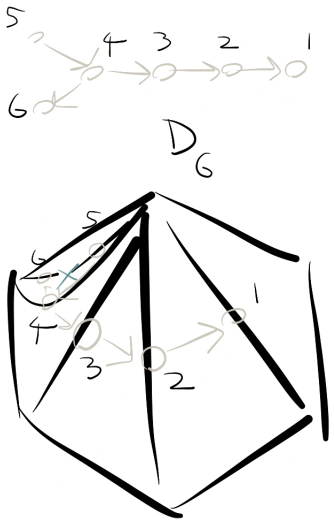


Figure 6: The cluster algebra of type D_6 from a punctured hexagon.

places for different reasons. For example, D_4, which we obtain from a punctured 4-gon, also occurs when considering a cluster algebra structure on 3×3 matrices obtained from considering minors.

Theorem 10.1. *(Felikson, Shapiro, Tumarkin) The mutation-finite cluster algebras are the following:*

1. *Surface (or orbifold) cluster algebras (including A_n, D_n),*

2. *Rank* 2 *cluster algebras,*

3. $11 + 7$ *additional diagrams, some of which have different possible assignments of short and long roots.*

Question from the audience: is it obvious that surface cluster algebras are mutation-finite? They may have infinitely many triangulations.

Answer: what we care about is not the number of triangulations but the number of triangulations modulo the action of the mapping class group (since this is already enough to determine the quiver). Up to the action of the mapping class group, a triangulation is specified by the combinatorial data of how to glue a fixed set of triangles together, and there are finitely many ways to do this. Alternately, there are finitely many possible B-matrices because their entries are at most 2 in absolute value.

Exercise 10.2. *Find surfaces that give the affine Dynkin diagrams.*

(Orientation of the arrows matters for $\tilde{A}_n$ since there is a cycle. But $\tilde{D}_n$ is acyclic so orientation doesn't matter there.)

Exercise 10.3. *What's a geometric surface model for D_4 in the form of a quiver with all the arrows pointing outwards? Can you see the triality symmetry?*

Previously we asked for a surface cluster algebra giving the recurrence $x_{n+1}x_{n-1} = x_n^2 + 1$. This can be obtained from a (punctured) torus one of whose side lengths is equal to 1, or alternately (by cutting the torus open) from an annulus.

Consider again the triangulation of a punctured n-gon near the puncture. We had to cancel an edge to get D_n, and this causes the corresponding mutation to disagree with the result we get from the Ptolemy relation.

Namely, changing triangulations at D, the Ptolemy relation gives

$$\lambda(E)\lambda(D) = \lambda(A)\lambda(C) + \lambda(B)\lambda(C) \tag{2}$$

but on the cluster algebra side the exchange relation is

$$x(D)x(D') = x(B) + x(A) \Rightarrow x(D') = \frac{\lambda(E)}{\lambda(C)} \tag{3}$$

where x denotes a cluster variable; this differs by a factor of $\frac{1}{\lambda(C)}$ from the expected answer $\lambda(E)$.

We can geometrically interpret the above relation as follows. For p a cusp in a hyperbolic surface and h a horocycle around p with a the length of the horocycle, the *conjugate horocycle* $\bar{h}$ is the corresponding horocycle of length $\frac{1}{a}$.

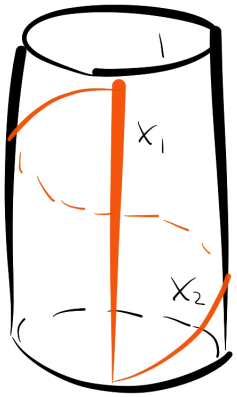


Figure 7: A cylinder / annulus giving the recurrence $x_{n+1}x_{n-1} = x_n^2 + 1$.

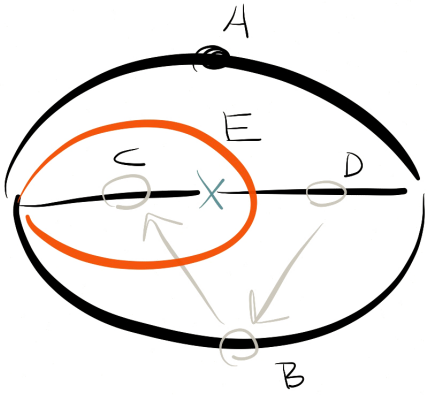


Figure 8: A disagreement between the Ptolemy relation and the exchange relation.

Lemma 10.4. *In a punctured monogon (a punctured ideal triangle with edges identified),*

$$\lambda(A)\lambda(A') = \lambda(B) \Rightarrow \lambda(A') = \frac{\lambda(B)}{\lambda(A)}. \tag{4}$$

(A' is the same geodesic as A, but the λ-length is measured with respect to the conjugate horocycle.)

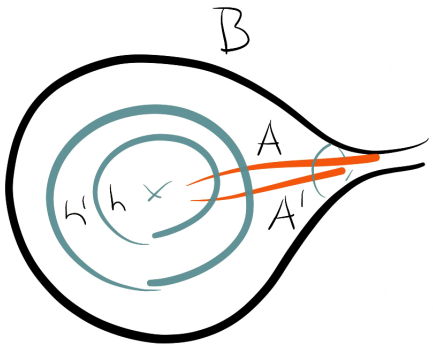


Figure 9: Conjugate horocycles.

Proof. Let a be the length of the original horocycle h. We have a formula describing this length which shows that

$$a = \frac{\lambda(B)}{\lambda(A)^2}. \tag{5}$$

The same formula applied to the conjugate horocycle $\bar{h}$ shows that

$$\frac{1}{a} = \frac{\lambda(B)}{\lambda(A')^2}. \tag{6}$$

Multiplying these gives the conclusion. □

We can now give a geometric interpretation to the cluster variable computation, which is that we are computing the λ-length of the new diagonal E with respect to a conjugate horocycle.

Exercise 10.5. *Another way of writing the relation for conjugate horocycles is as follows. Let h, h' be conjugate horocycles and let A, A' be parallel arcs, with the length of A' being measured with respect to h'. Then*

$$\lambda(A') = \lambda(A)\ell(h) \tag{7}$$

where $\ell(h)$ is the hyperbolic length of h. Give a direct geometric proof of this relationship.

274 Curves on Surfaces, Lecture 10

Dylan Thurston

Fall 2012

11 More about cluster algebras

Last time we discussed conjugate horocycles. This gave a relation $\lambda(A)\lambda(A') = \lambda(B)$ where $\lambda(A')$ is a λ-length measured with respec to the conjugate horocycle. On the other hand, we know that $\ell(h) = \frac{\lambda(B)}{\lambda(A)^2}$, which gives

$$\lambda(A') = \lambda(A)\ell(h) \tag{1}$$

or equivalently taking logarithms,

$$\ell(A') - \ell(A) = 2\ln \ell(h). \tag{2}$$

This can be proven using a scaling argument. The result is clear when $\ell(h) = 1$, since then the horocycle is its own conjugate. In general, a suitable scaling multiplies $\ell(h)$ by c, multiplies $\lambda(A)$ by $\frac{1}{\sqrt{c}}$, and multiplies $\lambda(A')$ by $\sqrt{c}$, so the conclusion follows.

Last time we also asked for a surface giving rise to the affine Dynkin diagrams as quivers. To get $\tilde{A}_{k,\ell}$ we can triangulate an annulus.

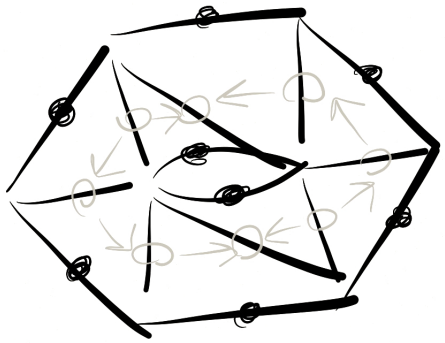


Figure 1: A triangulation giving $\tilde{A}2, 4$.

We also asked for a surface giving rise to D_4 in the orientation where all of the arrows point outward. On the quiver level this can be obtained from the other D_4 we had by mutating twice.

The corresponding geometric exchange relation for the first mutation is

$$x_1 y = x_4 x_3 + x_3 \tag{3}$$

but the actual exchange relation is

$$x_1 x_1' = x_4 + 1. \tag{4}$$

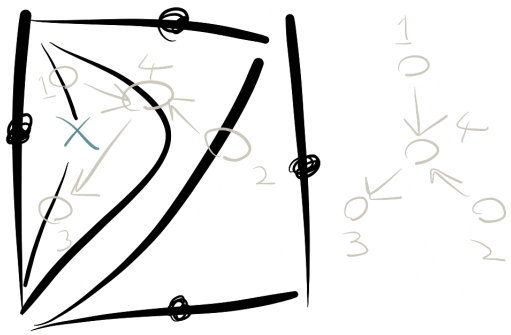


Figure 2: A D_4 with two arrows pointing inward.

As before, this suggests measuring a λ-length with respect to some conjugate horocycle.

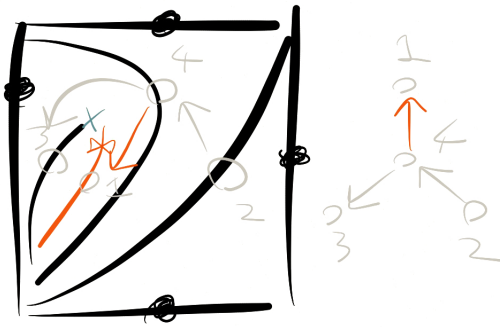


Figure 3: A corrected version of the first mutation giving the correct exchange relation.

Question from the audience: where is the triality symmetry here?

Answer: it appears to be somewhat hidden and is not readily accessible geometrically. Note that quotienting D_4 by triality gives G_2, which is exceptional and does not come from a surface at all.

Another example with hidden symmetry is the 4-punctured sphere. With a tetrahedral triangulation, the corresponding quiver is the octahedron with a certain triangulation. This octahedral quiver can be obtained from a triangulation in a second

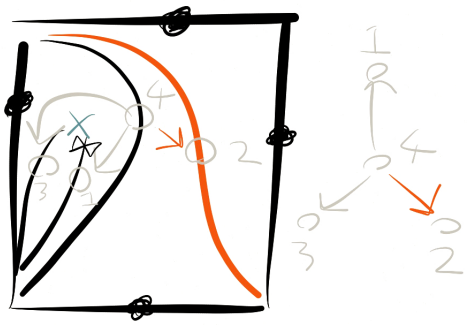


Figure 4: The second mutation.

way, which gives a hidden symmetry (related to Regge symmetry?). More precisely, it can be glued from Type II blocks (see below) in two different ways.

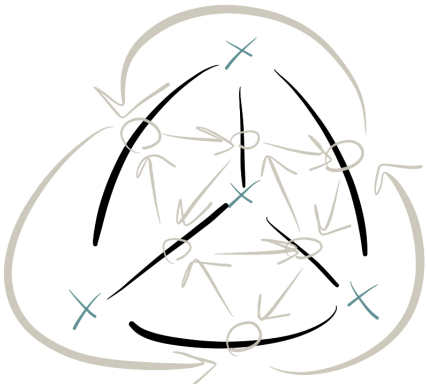


Figure 5: The octahedral quiver.

We will now clarify the geometric meaning of what we have been doing.

A *tagged simple arc* is an arc with one or both ends marked with a notch which does not self-intersect and which does not bound a monogon or a 1-punctured monogon. Notches can only appear at punctures in the interior and should agree at common endpoints if an arc goes from a puncture to itself. Geometrically, a notch indicates that λ-lengths should be measured with respect to the conjugate horocycle. Two tagged arcs are *compatible* if they don't cross and if either

1. the tags agree at common endpoints or

2. the arcs are parallel, one is notched, and one is plain.

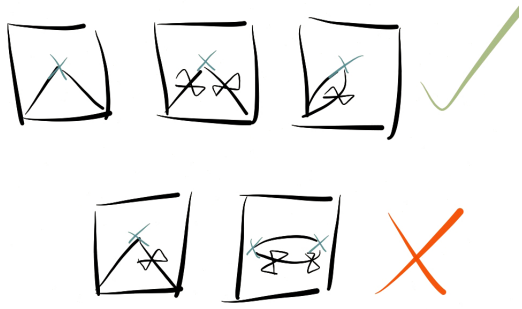


Figure 6: Compatible and incompatible tagged arcs.

A *tagged triangulation* on a surface with a fixed set of marked points is a maximal collection of (distinct) compatible tagged arcs between marked points.

Theorem 11.1. *Any tagged triangulation may be obtained from an ordinary triangulation T by*

1. *replacing self-folded triangles with parallel arcs and*
2. *flipping all tags at some vertices.*

We can construct quivers from a tagged triangulation. The way to remember how this construction works is to remember the relation $\lambda(A)\lambda(A') = \lambda(B)$ for A' a tagged arc parallel to A and B an arc around them. This suggests that when we replace a self-folded triangle with parallel arcs, we effectively double the corresponding vertex in the quiver.

Conversely, to determine when a quiver can come from a tagged triangulation, we can glue *blocks* together (not to themselves) along vertices in such a way that we cancel edges of opposite orientations. Blocks can only be glued along vertices which have not been previously glued.

Any cluster algebra occurring in this way is mutation-finite. However, we don't get some interesting examples, such as the exceptional series.

Exercise 11.2. *Show that it is not possible to obtain E_6, E_7, E_8 by gluing blocks.*

Here is a more precise statement of the classification theorem we stated previously.

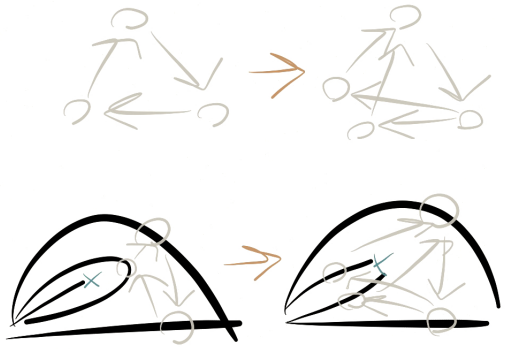


Figure 7: Removing a self-folded triangle and doubling the corresponding vertex.

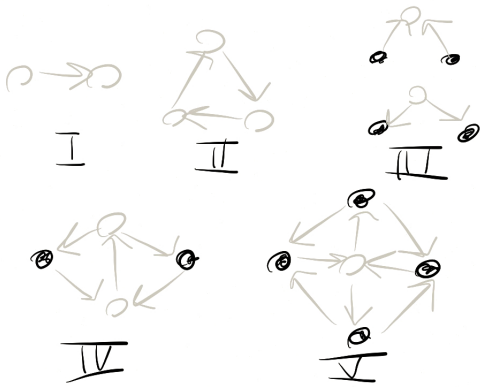


Figure 8: Blocks which glue together to form quivers coming from tagged triangulations.

Theorem 11.3. *Every mutation-finite skew-symmetric cluster algebra is either*

1. *rank* 2,
2. *a surface cluster algebra, or*
3. $E_6, E_7, E_8, \tilde{E}_6, \tilde{E}_7, \tilde{E}_8, E_6^{(1,1)}, E_7^{(1,1)}, E_8^{(1,1)}, X_6, X_7$.

It would be interesting to find a better proof of this.

Exercise 11.4. *Where is the default quiver in Bernhard Keller's applet on the above list? Can you mutate it to get to a standard form?*

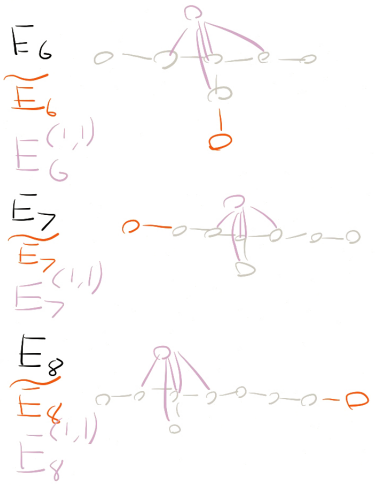


Figure 9: The exceptional diagrams $E_n, \tilde{E}_n$, and $E_n^{(1,1)}$.

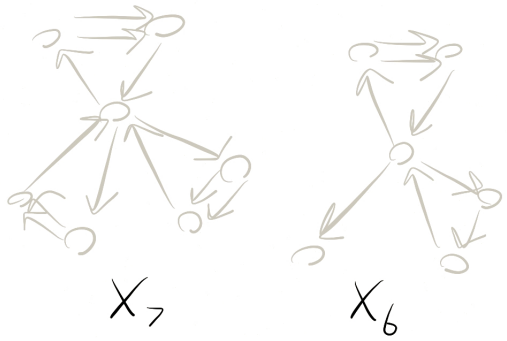


Figure 10: The exceptional diagrams X_7, X_6.

Some of the entries in the above list, such as E_6, E_7, E_8, are not only mutation-finite but of finite type (finitely many cluster variables). The affine ones $\tilde{E}_6, \tilde{E}_7, \tilde{E}_8$ are not mutation-finite, but the number of clusters reachable after n mutations is $O(n)$ rather than exponential for most quivers.

Exercise 11.5. *Mutate the punctured hexagonal quiver to obtain the D_6 quiver.*

Figure 11: Bernhard Keller's default quiver.

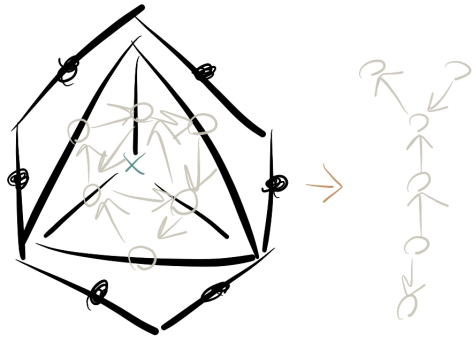


Figure 12: The punctured hexagon and D_6.

274 Curves on Surfaces, Lecture 11

Dylan Thurston

Fall 2012

12 Cross-ratio coordinates

Last time we asked for a quiver mutation-equivalent to the default one in Bernhard Keller's applet. As stated on Keller's page, this is $E_8^{(1,1)}$. One way to see this would be to count vertices and then attempt to rule out that the quiver comes from a surface using the classification in terms of blocks.

Last time we also asked for a proof that E_8 cannot be obtained by gluing blocks. First, type IV and V blocks cannot occur because E_8 is acyclic and the cycles in type IV and V cannot be removed by gluing. Type III blocks cannot occur because E_8 does not have a tail of two black vertices. This leaves blocks of type I and II from which it is impossible to construct a trivalent vertex of the type that occurs in E_8 (after some additional case analysis).

Recall that the cross-ratio of four points a, b, c, d is defined as follows: we send $a \to 0, b \to 1, c \to \infty$, and examine where d goes. A straightforward argument shows that

$$d \mapsto \frac{(d-a)(b-c)}{(d-c)(b-a)} = [d, b; a, c]. \tag{1}$$

We will use the negative of the cross ratio τ because positivity properties are important.

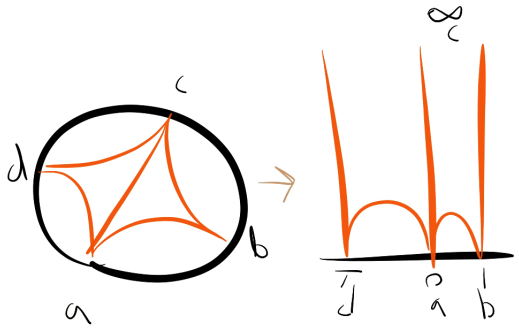


Figure 1: Construction of the cross ratio. (τ should be $-\tau$.)

Some basic properties:

1. $[d, b; a, c]$ is a projective invariant of a, b, c, d; every other such projective invariant is a function of it.

2. $[d, b; a, c] = [b, d; a, c]^{-1} = [d, b; c, a]^{-1}$. As a corollary, $[d, b; a, c] = [b, d; c, a]$, and so there is no need to orient the diagonal edge.

3. The action of S_4 permutes the six possible values of τ generated by $\tau \mapsto \frac{1}{\tau}$ and $\tau \mapsto 1 - \tau$.

$\tau = -[d, b; a, c]$ has other interpretations besides the cross-ratio. It is also a ratio of Euclidean lengths (and phases) $\frac{B \cdot D}{A \cdot C}$ in both the upper half plane and the disk model.

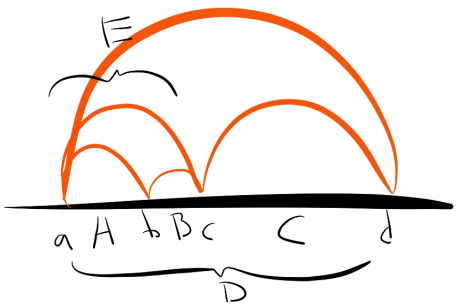


Figure 2: Euclidean distances.

It is also a ratio of λ-lengths

$$\tau(E) = \frac{\lambda(B)\lambda(D)}{\lambda(A)\lambda(C)}. \tag{2}$$

For a more geometric interpretation, we can look at where two angle bisectors intersect on the diagonal. It is reasonable to call these midpoints. They don't intersect at the same point, and the distance between them is the shear $\ln \frac{d}{b}$ if a is sent to the origin.

So in general we have

$$\tau = e^{\text{shear}}. \tag{3}$$

In general, any geometric quantity is some function of the cross-ratio, so for example we also have the following.

Exercise 12.1. *If ℓ is the length between two opposite sides, then $\cosh^2\left(\frac{\ell}{2}\right) = \tau + 1$, whereas for the other pair of opposites sides, $\cosh^2\left(\frac{\ell'}{2}\right) = 1 + \frac{1}{\tau}$.*

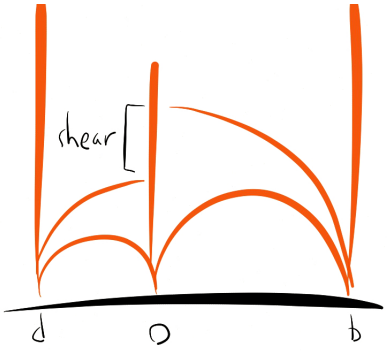


Figure 3: Shear and its relationship to the cross-ratio.

Hint: put the quadrilateral in a more symmetric form.
If θ denotes the angle between the two diagonals, we also have

$$\cos\theta = \frac{x-y}{x+y} = \frac{1-\tau}{1+\tau} \tag{4}$$

or, slightly more nicely,

$$\cos^2\left(\frac{\theta}{2}\right) = \frac{1}{1+\tau}. \tag{5}$$

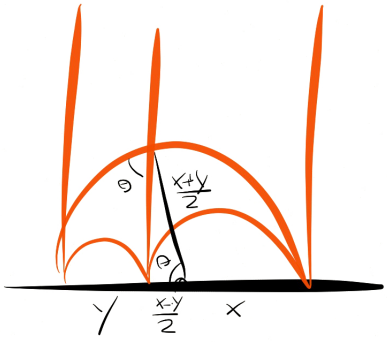


Figure 4: The angle between the two diagonals.

What happens to τ when changing triangulations? Using the interpretation in terms of λ-lengths, we have

$$\tau(E) = \frac{\lambda(B)\lambda(D)}{\lambda(A)\lambda(C)}, \tau(F) = \frac{\lambda(C)\lambda(A)}{\lambda(B)\lambda(D)} \tag{6}$$

so τ gets sent to $\frac{1}{\tau}$.

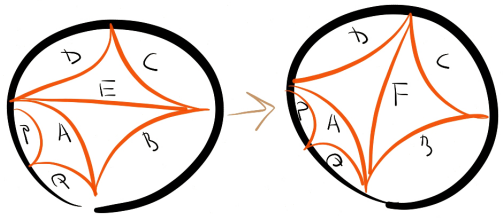


Figure 5: Changing triangulations.

This seems very nice. But when we try to generalize these coordinates to polygons with more sides, shear coordinates depend not only on how an edge changes but on how its neighbors change. So when we add more edges, we get shear coordinates like

$$\tau(A) = \frac{\lambda(Q)\lambda(E)}{\lambda(P)\lambda(F)} \tag{7}$$

which gets transformed to

$$\tau'(A) = \frac{\lambda(Q)\lambda(D)}{\lambda(P)\lambda(F)} = \tau(A)\frac{1}{1+\tau(E)^{-1}}. \tag{8}$$

This factor measures the exponential of the distance between two angle bisectors. Similarly, we find that

$$\tau'(B) = \tau(B)(1+\tau(E)) \tag{9}$$

$$\tau'(C) = \tau(C)\frac{1}{1+\tau(E)^{-1}} \tag{10}$$

$$\tau'(D) = \tau(D)(1+\tau(E)). \tag{11}$$

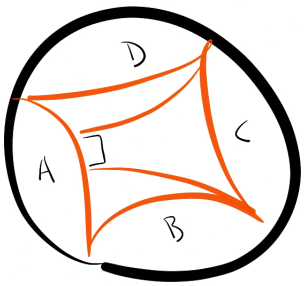


Figure 6: The distance between two angle bisectors.

We have a map from λ-coordinates to τ-coordinates. We know that the former parameterizes decorated Teichmüller space. τ-coordinates do not depend on a choice of horocycle and hence parameterize ordinary Teichmüller space.

But there is someting funny going on. Consider the once-punctured torus.

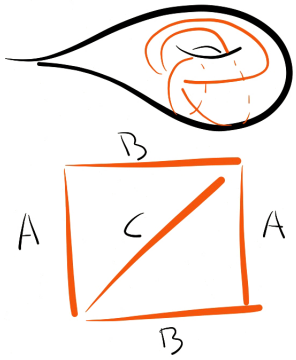


Figure 7: A once-punctured torus.

We parameterized this by three λ-lengths $\lambda(A), \lambda(B), \lambda(C)$ and now we are parameterizing it by three shear coordinates $\tau(A), \tau(B), \tau(C)$ and we are forgetting a horocycle, so this map cannot be surjective. In fact,

$$\tau(A) = \frac{\lambda(B)^2}{\lambda(C)^2}, \tau(B) = \frac{\lambda(C)^2}{\lambda(A)^2}, \tau(C) = \frac{\lambda(A)^2}{\lambda(B)^2} \tag{12}$$

so it follows that $\tau(A)\tau(B)\tau(C) = 1$.

In general, for every puncture, we forget a horocycle, so we should expect a corresponding relation.

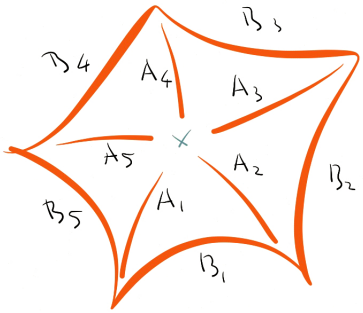


Figure 8: A punctured pentagon.

We get that

$$\prod_i \tau(A_i) = 1 \tag{13}$$

where A_i are the arcs incident to a puncture. What does this mean geometrically, and what do the corresponding τ-coordinates parameterize if we drop this restriction?

Inside each triangle adjacent to a puncture, we can measure the distance between the horocycle and midpoints (we will do this in the upper half-plane with the horocycle at infinity). As we move from triangle to triangle, these distances change by shears, and moving all the way around, the sum of the shears must be 0. The above is the exponential of this relation.

Alternately, on the upper half-plane, shear coordinates are ratios of Euclidean lengths $\tau(A_i) = \frac{x_i}{x_{i+1}}$. The condition that these ratios multiply to 1 is precisely the condition that the two possible lifts of the initial triangle to the upper half-plane have the same width, which in turn expresses the fact that the monodromy around the puncture is a parabolic element (translation).

What if we drop this condition $\prod \tau(A_i) = 1$? Consider a single triangle with two sides identified (alternately, a monogon). There are two midpoints involved, and we do not need to identify the sides so that the midpoints match; in general they can have shear. Correspondingly we get a geometric series of lifts to the upper half-plane which get smaller and smaller.

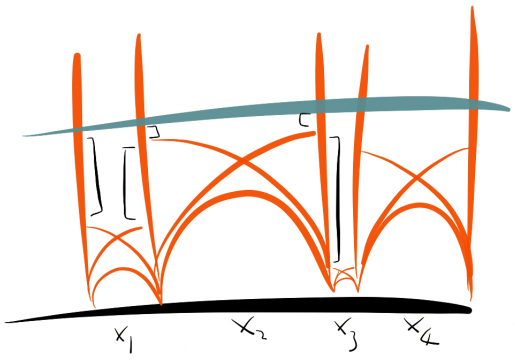


Figure 9: Shears along a horocycle.

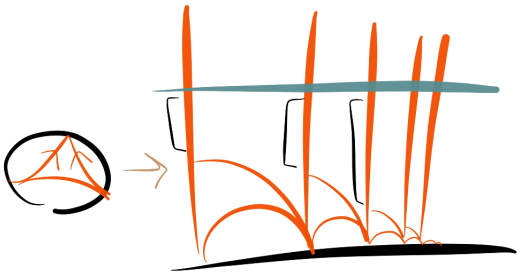


Figure 10: A triangle with two sides identified whose midpoints don't match.

We still get a hyperbolic metric on the punctured monogon, but it is not complete. Trying to draw a horocycle in the upper half-plane will give a spiral which does not close in on itself.

However, we can fix this by completing. In the upper half-plane we do this by adding the limit geodesic of the geometric series of lifts, which appears as extra geodesic boundary.

Exercise 12.2. *What is the length of the new boundary?*

Exercise 12.3. *What determines the direction of spiraling? (There are two things that are spiraling, namely the horocycle and the geodesic arc in the triangulation.)*

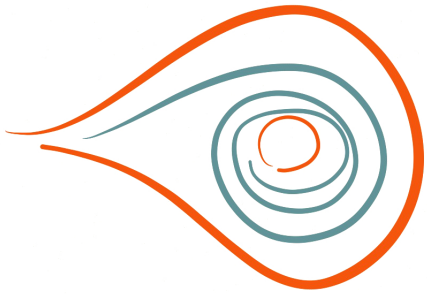


Figure 11: Geodesic boundary and a spiraling horocycle.

Hence the map from λ-coordinates to τ-coordinates has image in a Teichmüller space which includes the possibility of geodesic boundaries where the cusps were. It factors through a Teichmüller space which only describes cusps.

274 Curves on Surfaces, Lecture 12

Dylan Thurston

Fall 2012

13 Laminations and compactifying Teichmüller space

We want to compactify (decorated) Teichmüller space $\tilde{\mathcal{T}}_{g,n}$. One reason we might want to do this is the following: the mapping class group acts on Teichmüller space and we would like to understand the resulting dynamics. However, on a noncompact space this is difficult to do. We have an embedding from decorated Teichmüller space to $\mathbb{R}^I$ given by taking lengths of arcs in a triangulation, but this uses a choice of triangulation and we would like something that doesn't depend on choices so that we can get a natural action of the mapping class group.

We therefore take the lengths of all arcs rather than arcs in a fixed triangulation. This gives now an embedding in an infinite-dimensional Euclidean space. This space in turn embeds into a corresponding projective space, and we define the compactification to be the closure in this projective space. (We can also do this for undecorated Teichmüller space by taking the lengths of all simple closed curves instead of the lengths of arcs in a triangulation, which gives a compactification $\overline{\mathcal{T}}_{g,n}$.) This is the *Thurston compactification*.

The mapping class group still acts on the Thurston compactiication. Moreover:

Theorem 13.1. *$\overline{\mathcal{T}}_{g,n}$ is a ball.*

It follows by the Brouwer fixed point theorem that any element of the mapping class group acting on $\overline{\mathcal{T}}_{g,n}$ has a fixed point.

What happens to the Ptolemy relation in the compactification?

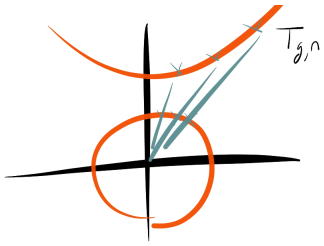


Figure 1: A heuristic picture of a sequence in Teichmüller space heading to infinity in its compactification.

Working in a projective space corresponds to rescaling lengths by some constant:

$$\ell^R(A) = \frac{\ell(A)}{k}. \tag{1}$$

The effect on λ-lengths is

$$\lambda^R(A) = \lambda(A)^{1/k}. \tag{2}$$

The Ptolemy relation now looks like

$$\lambda^R(E)\lambda^R(F) = \lambda^R(A)\lambda^R(C) \oplus_k \lambda^R(B)\lambda^R(D) \tag{3}$$

where

$$x \oplus_k y = \sqrt[k]{x^k + y^k}. \tag{4}$$

As $k \to \infty$, we have

$$\lim_{k\to\infty} x \oplus_k y = \max(x, y). \tag{5}$$

Rewriting in terms of renormalized lengths, we get

$$\ell^R(E) + \ell^R(F) = \max(\ell^R(A) + \ell^R(C), \ell^R(B) + \ell^R(D)). \tag{6}$$

Proposition 13.2. *The above identity holds at all points in compactified Teichmüller space which do not lie in Teichmüller space.*

In other words, compactifying tropicalizes the Ptolemy relation.

Earlier we saw this relation when looking at intersection numbers of curves crossing a triangulation. This suggests that simple closed curves give points in compactified Teichmüller space. We would like to see this geometrically.

Begin with a surface with a hyperbolic structure and a simple closed curve L in that surface. We will construct a sequence of hyperbolic structures going to infinity by placing a long neck where L is.

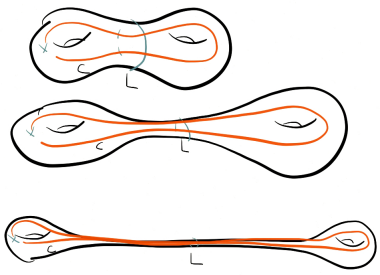


Figure 2: A lengthening neck.

As the neck gets longer, the length of any curve C is dominated by the number of times it crosses L (up to a constant); that is,

$$\ell(C) = k \cdot i(C, L) + O(1). \tag{7}$$

As $k \to \infty$, the lengths of curves approach their intersection numbers in projective space.

More formally, fix a hyperbolic metric Σ_0. Find a geodesic representative for L. Insert a Euclidean cylinder of width t at L. This is enough to get a conformal structure, and then we can uniformize to get a new hyperbolic structure Σ_t.

Example Consider the punctured torus, so the Teichmüller space $\mathcal{T}_{1,1}$. Recall that this can be naturally identified with the hyperbolic plane (which is mildly confusing), and the mapping class group is the usual action of $\mathrm{SL}_2(\mathbb{Z})$ (so the moduli space can be identified with the usual fundamental domain). Punctured tori can be represented as the quotient of $\mathbb{C}$ by the discrete subgroup spanned by two linearly independent $z, w \in \mathbb{C}$.

The boundary of Teichmüller space should be the usual boundary of the hyperbolic plane. Simple closed curves on the torus with rational slope can then be identified with points on the boundary by looking at how inserting cylinders increases the length of the vectors z, w.

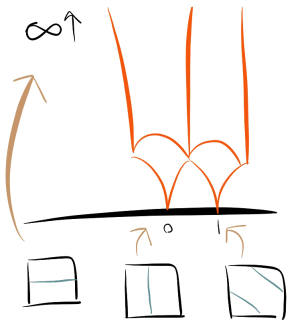


Figure 3: Simple closed curves in a torus and the corresponding boundary points in $\overline{\mathcal{T}}_{1,1}$.

Exercise 13.3. *What simple closed curve corresponds to the limit point at* $\frac{3}{5}$*?*

We now know that limit points of the compactification of Teichmüller space correspond to tropical solutions to the Ptolemy relations. Integral (rational) limit points correspond to simple closed curves, but this only gives countably many limit points. The real limit points correspond to *laminations*.

More precisely, an *integral decorated lamination* L is a collection of distinct simple closed nontrivial curves with positive weights except that curves surrounding a marked point can have negative weight. These weights describe possible limit behaviors of sequences of decorated hyperbolic structures.

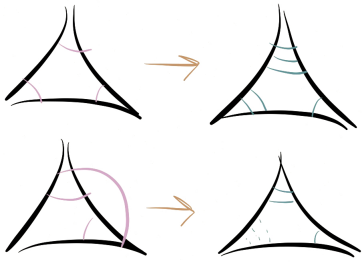


Figure 4: A lamination. The dashed line indicates a negative weight.

We can measure lengths by counting intersection numbers with weights:

$$\ell(A) = i(L, A). \tag{8}$$

Exercise 13.4. *What should the weight of a notched arc be?*

Before we move on to laminations, we should return to the punctured torus with a line of rational slope. $\mathbb{R}^2$ minus the punctures has a hyperbolic metric, and hyperbolically, geodesics will avoid punctures.

More precisely:

Theorem 13.5. *Around each cusp in a hyperbolic surface, simple geodesics do not enter in a horocycle of circumference* 2*.*

In other words, although lines of rational slope look straight in the Euclidean sense, they clump up dramatically in the hyperbolic metric. These clumps look like Cantor sets.

We are now ready to discuss the complete definition of a lamination. A *geodesic lamination* of a hyperbolic surface is a collection of simple geodesics $\{\gamma_\alpha\}$ whose union $\bigcup \gamma_\alpha$ is closed.

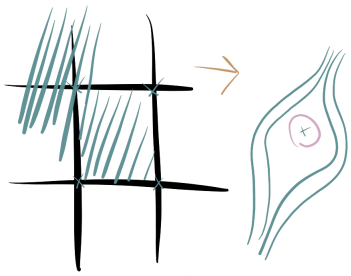


Figure 5: Geodesics avoiding a puncture.

Example Any simple closed curve is a geodesic lamination.

Example Any geodesic that spirals to a closed geodesic (together with the limiting closed geodesic) is a geodesic lamination.

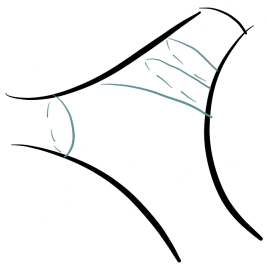


Figure 6: Some atypical examples of laminations.

These are not typical examples. Typically a lamination looks locally like a Cantor set cross the interval.

We want weights as above. A *measured lamination* L is a lamination with a *transverse measure*, namely, a measure on each transverse arc supported on its intersection of L which is transversely invariant in the sense that transverse arcs which are transverse isotopic have the same measure.

Measured laminations give rise to coordinates as follows: we associate to an arc between punctures its total measure.

It is not obvious how to write down measured laminations.

Example A spiraling lamination has no positive measure. Isotoping an arc around the spiral shows that no part of the spiral has positive weight.

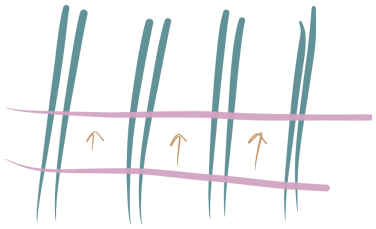


Figure 7: A transverse isotopy between transverse arcs.

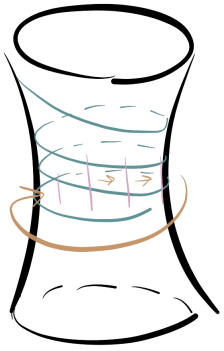


Figure 8: A spiraling lamination.

However, we can write down measured laminations using *train tracks*.

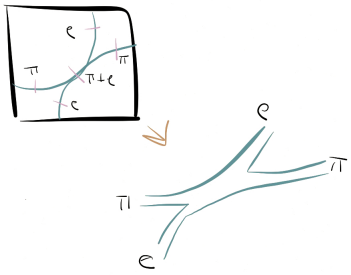


Figure 9: A train track.

Train tracks have an operation defined on them called *splitting*, and by repeatedly splitting a train track we get something approximating a measured lamination.

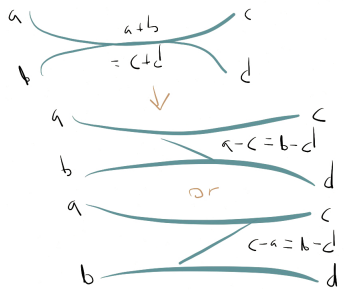


Figure 10: Splitting a train track.

274 Curves on Surfaces, Lecture 13

Dylan Thurston

Fall 2012

14 More about laminations

Recall that last time we were trying to understand the Thurston compactification of (decorated) Teichmüller space. When we did this we got points whose coordinates satisfied the tropical Ptolemy relations. We can imagine doing this for other kinds of Teichmüller space:

Teichmüller space	Coordinates	Integer limits	Coordinates
$\tilde{\mathcal{T}}_{g,n}$	λ-lengths	Integer laminations	Intersections
$\mathcal{T}_g$	Lengths of closed curves	Integer laminations	Intersections
$\mathcal{T}_{g,0,n}$	Shear coordinates	?	?

Here $\mathcal{T}_{g,0,n}$ denotes the Teichmüller space of surfaces with no punctures but with geodesic boundary, and we should take logarithms of shear coordinates before projectivizing because they behave multiplicatively. We do not yet know what the corresponding limit points of the Thurston compactification are.

Recall that one geometric interpretation of the cross-ratio was in terms of the distance (shear) between two midpoints obtained by dropping angle bisectors: if d is this distance, then $\tau = e^d$. Recall also that we obtained limit points by inserting a simple multicurve and inserting necks; then in the limit, lengths become proportional to intersection numbers.

Accordingly, we should count intersections of the segment between two midpoints; these are also called shear coordinates. We will count these with sign: in one direction they will be assigned $+1$ while in another direction they will be assigned -1.

Alternately, for any choice of decoration we had

$$\tau(E) = \frac{\lambda(B)\lambda(D)}{\lambda(A)\lambda(C)}. \tag{1}$$

Tropicalizing this relation gives almost, but not quite, the right answer: since $\lambda(A) = e^{\frac{\ell}{2}}$, we should take

$$\text{shear}(E) = \frac{\ell(B) + \ell(D) - \ell(A) - \ell(C)}{2}. \tag{2}$$

If we don't want to use decorations, we should work with infinitely many curves since ideal polygons have sides of infinite length. This gives in the limit a kind of metric space where the distance between two points is given by the intersection number of a geodesic between them. We can identify points at distance zero, and then a triangle with infinitely many curves at the vertices (since an ideal triangle has infinite length) becomes an infinite tripod tree. Gluing two such triangles together to obtain a quadrilateral gives us a metric tree.

This gives limit points described by multicurves such that there are infinitely many curves around punctures, some of which may be spiraling (to account for geodesic boundary). These are *unbounded laminations*.

Exercise 14.1. *How do these shear coordinates transform under change of triangulation? Compare to the geometric answer (from points in noncompactified Teichmüller space).*

When we draw a geodesic representative of a complicated multicurve on, say, the 4-punctured sphere, it does not look much like a curve because it becomes very close to itself. What we get looks more like a train track. As the curve becomes more complicated (under the action of the mapping class group) we get something which should describe a limit point in the compactification, which corresponds to a train track with real weights. In this particular case we should replace Fibonacci numbers with the corresponding powers of the golden ratio.

Exercise 14.2. *Apply the mapping class group element we have been applying to the limit train track. Check that you get the same thing up to scale and splitting.*

(It is probably a better idea to apply the inverse, which will make the train track simpler rather than more complicated.)

This reflects the fact that the Thurston compactification is a ball, so any element of the mapping class group has a fixed point. In general, studying the action of the mapping class group on the Thurston compactification led Thurston to the following theorem.

Theorem 14.3. *Every element ϕ of the mapping class group is in one of the following three categories:*

1. *ϕ has finite order.*
2. *ϕ is* reducible*: it fixes a finite collection of closed curves (e.g. Dehn twist) (possibly permuting them).*
3. *ϕ is* pseudo-Anosov*: it fixes exactly two points in the Thurston compactification, neither of which are points in Teichmüller space. Call these two projective measured laminations L_+, L_-. The lamination L_+ is, in an appropriate sense, $\lim_{n\to\infty} \phi^n(c)$ for any nontrivial multicurve c, and L_- is, in an appropriate sense, $\lim_{n\to\infty} \phi^{-n}(c)$.*

This is analogous to the classification of hyperbolic isometries into elliptic, parabolic, and hyperbolic elements. For example, ϕ finite order turns out to be equivalent to the claim that ϕ fixes a hyperbolic structure.

274 Curves on Surfaces, Lecture 15

Dylan Thurston
Notes by Qiaochu Yuan

Fall 2012

16 Orbifolds (Felikson)

Roughly speaking, a (2-dimensional) orbifold is the quotient of a surface by the action of a discrete group. These look like surfaces, possibly with boundary, some of whose points are orbifold points.

Example Consider the quotient of the disk D^2 by negation. This is just a cone; the cone point (orbifold point) is at the origin and has an angle of π.

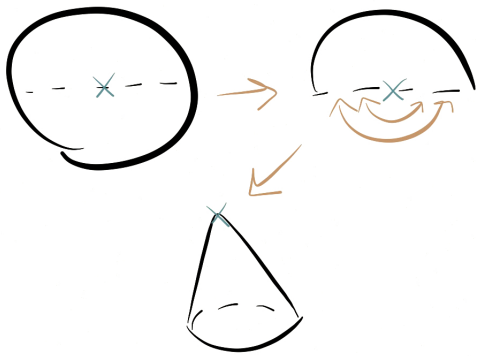


Figure 1: Quotienting a disk to obtain an orbifold.

Orbifolds are relevant to the complete statement of the classification of cluster algebras of finite mutation type. Recall that this classification was previously stated as follows:

Theorem 16.1. *(Shapiro, Tumarkin, Felikson) A cluster algebra with skew-symmetric matrix of finite mutation type is one of the following types:*

1. *Rank* 2,
2. *A surface cluster algebra,*
3. 11 *other exceptions.*

However, the full classification is for skew-symmetrizable matrices (matrices B such that there exists a diagonal matrix D such that BD is skew-symmetric). Now, we know that skew-symmetric matrices can be represented by quivers and quivers coming from surfaces are the block-decomposable ones.

Skew-symmetrizable matrices may be studied using certain diagrams (which do not capture everything about the matrix). We associate to an entry $b_{ij} > 0$ in such a matrix a pair of vertices i, j and an edge between them labeled $-b_{ij}b_{ji}$. These diagrams may be mutated like quivers as follows: a mutation at k

1. reverses arrows incident to k,
2. modifies triangles incident to k and their labels in such a way that $\sqrt{r} + \sqrt{r'} = \sqrt{pq}$ in the image below.

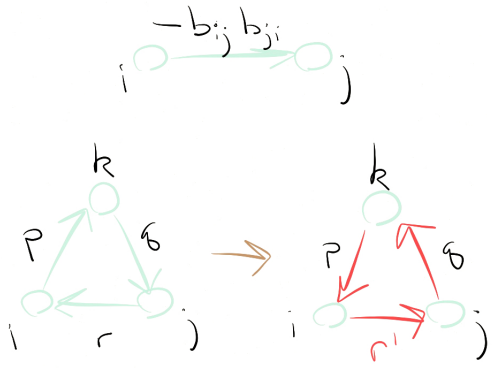


Figure 2: A diagram mutation.

The more general classification is the following.

Theorem 16.2. *(Shapiro, Tumarkin, Felikson) A cluster algebra with skew-symmetrizable matrix of finite mutation type is one of the following types:*

1. *Rank* 2,
2. *Obtained from blocks,*
3. $11 + 7$ *other exceptional diagrams.*

The blocks used here include the blocks we used to construct quivers coming from surfaces, but with five new blocks. They may be thought of as coming from triangulations of orbifolds.

We require all of the orbifold points to be cone points of angle π. A triangulation of such an orbifold is a maximal compatible set of arcs, where an arc may have endpoints either marked points or orbifold points, and compatibility means non-crossing and

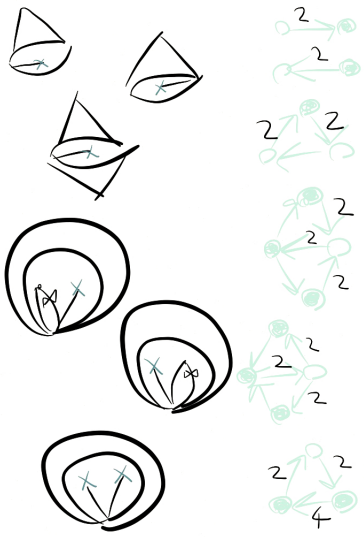


Figure 3: Blocks and the triangulations of orbifolds from which they arise.

also means that we forbid two arcs from having endpoint the same orbifold point. We allow our arcs to be tagged on ends which are not attached to orbifold points (since orbifold points do not have horocycles associated to them).

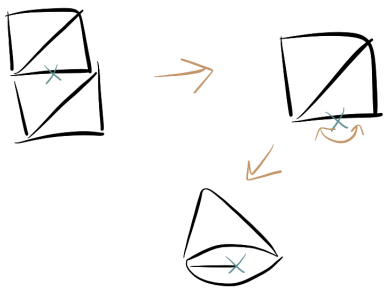


Figure 4: Quotienting a triangulation to obtain an orbifold triangulation.

We also need to decide what flips between triangulations are for arcs with an orbifold endpoint. These are obtained by removing the arc and drawing the only other compatible arc.

Theorem 16.3. *Flips act transitively on tagged triangulations of orbifolds.*

We can build diagrams from tagged triangulations of orbifolds in the same way as before, except that edges coming to or from an arc with an orbifold endpoint should be labeled by a 2 (and edges coming both to and from an arc with an orbifold endpoint should be labeled by a 4). Mutation of diagrams then corresponds to flips of triangulations, and any diagram we obtain in this way decomposes into blocks.

We cannot reconstruct a skew-symmetrizable matrix from a diagram because we have lost information in only labeling by $-b_{ij}b_{ji}$. This can be fixed by assigning weights to vertices given by the corresponding entries of the diagonal matrix D symmetrizing the matrix.

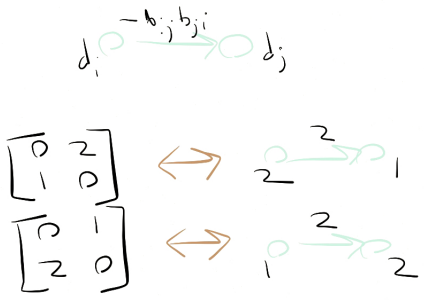


Figure 5: Using weights on vertices to reconstruct matrix entries.

In a block decomposition with weights, all white vertices necessarily have the same weight w, and all black vertices have weight either $2w$ or $\frac{w}{2}$.

Hence we should use weighted orbifolds, namely orbifolds such that their orbifold points are marked either 2 or $\frac{1}{2}$. These are the geometric objects associated to the second type of cluster algebra in the theorem above. From such data we can choose a triangulation, then build a weighted diagram, from which we can recover a skew-symmetrizable matrix.

We would now like to introduce hyperbolic structure (so we can assign λ-lengths, etc.). We first restrict our attention to the case that all of the orbifold points have weight $\frac{1}{2}$. We make all ordinary triangles ideal hyperbolic triangles, and we obtain orbifold points by symmetrically gluing ideal hyperbolic triangles (dropping an angle bisector). We may now decorate with horocycles (keeping in mind that there are no horocycles around orbifold points), and we define λ-lengths as before, with the λ-length of an arc with an orbifold endpoint the λ-length of its lift to a symmetrically glued ideal hyperbolic triangle.

The Ptolemy relation can look different here.

To accommodate orbifold points of weight 2, it is necessary to replace them by special marked points. These marked points are special because they assigned a self-conjugate horocycle.

With the above convention, λ-lengths parameterize the Teichmüller space associated to the orbifold.

Example The quiver B_n can be realized using one orbifold point. It is a quotient of D_n.

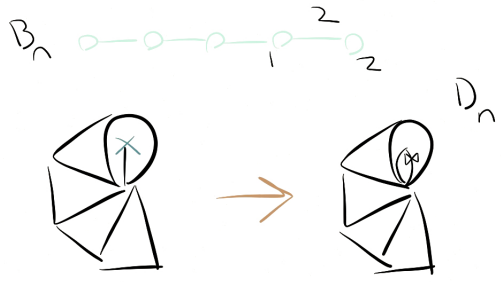


Figure 6: The B_n quiver as a quotient of the D_n quiver.

Example The quiver C_n can be realized using one orbifold point. It is a quotient of A_{2n-1}.

It is a natural question to ask whether any orbifold can be realized as a quotient of an ordinary surface. If there are an even number of orbifold points of weight $\frac{1}{2}$, we can pair them up and cut along lines connecting them. If there is more than one orbifold point of weight $\frac{1}{2}$, then we can pair up two of them and cut, and this doubles the remaining number of orbifold points. If the boundary is nonzero, we can cut along a line connecting an orbifold point and a point on the boundary. Finally, if our surface is not S^2, we can cut along a non-contractible loop.

The remaining case is when the surface is S^2 with one orbifold point of weight $\frac{1}{2}$; this orbifold cannot be realized as a quotient.

Laminations and skein relations work well in this formalism.

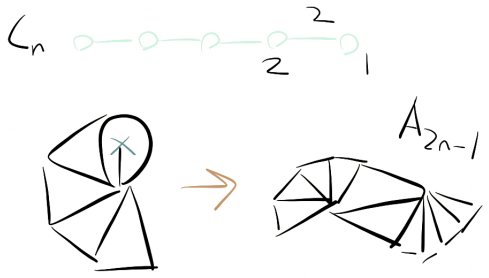


Figure 7: The C_n quiver as a quotient of the A_{2n-1} quiver.

274 Curves on Surfaces, Lecture 15

Dylan Thurston
Notes by Qiaochu Yuan

Fall 2012

16 Orbifolds (Felikson)

Roughly speaking, a (2-dimensional) orbifold is the quotient of a surface by the action of a discrete group. These look like surfaces, possibly with boundary, some of whose points are orbifold points.

Example Consider the quotient of the disk D^2 by negation. This is just a cone; the cone point (orbifold point) is at the origin and has an angle of π.

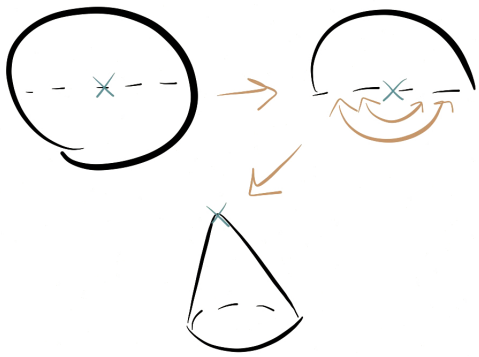


Figure 1: Quotienting a disk to obtain an orbifold.

Orbifolds are relevant to the complete statement of the classification of cluster algebras of finite mutation type. Recall that this classification was previously stated as follows:

Theorem 16.1. *(Shapiro, Tumarkin, Felikson) A cluster algebra with skew-symmetric matrix of finite mutation type is one of the following types:*

1. *Rank* 2,
2. *A surface cluster algebra,*
3. 11 *other exceptions.*

However, the full classification is for skew-symmetrizable matrices (matrices B such that there exists a diagonal matrix D such that BD is skew-symmetric). Now, we know that skew-symmetric matrices can be represented by quivers and quivers coming from surfaces are the block-decomposable ones.

Skew-symmetrizable matrices may be studied using certain diagrams (which do not capture everything about the matrix). We associate to an entry $b_{ij} > 0$ in such a matrix a pair of vertices i, j and an edge between them labeled $-b_{ij}b_{ji}$. These diagrams may be mutated like quivers as follows: a mutation at k

1. reverses arrows incident to k,
2. modifies triangles incident to k and their labels in such a way that $\sqrt{r} + \sqrt{r'} = \sqrt{pq}$ in the image below.

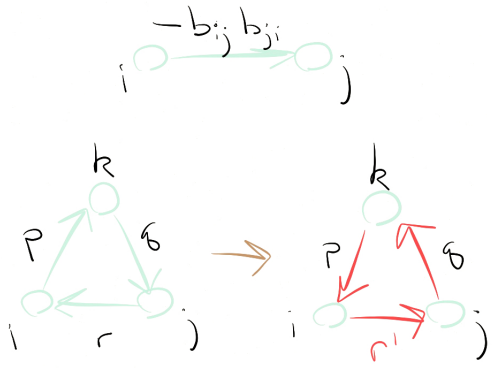


Figure 2: A diagram mutation.

The more general classification is the following.

Theorem 16.2. *(Shapiro, Tumarkin, Felikson) A cluster algebra with skew-symmetrizable matrix of finite mutation type is one of the following types:*

1. *Rank* 2,
2. *Obtained from blocks,*
3. $11 + 7$ *other exceptional diagrams.*

The blocks used here include the blocks we used to construct quivers coming from surfaces, but with five new blocks. They may be thought of as coming from triangulations of orbifolds.

We require all of the orbifold points to be cone points of angle π. A triangulation of such an orbifold is a maximal compatible set of arcs, where an arc may have endpoints either marked points or orbifold points, and compatibility means non-crossing and

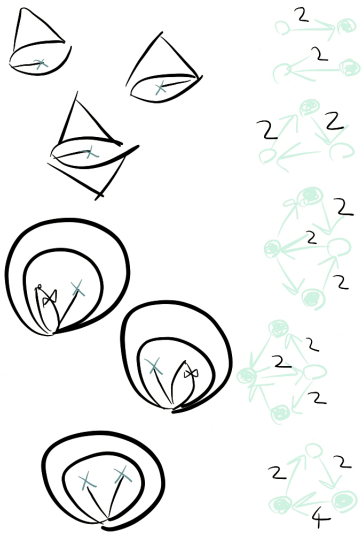


Figure 3: Blocks and the triangulations of orbifolds from which they arise.

also means that we forbid two arcs from having endpoint the same orbifold point. We allow our arcs to be tagged on ends which are not attached to orbifold points (since orbifold points do not have horocycles associated to them).

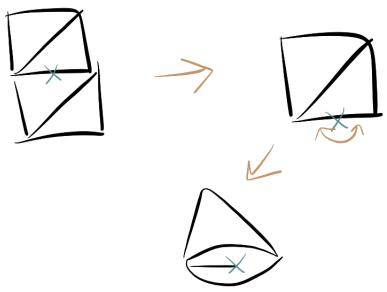


Figure 4: Quotienting a triangulation to obtain an orbifold triangulation.

We also need to decide what flips between triangulations are for arcs with an orbifold endpoint. These are obtained by removing the arc and drawing the only other compatible arc.

Theorem 16.3. *Flips act transitively on tagged triangulations of orbifolds.*

We can build diagrams from tagged triangulations of orbifolds in the same way as before, except that edges coming to or from an arc with an orbifold endpoint should be labeled by a 2 (and edges coming both to and from an arc with an orbifold endpoint should be labeled by a 4). Mutation of diagrams then corresponds to flips of triangulations, and any diagram we obtain in this way decomposes into blocks.

We cannot reconstruct a skew-symmetrizable matrix from a diagram because we have lost information in only labeling by $-b_{ij}b_{ji}$. This can be fixed by assigning weights to vertices given by the corresponding entries of the diagonal matrix D symmetrizing the matrix.

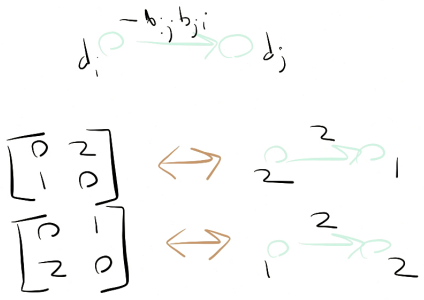


Figure 5: Using weights on vertices to reconstruct matrix entries.

In a block decomposition with weights, all white vertices necessarily have the same weight w, and all black vertices have weight either $2w$ or $\frac{w}{2}$.

Hence we should use weighted orbifolds, namely orbifolds such that their orbifold points are marked either 2 or $\frac{1}{2}$. These are the geometric objects associated to the second type of cluster algebra in the theorem above. From such data we can choose a triangulation, then build a weighted diagram, from which we can recover a skew-symmetrizable matrix.

We would now like to introduce hyperbolic structure (so we can assign λ-lengths, etc.). We first restrict our attention to the case that all of the orbifold points have weight $\frac{1}{2}$. We make all ordinary triangles ideal hyperbolic triangles, and we obtain orbifold points by symmetrically gluing ideal hyperbolic triangles (dropping an angle bisector). We may now decorate with horocycles (keeping in mind that there are no horocycles around orbifold points), and we define λ-lengths as before, with the λ-length of an arc with an orbifold endpoint the λ-length of its lift to a symmetrically glued ideal hyperbolic triangle.

The Ptolemy relation can look different here.

To accommodate orbifold points of weight 2, it is necessary to replace them by special marked points. These marked points are special because they assigned a self-conjugate horocycle.

With the above convention, λ-lengths parameterize the Teichmüller space associated to the orbifold.

Example The quiver B_n can be realized using one orbifold point. It is a quotient of D_n.

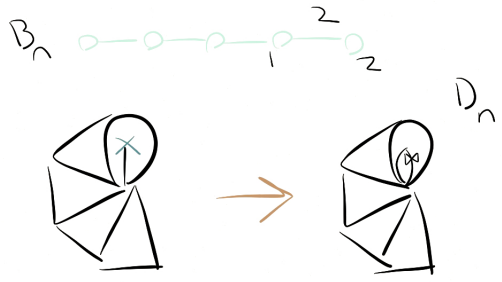


Figure 6: The B_n quiver as a quotient of the D_n quiver.

Example The quiver C_n can be realized using one orbifold point. It is a quotient of A_{2n-1}.

It is a natural question to ask whether any orbifold can be realized as a quotient of an ordinary surface. If there are an even number of orbifold points of weight $\frac{1}{2}$, we can pair them up and cut along lines connecting them. If there is more than one orbifold point of weight $\frac{1}{2}$, then we can pair up two of them and cut, and this doubles the remaining number of orbifold points. If the boundary is nonzero, we can cut along a line connecting an orbifold point and a point on the boundary. Finally, if our surface is not S^2, we can cut along a non-contractible loop.

The remaining case is when the surface is S^2 with one orbifold point of weight $\frac{1}{2}$; this orbifold cannot be realized as a quotient.

Laminations and skein relations work well in this formalism.

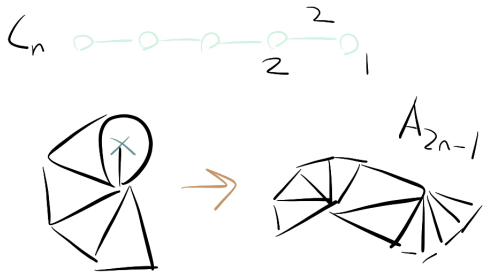


Figure 7: The C_n quiver as a quotient of the A_{2n-1} quiver.

274 Curves on Surfaces, Lecture 17

Dylan Thurston

Fall 2012

18 The Laurent phenomenon (Kalman)

The Laurent phenomenon is a phenomenon by which various recurrences defined by rational functions turn out to be Laurent polynomials in the first few terms. If the first few terms are set to 1, then the remaining terms become integers even though the recurrence divides by previous members of the term. For example, consider the sequence

$$y_k = \frac{y_{k-3}y_{k-1} + y_{k-2}^2}{y_{k-4}} \tag{1}$$

with initial conditions $y_1 = y_2 = y_3 = y_4$ (Somos-4).

Cluster algebras provide a natural setting for studying the Laurent phenomenon via the exchange relation (Fomin, Zelevinsky).

Theorem 18.1. *In a cluste ralgebra, any cluster variable is expressed in terms of any given cluster as a Laurent polynomial with coefficients in the group ring $\mathbb{ZP}$.*

We will instead prove a more general result, the Caterpillar Lemma. This is a statement about a sequence $\mathbb{T}_{n,m}$ of trees. This has m vertices of degree n in its spine.

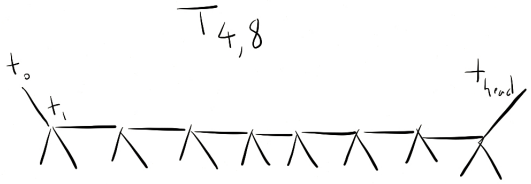


Figure 1: The caterpillar $T_{4,8}$.

We will label the edges emanating from each vertex with different labels and we will associate an exchange polynomial $P \in \mathbb{A}[x_1, ...x_n]$, not involving x_k, to every edge (here $\mathbb{A}$ is a UFD). This is a generalized exchange pattern.

Associate to t_0 the initial cluster $\mathbf{x}(t_0)$ of n independent variables. To each vertex $t \in \mathbb{T}_{n,m}$, we associate a cluster $\mathbf{x}(t)$. The variables in this cluster are uniquely determined by the exchange relations, for an edge labeled by k and P:

$$x_i(t) = x_i(t'), i \neq k \tag{2}$$

$$x_k(t)x_k(t') = P(\mathbf{x}(t)). \tag{3}$$

Lemma 18.2. *(Caterpillar) Suppose a generalized exchange pattern on $\mathbb{T}_{n,m}$ satisfies the following conditions:*

1. *for any edge labeled by k and P, the polynomial P does not involve x_k and is not divisible by any x_i.*
2. *For any two consecutive edges labeled by i, P and j, Q, the polynomials P and $Q_0 = Q|_{x_i=0}$ are coprime.*
3. *For any three consecutive edges labeled by i, P and j, Q and i, R, we have*

$$LQ_0^b P = R_{x_j \leftarrow \frac{Q_0}{x_j}} \tag{4}$$

where b is a non-negative integer and L is a Laurent monomial whose coefficients lie in $\mathbb{A}$ and which is coprime to P.

Then each $x_i(t), t \in \mathcal{T}_{n,m}$ is a Laurent polynomial in $x_1(t_0), ...x_n(t_0)$ with coefficients in $\mathbb{A}$.

Proof. For $t \in \mathcal{T}_{n,m}$ let $L(t)$ be the ring of Laurent polynomials in $\mathbf{x}(t)$. Abbreviate $L_0 = L(t_0)$. We proceed by induction on m (the length of the spine). This is straightforward for $m = 1$, so assume $m \geq 2$ and that the statement is true for all caterpillars with smaller spine.

We will need a lemma. Suppose the path from t_0 to t_{head} starts with edges labeled by i and j and consider the unique next edge labeled by i. Then the clusters $\mathbf{x}(t_1), \mathbf{x}(t_2), \mathbf{x}(t_3)$ are all in L_0, and in addition $\gcd(x_i(t_3), x_i(t_1)) = \gcd(x_j(t_2), x_i(t_1)) = 1$. The proof of this last statement involves the third condition.

Returning to the proof, by the inductive hypothesis $X = x_k(t_{\text{head}})$ belongs to both $L(t_1)$ and $L(t_3)$. We further claim that $X = \frac{f}{x_i(t_1)^a}$ for some $f \in L_0$ and some $a \geq 0$. This follows from the fact that $X \in L(t_1)$ and that $x_i(t_1) = \frac{P}{x_i} \in L_0$.

Similarly, we claim that $X = \frac{g}{x_j(t_2)^b x_i(t_3)^c}$ for some $g \in L_0$ and some $b, c \geq 0$. This follows from the fact that $X \in L(t_3)$, the fact that $x_i(t_3), x_j(t_3) \in L_0$ by the lemma, and the fact that $x_j(t_3) = x_j(t_2) \in L_0$.

We conclude that

$$X = \frac{f}{x_i(t)^a} = \frac{g}{x_j(t_2)^a x_i(t_3)^c}. \tag{5}$$

From the second part of the lemma, $\gcd(x_i(t_3), x_i(t_1)) = 1, \gcd(x_j(t_2), x_i(t_1)) = 1$, so $X \in L_0$ as desired. □

Good news: the Caterpillar Lemma can prove Laurentness in many situations. Bad news: it is often not trivial to rephrase a given problem in the Caterpillar Lemma framework. For example, when describing the caterpillar graph for the Somos-4 sequence, it is not obvious what the exchange polynomials on the legs should be. See Fomin-Zelevinsky for details.

19 Miscellaneous

Some sequences exhibiting the Laurent phenomenon like the Somos-7 sequence

$$x_k x_{k+7} = x_{k+1}x_{k+6} + x_{k+2}x_{k+5} + x_{k+3}x_{k+4} \tag{6}$$

cannot be described using cluster algebras (we would need the RHS to consist of two terms). Lam and Pylyavskyy have a notion of Laurent phenomenon algebra that addresses this.

Consider surfaces with boundary components and no marked points.

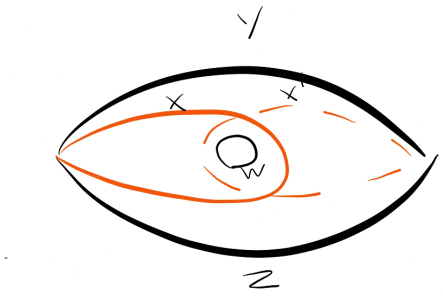


Figure 2: A surface with boundary components and no marked points.

We can write down what the exchange relations for the corresponding cluster algebra looks like using the skein relation.

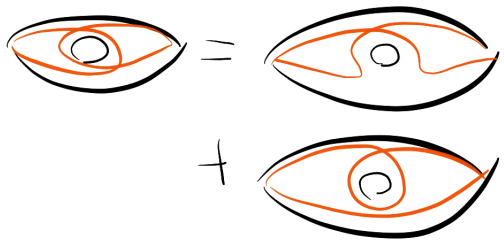


Figure 3: An application of the skein relation.

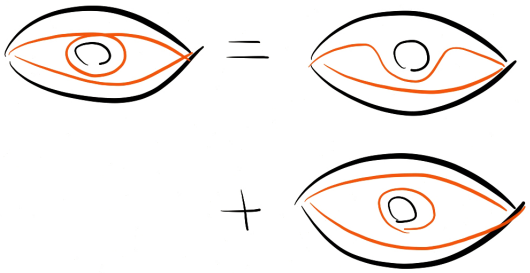


Figure 4: Another application of the skein relation.

This gives an exchange relation with three terms on the RHS (so not a cluster algebra): more specifically,

$$xx' = y^2 + z^2 + yzw. \tag{7}$$

There are two special cases. When $w = 0$, this has two terms. When $w = 2$, we should take $\sqrt{x}, \sqrt{x'}$ as the cluster variables, and then their product is $y + z$.

More geometrically, recall that for closed curves the λ-length is $2\cosh\left(\frac{\ell}{2}\right)$, or the trace of the monodromy. When $w = 2$ we have $\ell = 0$ (a cusp) and when $w = 0$ we are instead at a cone point and we should replace the hyperbolic cosine with the usual cosine of the angle, which is π at an order 2 cone point. In the case that $w = 2$, we

added square roots because we are looking at a self-conjugate horocycle.

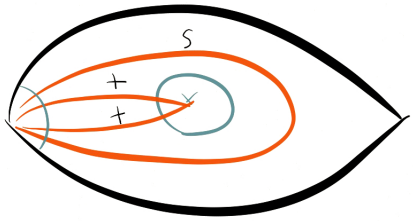


Figure 5: Self-conjugate horocycles and square roots. Here $t = \sqrt{s}$.

Let's return to skein theory. Last time we forgot an extra relation, namely that a loop around a puncture evaluates to 2.

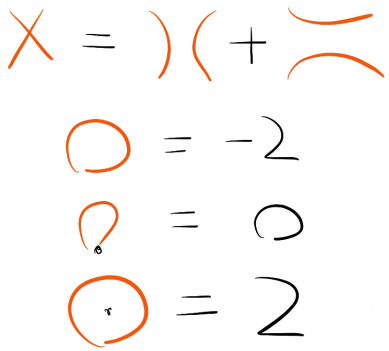


Figure 6: Relations for skein theory.

Not including this relation gives a sensible skein algebra, but one which has a zero divisor.

With this extra relation it is not completely obvious that these relations are consistent (that is, that the basis is still what we expect it to be). This can be proven using the diamond lemma (Bergman). This is a lemma about a system of reductions (e.g. relations replacing some terms by other terms in a presentation of an algebra). This system of reductions is required to have the following properties:

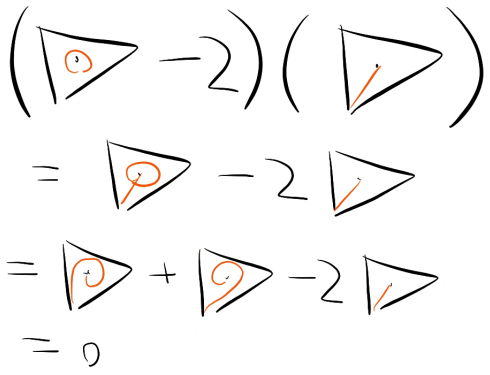


Figure 7: A zero divisor.

1. Local confluence: if R_1, R_2 are two reductions of some x, there is some further chain of reductions which makes them equal.

2. Any chain of reductions terminates.

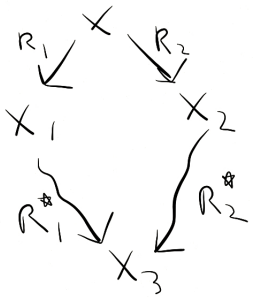


Figure 8: A diamond describing local confluence.

With these hypotheses, any two chains of reductions from x ends at the same place. This proceeds by induction on the length of chains.

These hypotheses are straightforward to verify for the skein relations.

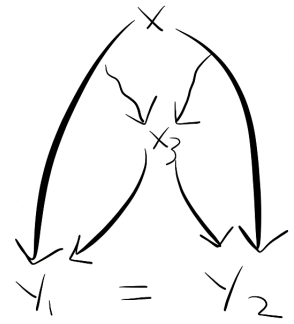


Figure 9: The inductive step.

274 Curves on Surfaces, Lecture 18

Dylan Thurston

Fall 2012

20 Skein theory and the Laurent phenomenon

Skein theory for tagged arcs requires an extra relation beyond the ones we have used so far.

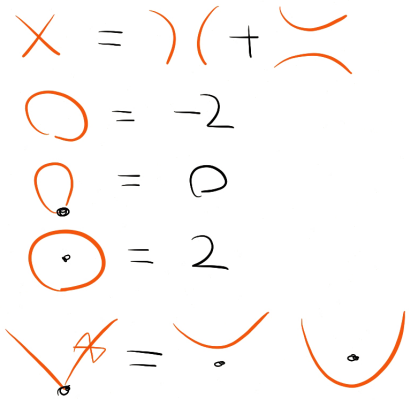


Figure 1: The complete set of relations.

Exercise 20.1. *Check that these reductions satisfy the hypotheses of the diamond lemma.*

We can induce an order on diagrams by considering its complexity, which is a triple $(t(D), c(D), r(D))$ where $t(D)$ is the number of tag mismatches, $c(D)$ is the number of crossings, and $r(D)$ is the number of reducible components (loops bounding disks and arcs bounding monogons), ordering diagrams by lexicographic order on their complexity, and ordering linear combinations by sorting the complexities of the components.

Theorem 20.2. *The skein algebra $Sk(\Sigma)$ of a surface Σ has the following properties:*

1. *It has a basis of simple curves (those with complexity $(0, 0, 0)$).*
2. *Elements are invariant under regular isotopy (RII, RIII, RIIb) and change by -1 under RI.*
3. *Multiplication is given by superimposing curves and is well-defined (independent of how the superposition is done).*

If T is a triangulation of Σ, we have a sequence of inclusions

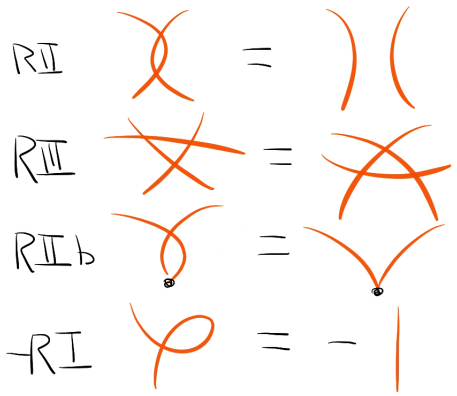


Figure 2: The various Reidemeister moves.

$$\mathbb{Z}[T] \subset A(\Sigma) \subset \mathrm{Sk}(\Sigma) \subset A^+(\Sigma) \subset \mathbb{Z}[T^{\pm 1}] \subset \mathbb{Q}(T) = K(\Sigma) \tag{1}$$

where $\mathbb{Z}[T]$ is the ring of polynomials in the arcs in T, $A(\Sigma)$ is the cluster algebra of Σ, $A^+(\Sigma)$ is the upper cluster algebra (the rational functions of the arcs that are integral Laurent polynomials in any cluster), $\mathbb{Z}[T^{\pm 1}]$ is the ring of integral Laurent polynomials in the arcs in T, and $\mathbb{Q}(T) = K(\Sigma)$ is the fraction field of any of the above.

The inclusion $\mathrm{Sk}(\Sigma) \subset A^+(\Sigma)$ is obtained as follows. As for quantum skeins, we have the following lemma.

Lemma 20.3. *If A is a simple arc and C any curve, then for sufficiently large n, $\langle C \cdot A^n \rangle \in Sk(\Sigma)$ has only terms which are compatible with (do not cross) A.*

It follows that if T is a triangulation consisting of arcs A_i, then we can find a monomial $\prod A_i^{n_i}$ such that $\langle C \cdot \prod A_i^{n_i} \rangle$ is compatible with T, hence is a polynomial in the A_i. Once we know that we can divide by the A_i, it follows that $\langle C \rangle \subset \mathbb{Z}[T^{\pm 1}]$ is a Laurent polynomial in every cluster.

Proof. Without tags, assume that C is simple and that C and A intersect n times. Expanding all of these intersections using the skein relations, we usually obtain two terms with $n-1$ intersections and all of the remaining terms have even fewer intersections. (However, sometimes it is possible to get fewer than $n-1$ intersections.) □

Example Consider an arc C in a triangulated annulus with arcs A_0, A_1. We compute that

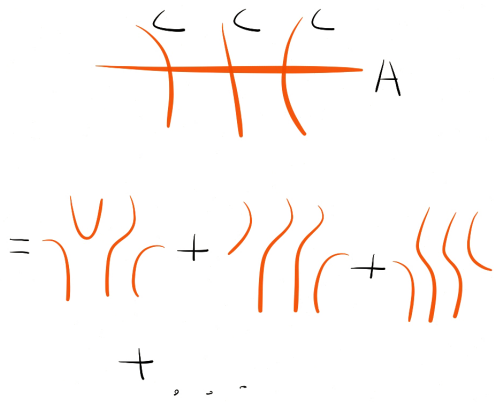


Figure 3: Expanding intersections.

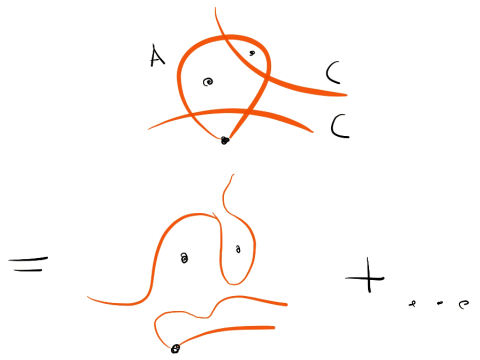


Figure 4: An example in which we get fewer than $n-1$ intersections.

$$\langle C \cdot A_0 \rangle = \langle A_1 \rangle + \langle A_{-1} \rangle \tag{2}$$

where

$$\langle A_1 \cdot A_{-1} \rangle = \langle A_0^2 \rangle + \langle B_1 \cdot B_2 \rangle. \tag{3}$$

Hence as a Laurent polynomial,

$$C = \frac{A_1}{A_0} + \frac{A_0}{A_1} + \frac{B_1 B_2}{A_0 A_1}. \tag{4}$$

This is an example of the following fairly general result.

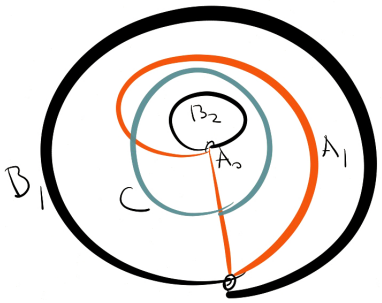


Figure 5: An arc in a triangulated annulus.

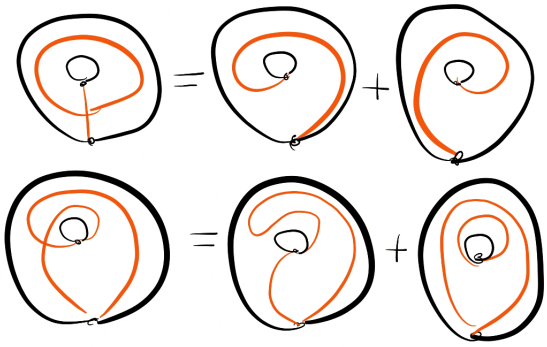


Figure 6: Skein relations in the annulus.

Theorem 20.4. *(Weak positivity) In many cases, any simple curve C expressed as a Laurent polynomial has all positive integer coefficients. These coefficients have combinatorial interpretations.*

Another condition we might ask for is strong positivity, namely that there is a basis x_i of some algebra (over some ordered ring) such that $x_i x_j = \sum n_{ij}^k x_k$ where $n_{ij}^k \geq 0$.

Exercise 20.5. *Find an example in the annulus where strong positivity fails with respect to the simple curves basis. More precisely, consider the curve A_k which is an arc wrapped around k times, and find the first k such that $A_0 A_k$ is not positive.*

It is an interesting problem to find natural strongly positive bases.

Theorem 20.6. *The map $Sk(\Sigma) \to A^+(\Sigma)$ defined above is injective.*

Proof. (Sketch) It is enough to show that multiplication by an arc E is invertible. We can order diagrams by their intersection with E, which has two dominant terms as above given by the left smooth and the right smoothing. We modify the order by approximately a shear; more precisely, we order by $-i(\cdot, A) + i(\cdot, B) - i(\cdot, C) + i(\cdot,)$ where A, B, C, D is a quadrilateral with E as its diagonal, and this picks out one dominant term. Now multiplication by E looks upper triangular. □

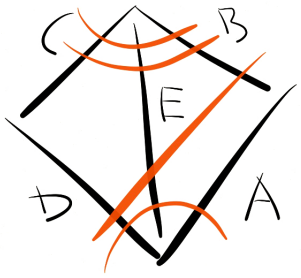


Figure 7: The modification of the order.

Conjecturally we in fact have $\mathrm{Sk}(\Sigma) = A^+(\Sigma)$ unless Σ has one puncture or has tagged arcs and one puncture.

274 Curves on Surfaces, Lecture 19

Dylan Thurston

Fall 2012

21 Strong positivity

Earlier we wrote down a basis (the bangles basis) of the skein algebra that is not strongly positive. Here we will use a different basis, the bracelet basis. It is composed of *bracelet basis curves*, which are diagrams where

1. No two components intersect,
2. No component intersects itself except for multiple covers of a single loop,
3. We do not have two powers of the same simple loop,
4. There are no tag mismatches, disks, punctured disks, or monogons.

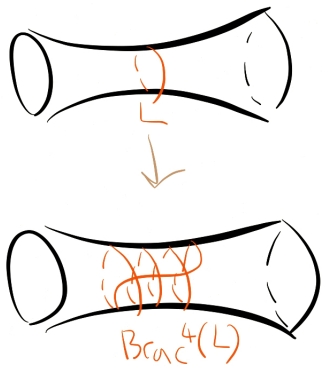


Figure 1: A bracelet basis curve.

Theorem 21.1. *The bracelets basis of $Sk(\Sigma)$ is strongly positive.*

Conjecturally the bracelets basis of $\mathrm{Sk}_q(\Sigma)$ is also strongly positive. There is also another basis, the bands basis, which is not known to be strongly positive for surfaces with punctures. It is strongly positive for the annulus with two or more marked points, and in this case it agrees with Lusztig's dual canonical basis (Lampe).

The Chebyshev polynomials $T_n(z)$ of the first kind are defined by the initial conditions $T_0(z) = 2, T_1(z) = z$, and

$$T_{n+1}(z) = zT_n(z) - T_{n-1}(z) \tag{1}$$

and satisfy

$$T_k(z)T_\ell(z) = T_{k+\ell}(z) + T_{|k-\ell|}(z). \tag{2}$$

They are also uniquely determined by the condition that

$$T_k(e^\ell + e^{-\ell}) = e^{k\ell} + e^{-k\ell}. \tag{3}$$

Lemma 21.2. *For $n \geq 1$, we have $\langle Brac^n(L)\rangle = T_n(\langle L\rangle)$.*

Proof. Check for $n = 1, 2$. For higher values of n apply the skein relation to the outermost crossing. □

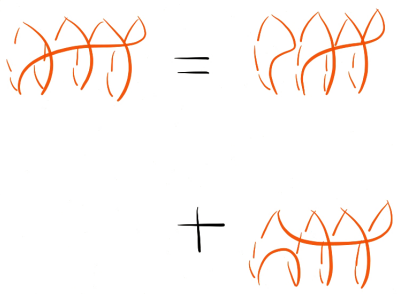


Figure 2: The skein relation applied to a bracelet.

Conjecturally the bracelets basis is an *atomic basis*: for any element of the skein algbra which is not positive in the bracelets basis, there exists a triangulation T with respect to which this element is not a positive Laurent polynomial.

Question from the audience: is the bands basis related to Jones-Wenzl idempotents?

Answer: yes, for $q = 1$. In representation-theoretic terms these correspond to symmeric powers of the defining representation.

For the unpunctured torus, there is a basis parameterized by pairs (k, ℓ) up to the identification $(k, \ell) = (-k, -\ell)$ given by taking

$$T_{(k,\ell)} = \mathrm{Brac}^{\gcd(k,\ell)}(\text{line of slope } \frac{\ell}{k}) \tag{4}$$

and this basis is strongly positive in $\mathrm{Sk}_q(\Sigma)$ (Frohman-Gelca).

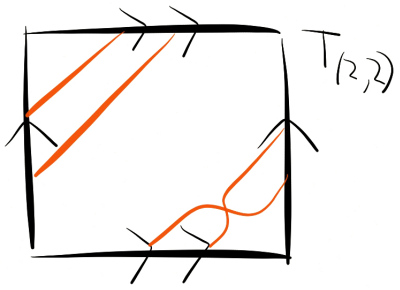


Figure 3: The basis element $T_{(2,2)}$.

Theorem 21.3. *Any diagram D is sign-definite in the bracelets basis in $Sk(\Sigma)$. Any* taut *diagram (immersed with minimal number of self-intersections, no disks, no monogons, no punctured disks) is positive in the bracelets basis.*

Say that a diagram is *positive* if it has no singular 0-gons or singular 1-gons. By this we mean the following. A *segment* of a diagram D is an interval between intersection points or endpoints. A *k-chain* is a sequence of k segments meeting cyclically at endpoints. A 0-*chain* is a loop component of a diagram. For a k-chain C, we can smooth out corners to obtain a loop C^0. A *singular k-gon* is a k-chain C such that C^0 is null-homotopic.

Theorem 21.4. *(Hass-Scott) Any curve diagram can be turned into a taut diagram by Reidemeister reductions (Reidemeister moves that make the diagram less complicated), together with the deletion of boundary loops and unknotted loops.*

This can be done greedily. The idea is to shorten the curves, but this is done combinatorially rather than geometrically.

Corollary 21.5. *A positive curve diagram can be turned into a taut curve diagram using only RIII, RII, RIIb (regular isotopy).*

Proof. By Hass-Scott, we can do this using all of the moves above. We need to show that when doing this we do not use RI, R0 (delete an unknotted loop), or R0b (delete a boundary loop). Before using RI, we have a singular 1-gon, and it suffices to show that going backwards we started out with a singular 1-gon. This follows from an inspection of each move. □

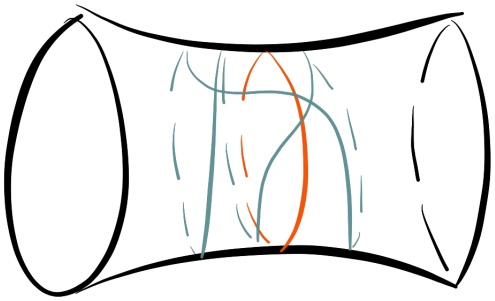


Figure 4: A counterexample.

It is not true that D is taut iff it has no singular 0-gons, 1-gons, or 2-gons. There is a counterexample in an annulus.

The correct salvage is to weaken the notion of singular 2-gon.

Exercise 21.6. *What goes wrong with the going backwards argument for 2-gons?*

We now return to the claim that every taut diagram is positive in the bracelets basis. Say that a *good crossing* of a taut diagram is a crossing where both resolutions are positive.

Lemma 21.7. *Every taut diagram that is not a bracelets basis curve has either a good crossing, two components that are powers of the same loop, or a tag mismatch.*

This is enough to prove the theorem. Let D be a taut diagram. If D is in the bracelets basis, we are done. Otherwise, we are in one of the cases above. In the first case, we resolve the crossing. In the second case, we apply the Chebyshev polynomial identity. The third case is similar to the first.

Question from the audience: is there an inner product with respect to which the bracelets basis is orthogonal?

Answer: that is an interesting question.

Question from the audience: does the bracelets basis admit a monoidal categorification?

Answer: this is not known. Conjecturally there should be a monoidal abelian category $C(\Sigma)$ whose simple objects correspond to the bracelets basis and whose monoidal product reproduces the structure constants of the skein algebra in the bracelets basis in the sense that the Grothendieck ring of $C(\Sigma)$ should be $\mathrm{Sk}(\Sigma)$.

274 Curves on Surfaces, Lecture 20

Dylan Thurston

Fall 2012

22 More about strong positivity

Today we will ignore tags.

Let D be a diagram. Recall that a crossing in D is positive if both of its resolutions are positive, where positivity means no singular 0-gons and 1-gons. Recall also that a diagram is taut if it has the minimum number of self-intersections.

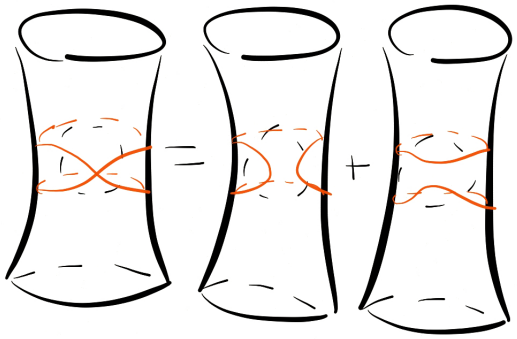


Figure 1: A crossing that is not positive.

Definition A *multi-bracelet* is a diagram where two components don't intersect and each component is a simple arc or a bracelet (and there are no 0-gons and 1-gons).

Lemma 22.1. *Any taut diagram that is not a multi-bracelet has a positive crossing.*

To show this we will use the following.

Lemma 22.2. *If D is taut but a crossing of D resolves into D_1, D_2 where D_1 is not positive, then D_1 has a singular 0-gon or 1-gon passing through the reducing disk (the disk surrounding the crossing in which we apply the skein relation) twice.*

Proof. Since D_1 is not positive, it has a singular 0-gon or 1-gon H. If H does not pass through the reducing disk, then we get a 0-gon or 1-gon for D, which contradicts tautness. If H passes through the reducing disk once, then we get a 1-gon or 2-gon for D, which also contradicts tautness. □

Exercise 22.3. *Find more examples of resolutions $D = D_1 + D_2$ where D is taut and D_1 is not positive. Check the lemma above in your examples.*

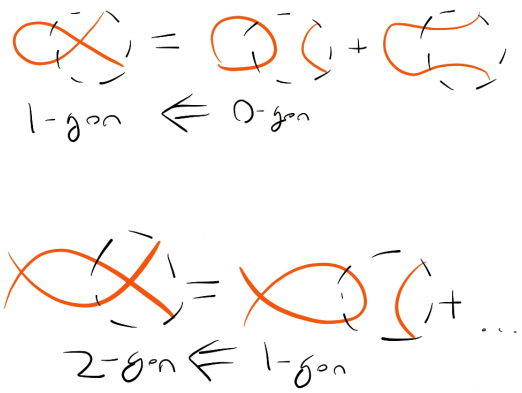


Figure 2: Singular polygons in D_1 and singular polygons in D.

Definition A *bracelet chain* in D is a 0-chain or 1-chain H such that the smoothing H^0 is homotopic to a bracelet. A *maximal bracelet chain* is a bracelet not contained in any larger bracelet chain for the same loop.

Lemma 22.4. *Every component of a taut diagram D with at least one self-crossing has a maximal bracelet chain.*

Proof. Let C be a component with a self-crossing. Then it has a 1-chain. Take a minimal 1-chain by inclusion. This is a 1-chain H such that H^0 is a simple loop L. Take H' to be a maximal bracelet chain containing H (which is also a bracelet for L). □

Lemma 22.5. *The crossing at the end of a maximal bracelet 1-chain is positive.*

Proof. Let H be such a maximal bracelet. Tne resolution of the crossing is not connected to the rest of the diagram locally, so a 0-gon or 1-gon cannot pass through the reducing disk twice. There are two possible cases a), b) involving the other resolution which must be ruled out.

To rule out case a), write the maximal bracelet as $\gamma^k \in \pi_1$ for k maximal and γ a loop. Write the rest of the diagram as $\rho \in \pi_1$. If we get a singular 1-gon in the first case, then $\rho\gamma^k = \mathrm{id}$, or $\rho = \gamma^{-k}$, which contradicts the maximality of H.

To rule out case b), with notation as above, we have $\rho\gamma^\ell = \mathrm{id}$ for some $\ell \leq k$. It follows that the entire component is a bracelet, which contradicts the maximality of H. □

We are getting close to the proof of the first lemma; it suffices now to consider crossings between components.

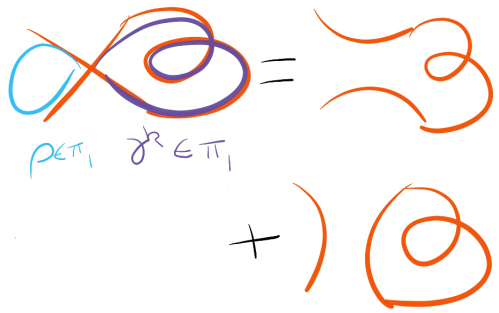


Figure 3: Part of a maximal bracelet.

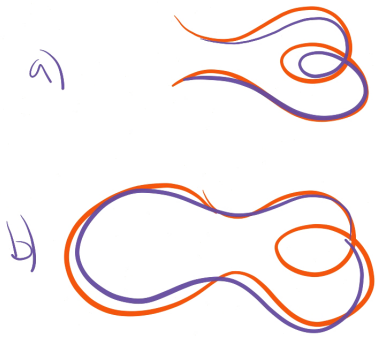


Figure 4: The two cases above.

Exercise 22.6. *Show that any crossing between two components C_1, C_2 of a taut diagram D where C_1, C_2 are simple arcs or bracelets is positive.*

Use the fact that roots are unique in $\pi_1(\Sigma)$: that is, if $\gamma^k = \rho^\ell$, then there exists σ, s, t such that $\gamma = \sigma^s, \rho = \sigma^t$, and $sk = t\ell$.

Exercise 22.7. *Prove the multiplication rule $T_{(a,b)}T_{(c,d)} = T_{(a+c,b+d)} + T_{(a-c,b+d)}$ for the basis for the unpunctured torus (at $q = 1$) from last time.*

Now we will discuss a geometric interpretation of the skein relations. Here we will ignore marked points and arcs. The skein relation should have something to do with SL_2. More precisely, it should have something to do with the following fact: if $A, B \in \mathrm{SL}_2$, then

$$\mathrm{tr}(A)\mathrm{tr}(B) = \mathrm{tr}(AB) + \mathrm{tr}(AB^{-1}). \tag{1}$$

This follows from the Cayley-Hamilton theorem, which gives $B^2 - \mathrm{tr}(B)B + I = 0$, after dividing by B and multiplying by A, then taking traces.

A, B should be the monodromy of two loops, except that the signs don't match.

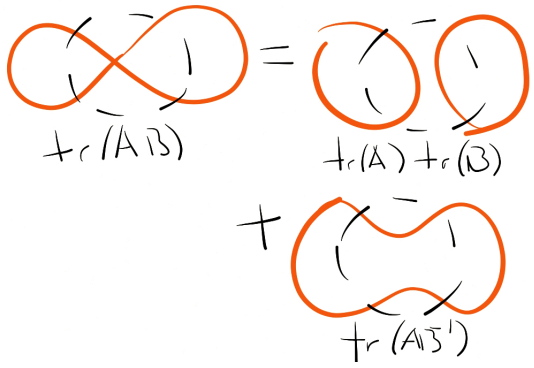


Figure 5: The signs are incorrect here.

This is because we were taking $q = 1$ and we actually need to take $q = -1$, which gives a skein relation in which three terms add up to 0. Geometrically we should take the negative of the trace. More precisely,

Proposition 22.8. *If Σ is a surface and $\rho : \pi_1(\Sigma) \to SL_2$ is a representation, then the assignment*

$$D \mapsto \prod_i -tr(\rho(D_i)) \tag{2}$$

where D is a diagram with components D_i satisfies the $q = -1$ skein relation.

However, taking $q = -1$ destroys positivity. To get back to $q = 1$ we need to twist. There is a fibration

$$S^1 \to \mathrm{UT}(\Sigma) \to \Sigma \tag{3}$$

where UT is the unit tangent bundle. This gives a long exact sequence in homotopy

$$\pi_2(\Sigma) \to \pi_1(S^1) \to \pi_1(\mathrm{UT}(\Sigma)) \to \pi_1(\Sigma) \to 0 \tag{4}$$

and if Σ is not S^2 then $\pi_2(\Sigma)$ vanishes, hence $\pi_1(\mathrm{UT}(\Sigma)) = \tilde{\pi}_1(\Sigma)$ is a canonical $\mathbb{Z}$-extension of $\pi_1(\Sigma)$. We define a twisted SL_2 representation to be a representation $\rho : \tilde{\pi}_1(\Sigma) \to \mathrm{SL}_2$ such that $\rho(360^\circ \text{ turn}) = -1$. The corresponding quotient map to PSL_2 gives an honest PSL_2-representation.

If γ is an immersed curve in Σ, by taking tangent vectors it lifts to a curve $\tilde{\gamma}$ in $\mathrm{UT}(\Sigma)$, and associated to this choice of lift is a trace. (If we had just used a PSL_2-representation, the trace would only be defined up to sign.)

Proposition 22.9. *For Σ a surface and $\rho : \tilde{\pi}_1(\Sigma) \to SL_2$ a twisted SL_2-representation, the assignment*

$$D \mapsto \prod_i tr(\rho(\tilde{D}_i)) \tag{5}$$

satisfies the $q = 1$ skein relation.

In the background here is the fact that a hyperbolic structure on Σ has a canonical twisted SL_2-representation lifting the PSL_2-representation given by considering the universal cover.

This is related to spin structures. Recall that $\pi_1(\mathrm{SO}(n)) \cong \mathbb{Z}/2\mathbb{Z}$ for $n \geq 3$; in particular it is not simply connected. When $n = 2$ we have $\pi_1(\mathrm{SO}(2)) \cong \mathbb{Z}$. In any case, for $n \geq 2$, it follows that $\mathrm{SO}(n)$ has a unique double cover called $\mathrm{Spin}(n)$, and for $n \geq 3$ this is the universal cover. There are exceptional isomorphisms $\mathrm{Spin}(3) \cong \mathrm{SU}(2)$ and $\mathrm{Spin}(4) \cong \mathrm{SU}(2) \times \mathrm{SU}(2)$.

A *spin structure* on a smooth oriented (Riemannian for simplicity but this is not necessary) manifold M is a lift of the frame bundle to a $\mathrm{Spin}(n)$-bundle. Concretely, this gives us some information about which loops in the frame bundle lift and which do not. On an oriented surface Σ, rather than thinking about frames we can think about tangent vectors, and then the question is whether or not a path of tangent vectors lifts. We can generate paths of tangent vectors using an immersed curve. This gives us a rule assigning connected immersed curves signs ± 1 satisfying some rules.

A twisted SL_2-representation can then be described as the product of an ordinary SL_2-representation and a spin structure.

274 Curves on Surfaces, Lecture 21

Dylan Thurston

Fall 2012

23 More about the geometry of skein relations

Recall that a twisted SL_2-representation of a surface Σ is an SL_2-representation of a $\mathbb{Z}$-extension $\tilde{\pi}_1(\Sigma) \cong \pi_1(\mathrm{UT}(\Sigma))$ of the fundamental group such that a 360° rotation acts by -1. Such a representation in particular descends to an ordinary PSL_2-representation of $\pi_1(\Sigma)$. Defining

$$\tilde{\rho}(D) = \prod_i \mathrm{tr}(\tilde{\rho}(\tilde{D}_i)) \tag{1}$$

where D is a curve diagram, $\tilde{D}_i$ are the lifts of the components of D to $\mathrm{UT}(\Sigma)$, and $\tilde{\rho}$ is a twisted representation, we claim that $\tilde{\rho}(D)$ satisfies the skein relations with $q = 1$.

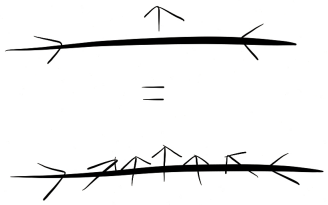


Figure 1: An immersed curve and its tangent vectors.

We should say more about paths in the unit tangent bundle. We can notate these by writing down an immersed curve in Σ together with arrows indicating how the unit tangent vectors should rotate. This notation satisfies some straightforward axioms.

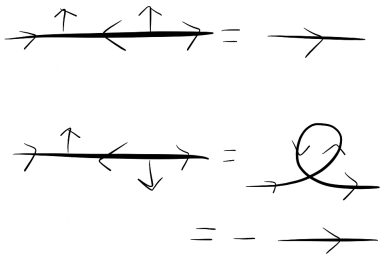


Figure 2: Some axioms.

To verify the skein relations, there are two cases depending on how the skeins close up into curves. We will apply the trace relation $\mathrm{tr}(A)\mathrm{tr}(B) = \mathrm{tr}(AB) + \mathrm{tr}(AB^{-1})$, but keeping track of tangent vectors.

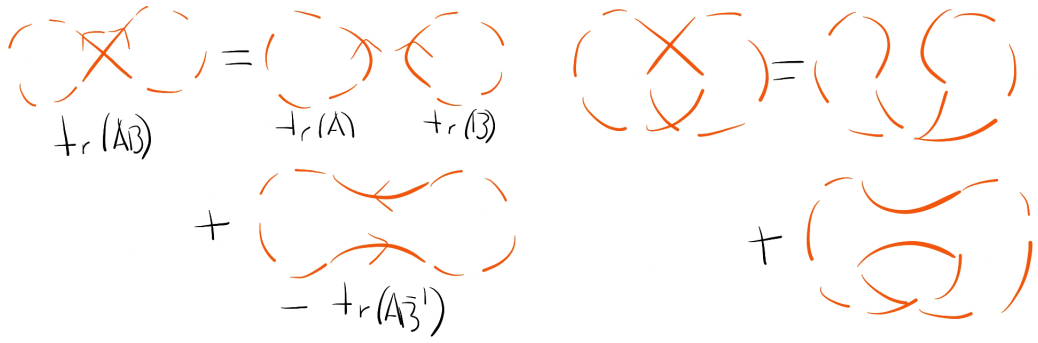


Figure 3: The two cases.

Exercise 23.1. *Verify the skein relation in the second case.*

The relationship to spin structures is the following. Concretely, a double cover of the frame bundle is a homomorphism η from π_1 of the frame bundle to $\mathbb{Z}/2\mathbb{Z}$. For a surface Σ the frame bundle is essentially the unit tangent bundle, so a spin structure on Σ assigns a sign to every path in $\mathrm{UT}(\Sigma)$, hence to immersed curves in Σ. This assignment satisfies various axioms.

Alternatively, we can think of η as an element of $H^1(\mathrm{UT}(\Sigma), \mathbb{Z}/2\mathbb{Z})$ which is non-trivial on the 360° rotation. There is an action of $H^1(\Sigma; \mathbb{Z}/2\mathbb{Z})$ on the above cohomology group (by translation), hence $H^1(\Sigma; \mathbb{Z}/2\mathbb{Z})$ acts on the set of spin structures of Σ. If spin structures exist, the set of all such spin structures is then a torsor over $H^1(\Sigma; \mathbb{Z}/2\mathbb{Z})$. More concretely, if η_1, η_2 are two spin structures, their difference is an element of $H^1(\Sigma; \mathbb{Z}/2\mathbb{Z})$.

Proposition 23.2. *If $\rho : \pi_1(\Sigma) \to SL_2$ is an SL_2-representation of the fundamental group and $\eta : \pi_1(UT(\Sigma)) \to \mathbb{Z}/2\mathbb{Z}$ is a spin structure, then $\rho\eta : \pi_1(UT(\Sigma)) \to SL_2$ is a twisted SL_2-representation.*

Conversely, we have an identification

$$\mathrm{Rep}^{\mathrm{Twist}}_{\mathrm{SL}_2}(\Sigma) = (\mathrm{Rep}_{\mathrm{SL}_2}(\Sigma) \times \mathrm{Spin}(\Sigma))/H^1(\Sigma; \mathbb{Z}/2\mathbb{Z}) \tag{2}$$

where $H^1(\Sigma; \mathbb{Z}/2\mathbb{Z}) \cong \mathrm{Hom}(\pi_1, \mathbb{Z}/2\mathbb{Z})$ acts diagonally.

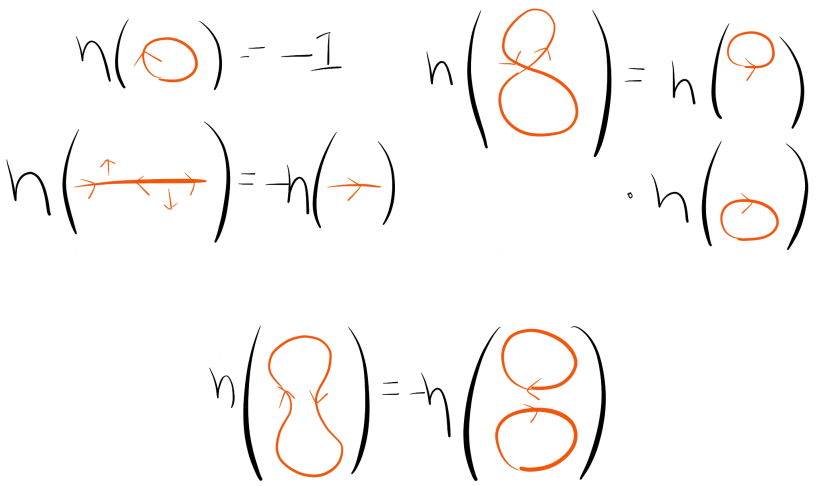


Figure 4: Some properties of spin structures.

Exercise 23.3. *What are the spin structures on the torus? What are their orbits under the action of the mapping class group?*

The skein algebra at $q = 1$ describes a class of functions on the set of twisted SL_2-representations. Moreover it has a strongly positive basis which is invariant under the action of the mapping class group. Passing to ordinary SL_2-representations by multiplication by a spin structure, we lose this invariance because not all spin structures are preserved by the action of the mapping class group.

Now suppose Σ is equipped with a hyperbolic structure. Then the universal cover of Σ is naturally identified with the hyperbolic plane $\mathbb{H}^2$, so we can write $\Sigma \cong \mathbb{H}^2/\Gamma$ where $\Gamma \cong \pi_1(\Sigma)$ is a discrete subgroup of $\mathrm{PSL}_2(\mathbb{R})$. Thus a hyperbolic structure determines a (faithful) representation

$$\pi_1(\Sigma) \to \mathrm{PSL}_2(\mathbb{R}). \tag{3}$$

This is part of the structure defining a twisted SL_2-representation. The different lifts of this representation to an SL_2-representation are classified by spin structures and people usually pick one.

However, there is a canonical such lift. To see this, write $\mathbb{H}^2$ as the quotient of $\mathrm{PSL}_2(\mathbb{R})$ by $\mathrm{SO}(2)$. This gives an identification of $\mathrm{UT}(\mathbb{H}^2)$ with $\mathrm{PSL}_2(\mathbb{R})$, hence an identification

$$\mathrm{UT}(\Sigma) \cong \mathrm{PSL}_2(\mathbb{R})/\Gamma \tag{4}$$

using the fact that the identification $\Sigma \cong \mathbb{H}^2/\Gamma$ respects tangent spaces. This in turn gives an identification

$$\mathrm{UT}(\Sigma) \cong \mathrm{SL}_2(\mathbb{R})/\tilde{\Gamma} \tag{5}$$

where $\tilde{\Gamma}$ is the preimage of Γ in $\mathrm{SL}_2(\mathbb{R})$ (some $\mathbb{Z}/2\mathbb{Z}$ central extension). This is not quite $\tilde{\pi}_1(\Sigma)$, but there is a diagram

$$\begin{array}{ccccc} \mathbb{Z} & \longrightarrow & \tilde{\pi}_1(\Sigma) & \longrightarrow & \pi_1(\Sigma) \\ \downarrow & & \downarrow & & \downarrow \cong \\ \mathbb{Z}/2\mathbb{Z} & \longrightarrow & \tilde{\Gamma} & \longrightarrow & \Gamma \end{array} \tag{6}$$

relating them.

Alternatively, consider the boundary of $\mathbb{H}^2$, which is naturally identified with $\mathbb{RP}^1$. This is naturally acted on by $\mathrm{PSL}_2(\mathbb{R})$. On the other hand, given a tangent vector to a point in $\mathbb{H}^2$ we can follow a unique geodesic to the boundary, which gives a natural identification

$$\mathrm{UT}_p(\Sigma) \cong \mathrm{UT}_p(\mathbb{H}^2) \cong \mathbb{RP}^1. \tag{7}$$

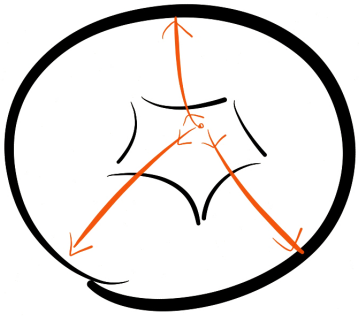


Figure 5: Identifying unit tangents with the boundary.

This identification can also be used to obtain the above result.

It follows that the Teichmuller space for Σ embeds into $\mathrm{Rep}^{\mathrm{Twist}}_{\mathrm{SL}_2}$.

Proposition 23.4. *The Teichmüller space for Σ is the totally positive part of $Rep_{SL_2}^{Twist}$: that is, it is the part where all elements in the positive basis for the skein algebra take positive values.*

To take decorations into account, we need a notion of twisted decorated SL_2-representation. We will first think of twisted SL_2-representations as twisted SL_2-local systems (2-dimensional real vector bundles V over $UT(\Sigma)$ such that the projectivization of V is the pullback of the unit tangent bundle over Σ). To decorate them, we want the additional data of a choice of vector in V_p for each outward-pointing tangent vector at a boundary point.

274 Curves on Surfaces, Lecture 22

Dylan Thurston

Fall 2012

24 Additive categorification of surface cluster algebras (Christof)

Surface cluster algebras can be categorified (additively) as follows. We consider certain triangulated 2-Calabi-Yau $\mathbb{C}$-linear categories C with a cluster tilting object $T = T_1 \oplus ... \oplus T_n$. Triangulated categories generalize derived categories: their main feature is the existence of a self-equivalence $\Sigma : C \to C$ and a collection of distinguished triangles

$$X \to Y \to Z \to \Sigma X \tag{1}$$

satisfying various axioms and generalizing exact sequences. One of these is that for every morphism $X \xrightarrow{f} Y$ there exists a distinguished triangle

$$X \xrightarrow{f} Y \to Z \to \Sigma X. \tag{2}$$

2-Calabi-Yau means that there is a natural isomorphism

$$\mathrm{Hom}(X, Y) \cong \mathrm{Hom}(Y, \Sigma^2 X)^* \tag{3}$$

where * denotes the linear dual. In particular,

$$\mathrm{Hom}(X, \Sigma Y) \cong \mathrm{Hom}(Y, \Sigma X)^*. \tag{4}$$

Cluster tilting means that for every X we have

$$\mathrm{Hom}(T, \Sigma X) = 0 \Leftrightarrow X \text{ is direct summand of } T. \tag{5}$$

In particular, there is a functor

$$F : C \ni X \mapsto \mathrm{Hom}(T, \Sigma X) \in \mathrm{End}(T)^{op}\text{-Mod} \tag{6}$$

whose kernel is given by morphisms which factor through T.

The above conditions imply that every Hom-space is finite-dimensional. In particular, $\mathrm{End}(T)^{op}$ is finite-dimensional. It can be written as $\mathbb{C}[Q]/I$ where $\mathbb{C}[Q]$ is the path algebra of a quiver Q and I is an admissible ideal. This quiver Q is canonically determined by the algebra.

The summands $T = T_1 \oplus ... \oplus T_n$ can be exchanged; for each i there exists $T_i' \neq T_i$ such that $T/T_i \oplus T_i'$ is again a cluster tilting object. The corresponding quiver Q changes according to quiver mutation.

There is a map called the cluster character sending an object Z to a certain sum

$$C^T(Z) = \underline{x}^{\underline{g}(Z)} \sum_{\underline{\ell}} \chi(\mathrm{Gr}_{\underline{\ell}}^{Q^{op}}(F(Z)))\underline{y}^{\underline{\ell}} \tag{7}$$

over Euler characteristics of quiver Grassmannians which evaluates to a Laurent polynomial in $\mathbb{Z}[x_1^{\pm}, ...x_n^{\pm}]$. Here $g(Z)$ is defined as follows: if

$$\Sigma^{-1}Z \to T_Z^{\underline{b}} \to T_Z^{\underline{a}} \to Z \tag{8}$$

is a distinguished triangle, where $\underline{a} = (a_1, a_2, ...)$ and

$$T_Z^{\underline{a}} = T_1^{a_1} \oplus T_2^{a_2} \oplus ... \tag{9}$$

then

$$g(Z) = \underline{a} - \underline{b}. \tag{10}$$

Furthermore,

$$y_k = \prod_{i=1}^{n} x_i^{|Q(i,k)|-|Q(k,i)|} \tag{11}$$

where $|Q(i,k)|$ is the number of edges from i to k.

Ideally $C^T(Z)$ is contained in the upper cluster algebra associated to Q. The indecomposable rigid objects (those satisfying $\mathrm{Hom}(Z, \Sigma Z) = 0$) give cluster variables.

Quiver Grassmannians are defined as follows. If M is a module over a quiver algebra $\mathbb{C}[Q]$ and $\underline{\ell}$ is a dimension vector, then $\mathrm{Gr}_{\underline{\ell}}^{Q}(M)$ is a projective variety parameterizing submodules of M with dimension vector $\underline{\ell}$. Any projective variety appears as some quiver Grassmannian, so they can be come arbitrarily complicated.

We can recognize the images of rigid objects in $\mathrm{End}(T)^{op}$-Mod as follows. For M a module, take a minimal projective presentation

$$P_n \xrightarrow{\pi} P_0 \to M \to 0 \tag{12}$$

and consider the cokernel

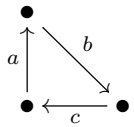
$$\mathrm{Hom}(P_0, M) \xrightarrow{\mathrm{Hom}(\pi, M)} \mathrm{Hom}(P_n, M) \to \mathcal{E}(M). \tag{13}$$

Then $M = F(Z)$ with Z rigid if and only if $\mathcal{E}(M) = 0$. In particular, we need $\mathrm{Ext}^1(M, M) = 0$.

To find categories and objects T of them satisfying the above conditions, we can start from a quiver with potential. If Q is a quiver, a potential W is a linear combination of cycles in Q. From this data one can construct a Ginzburg dg-algebra, and from this dg-algebra one can construct the required category and object. This

object T satisfies $\mathrm{End}(T)^{op} = \mathbb{C}[Q]/\langle \partial W \rangle$ where ∂W is the ideal generated by all cyclic derivatives of the potential W. Actually one needs to take a suitable completion of this in general.

Exercise 24.1. *Consider the quiver*



(14)

with potential $W = cba$. In this case the completed and uncompleted algebras are the same and $\langle \partial W \rangle = \{ba, cb, ac\}$.

274 Curves on Surfaces, Lecture 23

Dylan Thurston
Notes by Qiaochu Yuan

Fall 2012

25 More about additive categorification of surface cluster algebras (Christof)

Mutation of quivers with potential proceeds as follows. Consider again the quiver Q given by

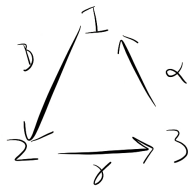


Figure 1: A quiver.

with potential $W = \gamma\beta\alpha$. We want to mutate at 1. The premutation $\tilde{Q}$ has new composite arrows for each composition of arrows through 1, and in addition we change the orientation of arrows starting or ending at 1, giving

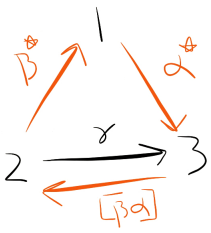


Figure 2: The premutation.

We now need to modify the potential by replacing $\beta\alpha$ with the composite arrow and adding a new term for each composite arrow, giving

$$\tilde{W} = \gamma[\beta\alpha] + [\beta\alpha]\alpha^*\beta^*. \tag{1}$$

The potential now has a 2-cycle, which we will try to eliminate as follows. There is an automorphism

$$\varphi : \gamma \mapsto \gamma + \beta^* \alpha^* \tag{2}$$

of the completed path algebra with the property that if $\tilde{W}' = \gamma[\beta\alpha]$, then $\varphi(\tilde{W}') = \tilde{W}$ up to cyclic rotation of summands. This allows us to remove the 2-cycle $\gamma[\beta\alpha]$, which trivializes the potential.

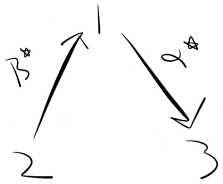


Figure 3: The mutation.

We can associate a quiver with potential to a triangulation of a surface as follows. We restrict our attention to ideal triangulations without tags or self-folded triangles. Details can be found in Labardini-Fragoso.

Let (Σ, M) be a surface with marked points M, a subset $P \subseteq M$ of which are punctures. Fix $x_p \in \mathbb{C}^*$ for each $p \in P$. We associate to a triangulation τ a non-reduced quiver $\tilde{Q}(\tau)$ with vertices given by the edges of the triangulation and edges associated to each interior triangle and each puncture.

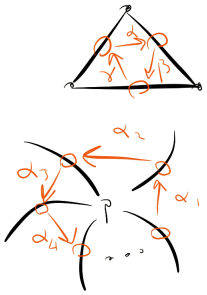


Figure 4: A triangle and a puncture.

We associate to each interior triangle a term $c_\Delta = \gamma\beta\alpha$ and to each puncture a term $c_p = \alpha_n ... \alpha_2 \alpha_1$, and we sum these up to obtain a potential

$$\tilde{W}(\tau) = \sum_{\Delta} c_{\Delta} + \sum_{p} x_p c_p. \tag{3}$$

Finally, we need to remove 2-cycles. The resulting definition of a quiver with potential is compatible with flips of triangulations except in a few cases (Labardini-Fragoso), namely

1. $g(\Sigma) \geq 1$, $\partial\Sigma = \emptyset$ and $|M| \in \{2, 3, 4, 5\}$ and
2. $\Sigma = S^2$ and $|M| \in \{5, 6, 7, 8\}$.

Example Let Σ be a once-punctured 4-gon. The corresponding quiver has a 2-cycle. Before it is removed the potential is $\tilde{W} = \beta\alpha\varepsilon_2 + \delta\gamma\varepsilon_1 + x\varepsilon_2\varepsilon_1$, and after it is removed the potential is $W = \delta\gamma\beta\alpha$.

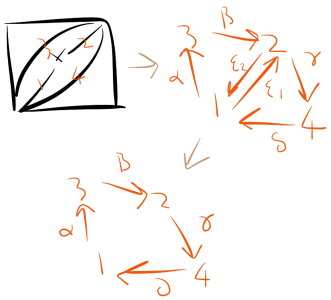


Figure 5: A quiver mutation on a surface.

To categorify surface cluster algebras we now need to associate modules to curves on the surface. We restrict our attention to the case that there are no punctures.

Example Consider a torus with a hole and a marked point on the boundary. The potential is

$$W = \beta^+\beta\beta^- + \alpha^+\alpha\alpha^-. \tag{4}$$

We want to associate modules to curves on this surface. Take, for example, a loop around the hole. We will follow the loop around and associate to each intersection we find a basis vector of the module. Each intersection is connected to the next intersection by some edge in the quiver, so the action of the path algebra is specified by sending intersections to the next intersection in this way.

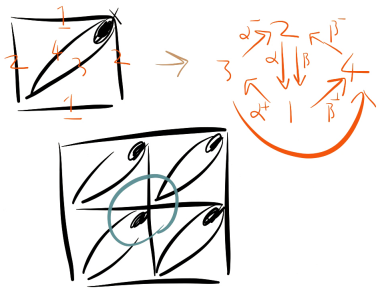


Figure 6: A curve on a surface.

274 Curves on Surfaces, Lecture 24

Dylan Thurston

Fall 2012

26 Monoidal categorification of skein algebras

Let C be an abelian category (an Ab-enriched category with certain nice properties). Its Grothendieck group $K_0(C)$ is the quotient of the free abelian group on isomorphism classes of objects in C by a relation $[B] = [A] + [C]$ for every short exact sequence

$$0 \to A \to B \to C \to 0 \tag{1}$$

in C. If moreover C is monoidal (with monoidal structure suitably compatible with the abelian structure) then $K_0(C)$ inherits a multiplication $[A][B] = [A \otimes B]$ and hence becomes a ring.

If furthermore C has simple objects and each object of C has a composition series of finite length, then $K_0(C)$ is freely generated by the simple objects, and if $X = X_i \supset X_{i-1} \supset ... \supset X_0 = 0$ is a composition series then

$$[X] = \sum_i [X_i/X_{i-1}]. \tag{2}$$

For example, $C = R$-Mod for R a commutative ring satisfies the above. We can also consider representations of groups. More generally we can take $C = H$-Mod where H is a Hopf algebra.

Under the above assumptions, $K_0(C)$ has a strongly positive basis given by simple objects. The structure constants are given by finding the composition series of the tensor product of two simple objects.

Example Consider the representation theory of $\mathrm{SL}_2(\mathbb{R})$. There is an obvious simple module $V = \mathbb{R}^2$. Its symmetric powers $V_k = \mathrm{Sym}^k(V)$ are also simple and can be identified with homogeneous polynomials of degree k in two variables. The tensor products are

$$V_k \otimes V_\ell \cong V_{k+\ell} \oplus V_{k+\ell-2} \oplus ... \oplus V_{|k-\ell|}. \tag{3}$$

Whenever we find an algebra with a positive integral basis, we can conjecture that that algebra has a monoidal categorification in the sense that we can identify it with $K_0(C)$ for some C such that the simple objects are sent to the integral basis. In particular we conjecture that skein algebras $\mathrm{Sk}(\Sigma)$ have this property with the simple objects sent to the bracelets basis.

In the monoidal categorification, product in $\mathrm{Sk}(\Sigma)$ corresponds to tensor product and unions of simple arcs correspond to simple objects. By contrast, in the additive categorification, product in $\mathrm{Sk}(\Sigma)$ corresponds to direct sum and simple arcs correspond to objects satisfying $\mathrm{Ext}^1(X, X) = 0$.

This conjecture is known for the surfaces whose quivers have type A_n and D_n but open in general. One reason to expect it to be true comes from the study of topological quantum field theories. A topological quantum field theory is a symmetric monoidal functor from a cobordism category to a category of vector spaces. Explicitly, for fixed n, it should assign vector spaces $Z(M)$ to $(n-1)$-manifolds M and, to an n-manifold with boundary $\bar{M} \sqcup N$, it should assign a linear map $Z(M) \to Z(N)$ in a way which is compatible with composition and tensor product.

For example, when $n = 2$ a topological quantum field theory is precisely a vector space $Z(S^1)$ together with a commutative product (coming from the pair of pants) and a linear functional (coming from half of a torus) such that

$$\langle ab, c\rangle = \langle a, bc\rangle \tag{4}$$

and which is nondegenerate, and this is precisely a (commutative) Frobenius algebra.

Figure 1: Structure maps of a Frobenius algebra.

When $n = 3$ there are a rich class of examples coming from quantum groups. Associated to any semisimple Lie group G is a quantum group $U_q(\mathfrak{g})$ depending on a prameter q, and setting q to a root of unity we can obtain Witten-Reshetikhin-Turaev TQFTs.

Rather than construct functors out of cobordism categories we can construct functors out of the category of framed tangles.

One such functor is essentially the Jones polynomial. That is, it is obtained by taking free linear combinations and quotienting by the skein relations, which gives an intermediate category called the Temperley-Lieb category, and then applying $\mathrm{Hom}(\emptyset, -)$. The Temperley-Lieb category is spanned by tangles with no crossings. This can be used to construct an $n = 3$ TQFT.

When $n = 4$ we have an interesting near-TQFT (modulo conjectures): Donaldson theory is dual to Seiberg-Witten theory is equivalent to embedded contact homology

Figure 2: A tangle.

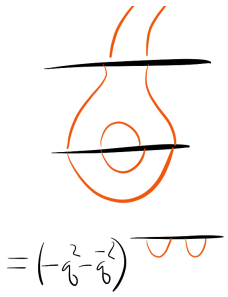


Figure 3: Composition in the Temperley-Lieb category.

is equivalent to Hegaard-Floer homology (these are all very difficult results). This is the only known way to distinguish smooth structures on topological 4-manifolds, and it would be nice if it were easier to compute.

One way to compute is to extend downward. Currently these TQFTs associate data to 4-manifolds and 3-manifolds, but if they also associated data to 2-manifolds we could cut up the 3-manifolds further. In fact they associate homologies to 3-manifolds and can be made to associate derived categories to 2-manifolds.

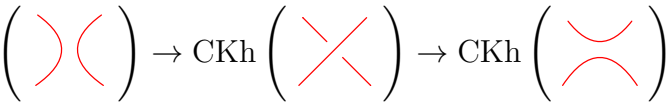
More functorially, there is a 2-category whose objects are $n-2$-manifolds, whose morphisms are $n-1$-dimensional cobordisms between them, and whose 2-morphisms are n-dimensional cobordisms between morphisms (hence in particular are manifolds with corners), and we want a (symmetric monoidal) 2-functor out of this category to some linear 2-category, e.g. a 2-category of linear categories. For example, bordered Heegaard-Floer theory assigns to a surface Σ the derived category of modules over some dg-algebra $A(\Sigma)$, assigns to a 3-manifold a dg-bimodule, and which assigns to a 4-manifold a morphism of bimodules.

Conjecturally cluster algebras can be used to construct a different 4-dimensional

TQFT, and monoidal categorifications might be relevant to doing this. Even more conjecturally we should be able to monoidally categorify $\mathrm{Sk}_q(\Sigma)$ (although a strongly positive basis is not known in general in this case).

This is related to categorifying the Jones polynomial to Khovanov homology. Khovanov homology can be thought of as a functor from the category of framed knots in S^3 and cobordisms to bigraded vector spaces which recovers the Jones polynomial after taking Euler characteristics (alternating sum of graded dimensions). It upgrades the skein relation to a short exact sequence

$$0 \to \mathrm{CKh}\left(\right) \to \mathrm{CKh}\left(\right) \to \mathrm{CKh}\left(\right) \to 0 \tag{5}$$



(with grading shifts) by constructing a chain complex using all possible resolutions of the crossings of a knot.

The 3-manifold TQFTs discussed above come from quantum groups. It is also possible to use quantum groups, together with a representation, to construct a knot polynomial, and Khovanov homology suggests that this story can be categorified. For example, quantum SL_2 and its standard representation gives the Jones polynomial, which is categorified to Khovanov homology. More generally, quantum SL_n and its standard representation gives the HOMFLYPT polynomial, which is categorified to Khovanov-Rozansky or HOMFLYPT homology. Similarly, quantum $\mathrm{SO}(n)$ and its standard representation gives the Kauffman polynomial (not to be confused with the Kauffman bracket), which is categorified to $\mathrm{SO}(n)$ homology (conjecturally).

We also get an invariant of tangles, at least for Khovanov homology, which should give a monoidal categorification for the skein algebra of the disk with n punctures.

274 Curves on Surfaces, Lecture 25

Dylan Thurston

Fall 2012

27 Types of multicurves

References for today: Mirzakhani, Growth of the number of simple closed geodesics on hyperbolic surfaces. Rivin, A simpler proof of Mirzakhani's simple curve asymptotics.

Thurston conjectured the following. Let M be a compact 3-manifold whose boundary is a union of tori. If M is irreducible, atoroidal, and has infinite π_1, then M has a finite cober which fibers over S^1. More generally, we might ask how common it is for a 3-manifold to fiber over S^1.

A 3-manifold has tunnel-number one if $M = H \cup (D^2 \times I)$ where H is an orientable handlebody of genus 2 and the two pieces have been glued along a simple closed curve γ on ∂H. We choose such a thing randomly by choosing Dehn-Thurston coordinates of the corresponding curve on ∂H randomly with size $\leq L$. As $L \to \infty$, it turns out that the probability that M fibers over the circle vanishes as $L \to \infty$.

Alternately, we could fix a set of generators of the mapping class group of ∂H (e.g. some Dehn twists) and randomly apply them to an initial curve γ_0. Conjecturally as $L \to \infty$ the probability that M fibers of the circle still vanishes as $L \to \infty$.

We want the curve γ above to be connected and non-separating. By this we mean the following. Consider multicurves in a surface of genus 2 up to the action of the mapping class group (types of multicurves). A connected such curve either divides the surface into two genus 1 pieces (the separating case) or loops around one of the two holes (the non-separating case), and the general multicurve is a union of such things.

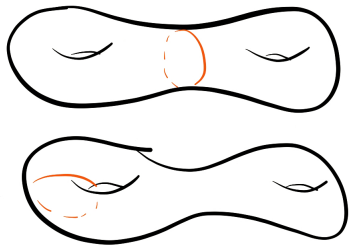


Figure 1: Connected curves on a two-holed torus.

Theorem 27.1. *(Mirzakhani) Fix a multicurve γ. The probability that a random multicurve in Dehn-Thurston coordinates is equivalent to γ under the action of the mapping class group approaches a limit $0 < c_\gamma < 1$ as $L \to \infty$. Furthermore, if $\Omega \subset \mathbb{R}^{6g-6}$ is a bounded region in the space of Dehn coordinates, the proportion of Dehn-Thurston coordinates of random curves that sit inside Ω after rescaling and*

that are equivalent to γ under the action of the mapping class group again, suitably rescaled, again approaches c_γ.

In the case that γ is connected and nonseparating we have $c_\gamma \approx \frac{1}{5}$.

Compare to the case $g = 1$. Then there is only one type of connected curve, and a simple multicurve up to the action of the mapping class group is a finite number of copies of this.

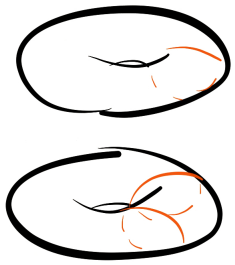


Figure 2: Curves on a torus.

Choosing a random multicurve on the torus means choosing a random pair (p, q) of positive integers, and the number of connected components of the resulting curve is $\gcd(p, q)$. Mirzakhani's result in this case (which is much older) says that there is a definite probability of obtaining $\gcd(p, q) = 1$, which is just $\frac{6}{\pi^2}$.

The appearance of π here is not surprising. Another part of Mirzakhani's result is that c_γ is proportional to the Weil-Petersson volume of the Teichmüller space of $S \setminus \gamma$. (We consider punctures at the boundary components of $S \setminus \gamma$.) The Teichmüller space of $S_{g,n}$ has a canonical symplectic form ω, and ω^{3g-3+n} gives a canonical volume form.

(Edit: there is a result of this form, but the result above is not true as stated.)

To obtain ω, there is another set of coordinates on Teichmüller space called Feichel-Nielsen coordinates obtained by choosing a pants decomposition and looking at lengths ℓ_i of each curve, then looking at the twists t_i around each curve. The symplectic form is then $\omega_{WP} = \sum d\ell_i \wedge dt_i$ (in particular the above does not depend on the choice of pants decomposition).

Alternately, if the surface has a triangulation, then consider shear coordinates s_i for each edge of a triangulation T. Then

$$\omega_{WP} = \sum_{\Delta_{ijk}} (ds_i \wedge ds_j + ds_j \wedge ds_k + ds_k \wedge ds_i) \tag{1}$$

where the sum runs over all triangles and i, j, k are the edges in clockwise order (in particular the above does not depend on the choice of triangulation).

Theorem 27.2. *(Mirzakhani) The Weil-Petersson volume of the Teichmuller space of $S_{g,n}$ is a rational multiple of $\pi^{6g-6+2n}$.*

A key ingredient is that the action of the mapping class group on measured laminations is ergodic with respect to Lebesgue measure. (We say that the action of a group G on a measure space (X, μ) is ergodic if any G-invariant set is either empty or has full measure, and moreover any G-invariant measure that is absolutely continuous with respect to μ is a constant multiple of μ.)

When we compactified Teichmüller space, we tropicalized λ-lengths and obtained bounded measured laminations. Alternately, we tropicalized shear coordinates and obtained unbounded measured laminations. The latter does not give a symplectic manifold, but we can consider the subspace where the sum of the shear coordinates around each puncture is 0 (no spiraling into punctures).

Mirzakhani's result above can be translated into the theory of cluster algebras as follows.

Theorem 27.3. *(Mirzakhani) Fix a surface cluster algebra, not of finite type, with some set of marked points $m_1, ...m_k$. Consider random basis elements x with $deg_{m_i}(x) = 0$. Then the probability that x is of some type (e.g. connected) is definite (strictly between 0 and 1).*

A similar statement should be true for other mutation-finite cluster algebras (neither finite type nor affine). Mutation sequences giving a cluster with the same quiver form a group analogous to the mapping class group, and we can study the orbits of some conjectural positive basis under this group. Conjecturally the orbits are finitely generated in a suitable sense, there is a definite probability of getting any orbit, and the ratios of these probabilities are rational.

What happens in the non-mutation-finite case? What is the analogue of cutting a surface along a simple curve?

274 Curves on Surfaces, Lecture 26

Dylan Thurston

Fall 2012

28 More about geometric interpretations of skein relations

Recall that giving a hyperbolic structure to a surface Σ gives a (discrete, faithful) representation of $\pi_1(\Sigma)$ in $\mathrm{Aut}(\mathbb{H}^2) \cong \mathrm{PSL}_2(\mathbb{R})$ up to conjugacy. If Σ has marked points, then we want cusps in the hyperbolic structure, which gives a collection of ideal points in $\partial\mathbb{H}^2$ (lifts of the cusps) acted on by $\pi_1(\Sigma)$. The number of such orbits should be finite and should satisfy some other conditions.

Decorating cusps gives a collection of horocycles in $\mathbb{H}^2$ acted on by $\pi_1(\Sigma)$. Recall that horocycles can be thought of as (positive) null vectors in the light-cone model or as elements of $\mathbb{R}^2/\{\pm 1\}$.

Now ignore twisting and suppose that we have an $\mathrm{SL}_2(\mathbb{R})$-representation of $\pi_1(\Sigma)$. These can be identified with certain $\mathbb{R}^2$-bundles over Σ with a flat connection (equivalently, an $\mathbb{R}^2$-local system). We want bundles whose transition functions lie in SL_2 (equivalently, SL_2-local systems).

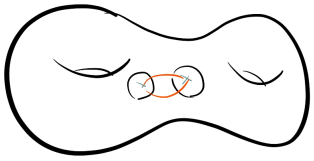


Figure 1: Monodromy of a bundle on a surface.

In this picture, a decorated cusp corresponds to a choice of vector in the fiber of the local system above each marked point.

$\mathrm{PSL}_2(\mathbb{R})$ acts on the unit tangent bundle $\mathrm{UT}(\mathbb{H}^2)$ freely and transitively, so $\mathrm{UT}(\mathbb{H}^2)$ can in fact be identified with $\mathrm{PSL}_2(\mathbb{R})$. Taking double covers gives $\mathrm{SL}_2(\mathbb{R}) \cong \mathrm{UT}^{(2)}(\mathbb{H}^2)$, and taking universal covers gives $\widetilde{\mathrm{SL}_2(\mathbb{R})} \cong \widetilde{\mathrm{UT}(\mathbb{H}^2)}$. (The universal cover of $\mathrm{SL}_2(\mathbb{R})$ is a good example of a group that has no faithful finite-dimensional representations.)

As we saw earlier, this story descends to Σ and gives both a canonical $\mathbb{Z}/2\mathbb{Z}$-extension and a canonical $\mathbb{Z}$-extension of $\pi_1(\Sigma)$, and this is how we define twisted representations. Also, as we saw earlier, a hyperbolic structure gives a PSL_2-representation which canonically lifts to a twisted SL_2-representation.

An immersed loop L on Σ gives a loop in $\mathrm{UT}(\Sigma)$. Given a twisted representation $\tilde{\rho}$, we can now extract a number $\mathrm{tr}(\tilde{\rho}(\tilde{L}))$.

Proposition 28.1. *If $\tilde{\rho}$ comes from a hyperbolic structure and L is taut, then $tr(\tilde{\rho}(\tilde{L})) > 2$.*

Proof. If L is taut, then it is regular isotopic to its geodesic representative L_2. Moreover, L_2 lifts to a curve $\tilde{L}_2$ in $\mathrm{UT}(\Sigma)$ which is still geodesic, and it lifts again to a geodesic (but not necessarily closed) curve in $\mathrm{UT}^{(2)}(\mathbb{H}^2) \cong \mathrm{SL}_2(\mathbb{R})$. A geodesic in a Lie group (with respect to a bi-invariant metric) is up to translation of the form $e^{tM}, M \in \mathfrak{sl}_2(\mathbb{R})$, and the number we want is $\mathrm{tr}(e^M)$. We know that e^M is hyperbolic, which means that the absolute value of the trace is greater than 2, and diagonalizing M the conclusion follows. □

Theorem 28.2. *The positive real points of Spec(Sk(Σ)) (the real points on which the bands or bracelets basis evaluate to positive numbers) are naturally identified with Teich(Σ).*

Proof. (Sketch) The interesting case is when Σ is closed. Let $\nu : \mathrm{Sk}(\Sigma) \to \mathbb{R}$ be a positive point. If L is a simple loop, then $\nu(L) > 0$, but we also have $\nu(\mathrm{Brac}^{(k)})(L) > 0$. But

$$\nu(\mathrm{Brac}^{(k)}(L)) = T_k(\nu(L)) \tag{1}$$

where T_k is the k^{th} Chebyshev polynomial. The condition that this is positive for all k implies that $\nu(L) \geq 2$. The complex points of $\mathrm{Spec}(\mathrm{Sk}(\Sigma))$ can be identified with twisted $\mathrm{SL}_2(\mathbb{C})$-representations of $\pi_1(\Sigma)$ (Bullock), and ν itself gives a representation into $\mathrm{PSL}_2(\mathbb{R})$ in which all elements are parabolic or hyperbolic. This is not quite enough to show that ρ is discrete; there is more work needed...

Some indication of why this should be true. The closure of the image of ρ in $\mathrm{PSL}_2(\mathbb{R})$ is a Lie subgroup. Its connected component of the identity is a connected Lie subgroup, hence corresponds to some Lie subalgebra of $\mathfrak{sl}_2(\mathbb{R})$. If the image of ρ consists of hyperbolic elements (and the identity) then the closure of the image cannot be all of $\mathrm{PSL}_2(\mathbb{R})$, so it suffices to rule out the other possible images. □

We now return to the case of marked points (on the boundary for simplicity). We would like to generalize twisted representations to this case. A twisted representation corresponds to an SL_2-local system, not on a surface Σ, but on its unit tangent bundle $\mathrm{UT}(\Sigma)$. We then associate to each marked point a vector in the fiber of the local system above the outward-pointing normal to the marked point.

We can now associate real numbers to arcs A between marked points p and q given a decorated local system as follows. By lifting the arc to $\mathrm{UT}(\Sigma)$ appropriately, we can

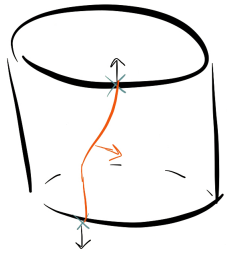


Figure 2: Vectors in the fibers above outward-pointing normals.

get a linear map $\tilde{\rho}(\tilde{A})$ from the fiber of the local system over $\tilde{p}$ (the outward-pointing normal at p) to $\tilde{q}$ (the outward-pointing normal at q). We now choose the real number

$$\tilde{\rho}(\tilde{A})v_p \wedge v_q \tag{2}$$

where v_p is the chosen vector over $\tilde{p}$ and v_q is the chosen vector over $\tilde{q}$. (We have chosen an identification of the determinant bundle with the trivial line bundle.)

Recall that the identity

$$\mathrm{tr}(A)\mathrm{tr}(B) = \mathrm{tr}(AB) + \mathrm{tr}(AB^{-1}) \tag{3}$$

for $A, B \in \mathrm{SL}_2$ gives us the skein relations for loops.

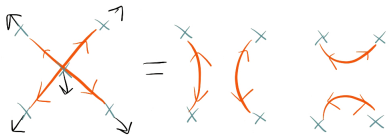


Figure 3: The skein relations again.

The skein relations for arcs can be obtained using the identity

$$Av \wedge w + A^{-1}v \wedge w = \mathrm{tr}(A)\,(v \wedge w) \tag{4}$$

and the Plücker relation

$$(v_1 \wedge v_3)(v_2 \wedge v_4) = (v_1 \wedge v_2)(v_3 \wedge v_4) + (v_1 \wedge v_4)(v_2 \wedge v_3). \tag{5}$$

For example, to prove the skein relation for two crossing arcs, we can translate everything to the fiber over the intersection point.

274 Curves on Surfaces, Lecture 14

Dylan Thurston
Notes by Qiaochu Yuan

Fall 2012

15 Compactifications, skein theory

Theorem 15.1. *The Thurston compactification is a ball for the following Teichmüller spaces:*

1. $\tilde{\mathcal{T}}_{g,n}$ *(λ-lengths),*
2. $\mathcal{T}_{g,n}$ *(cross-ratios),*
3. $\mathcal{T}_{g,0,n}$ *(cross-ratios).*

It is also a ball for $\mathcal{T}_g$ but this is harder (the coordinates are given by lengths of all simple closed curves).

Proof. Correction: when we discussed the Thurston compactification, we should have embedded not in a projective space but in a sphere (in other words instead of quotienting by nonzero reals we quotient by positive reals). If T denotes a triangulation, we can then measure λ-lengths to get a map $\tilde{\mathcal{T}}_{g,n} \to \mathbb{R}^{|T|}$. Embedding $\mathbb{R}^{|T|}$ into $\mathbb{R}^{|T|+1}$ gives a map to $S^{|T|}$.

Lemma 15.2. *The map $\overline{\tilde{\mathcal{T}}_{g,n}} \to S^{|T|}$ above is an embedding.*

Proof. $\tilde{\mathcal{T}}_{g,n}$ embeds into $\mathbb{R}^{|T|}$ due to the existence of the exchange relation: we can express the λ-length of any arc as a Laurent polynomial in the λ-lengths of a fixed triangulation.

To show that the Thurston compactification embeds into $S^{|T|}$, we use the fact that on the new points, our coordinates satisfy a tropical exchange relation, which expresses a new length as a piecewise-linear function of the old lengths. (The key point here is that the original exchange relation is a subtraction-free expression in positive variables. Without this condition, there is no guarantee that a new length is a function of old lengths.) □

The above does not always happen.

Exercise 15.3. *Consider decorated quadrilaterals with three boundary edges of λ-length 1. This moduli space embeds into $\mathbb{R}^2$, and the corresponding Thurston compactification does not embed into S^2.*

We conclude that the Thurston compactification of $\tilde{\mathcal{T}}_{g,n}$ is a hemisphere in $S^{|T|}$, hence homeomorphic to a ball. The argument for $\mathcal{T}_{g,0,n}$ is similar. For $\mathcal{T}_{g,n}$, we have additional relations $\prod \tau(E) = 1$ around punctures, which introduce linear relations after taking logarithms. The corresponding closure in $S^{|T|}$ is still homeomorphic to a ball. □

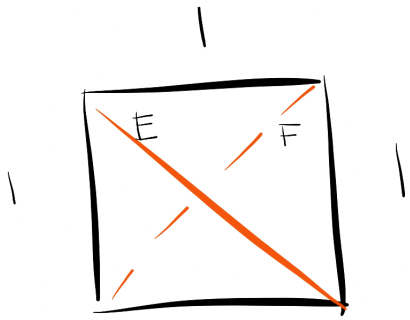


Figure 1: A decorated quadrilateral.

An alternate definition of a lamination, which will hopefully make it more concrete, is the following.

Definition A lamination on $\mathbb{H}^2$ is a collection of pairs of distinct points on the boundary such that the corresponding geodesics do not cross and which is closed as a subset of $(\partial\overline{\mathbb{H}})^2 \setminus \Delta$. A lamination on a hyperbolic surface $\Sigma = \mathbb{H}^2/\Gamma$ is a lamination on $\mathbb{H}^2$ which is invariant under Γ.

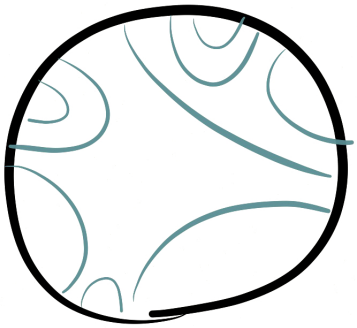


Figure 2: A lamination.

Onto skein theory.

Definition Skein theory for a decorated hyperbolic surface Σ and a multicurve on Σ with endpoints at the punctures is defined up to sign as follows (where ℓ denotes the length of a geodesic representative of a curve, or 0 if there is no such representative):

1. $\lambda(A) = e^{\frac{\ell(A)}{2}}$ for A an arc,
2. $\lambda(C) = 2\cosh\left(\frac{\ell(C)}{2}\right)$ for C a loop,
3. $\lambda(C_1 \cup C_2) = \lambda(C_1)\lambda(C_2)$.

Note that there is no assumption that the multicurve is simple.

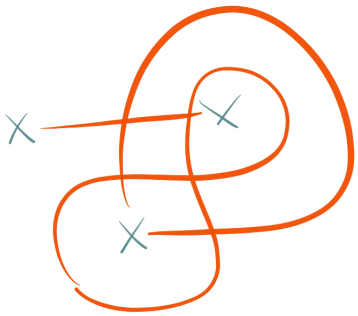


Figure 3: A multicurve, not necessarily simple.

Lemma 15.4. *Up to sign, the skein relation*

$$\lambda\left(\quad\right) = \lambda\left(\quad\right) + \lambda\left(\quad\right) \tag{1}$$

holds.

Some motivation for the above definition is the following.

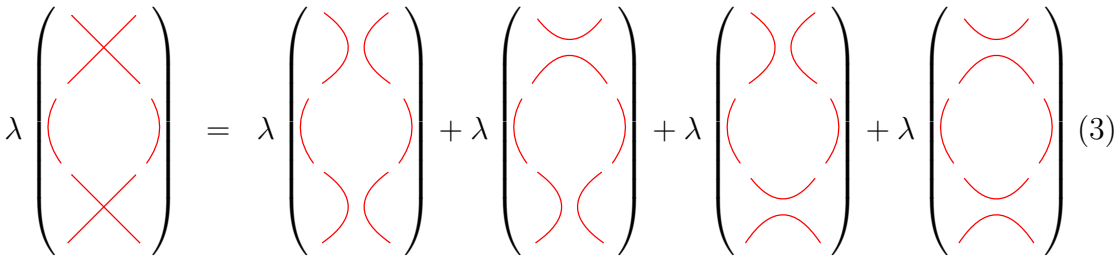
Exercise 15.5. *Let $\phi \in SL_2(\mathbb{R})$ be hyperbolic. Let $\ell(\varphi)$ be the translation length of ϕ (the distance by which ϕ translates on the geodesic between its two fixed points). Then*

$$tr(\phi) = \pm 2 cosh\left(\frac{\ell(\varphi)}{2}\right). \tag{2}$$

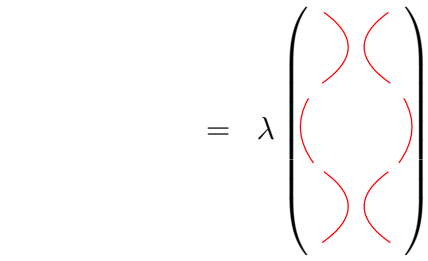
Definition A *curve diagram* on Σ is an immersion $C \to \Sigma = (S, M)$ (M the set of marked points) where C is a 1-manifold with boundary, the boundary of C is sent to marked points, the interior of C misses marked points, and there are no triple intersections. We consider curve diagrams up to isotopies not changing the crossings. The corresponding set is denoted $\mathrm{CD}(\Sigma)$.

Definition The *skein algebra* $\mathrm{Sk}(\Sigma)$ of Σ is $\mathbb{Z}[\mathrm{CD}(\Sigma)]$ modulo the skein relation, the relation that a closed loop has value -2, and the relation that a loop at a marked point has value 0.

Setting a closed loop to -2 is equivalent to invariance under Reidemeister II:

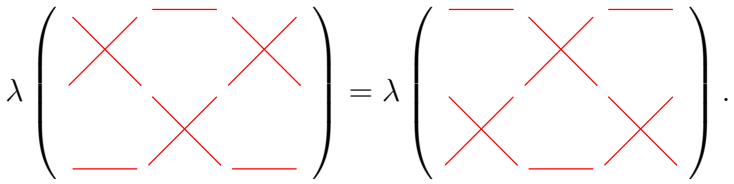


(3)



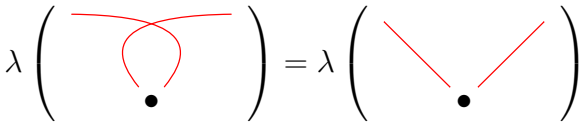
(4)

Exercise 15.6. *Check invariance under Reidemeister III. (Do not expand into* 8 *terms.) That is, check that*



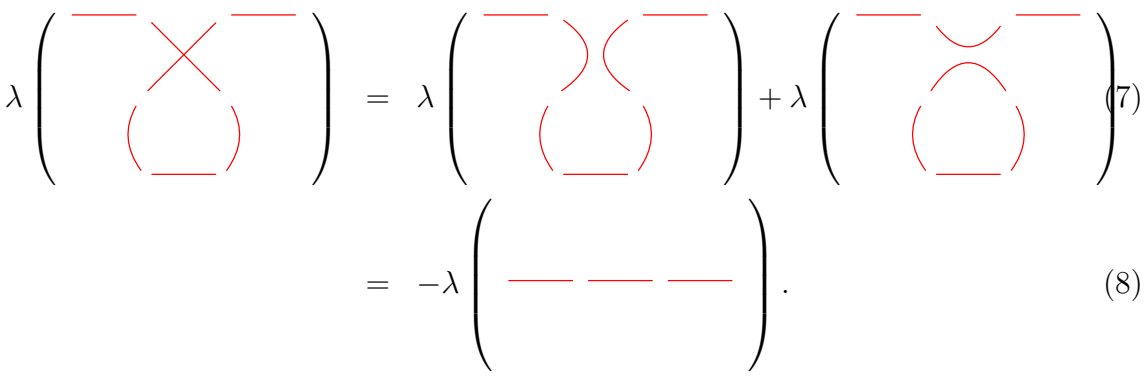
(5)

Setting a boundary loop to 0 is equivalent to invariance under a boundary version of Reidemeister II:



(6)

We do not have invariance under Reidemeister I:



(7)

(8)

Instead, adding a twist introduces a sign. In summary, we have the following.

Proposition 15.7. *The class of a curve in $Sk(\Sigma)$ is invariant under RII, RIIb, RIII, and changes by sign under RI.*

Curves up to RII, RIIb, RIII may be regarded as curves up to regular isotopy.

Theorem 15.8. *Simple multicurves (no loops, no boundary loops) form an integral basis for $Sk(\Sigma)$.*

Proof. There is an expansion map $\mathbb{Z}[\mathrm{CD}(\Sigma)] \to \mathbb{Z}[\mathrm{SC}(\Sigma)]$ (SC the set of simple multicurves) which resolves all crossings using the skein relation and removes all loops and boundary loops. There is also an inclusion map in the other direction, since simple multicurves are curve diagrams. There is also a map $\mathbb{Z}[\mathrm{CD}(\Sigma)] \to \mathrm{Sk}(\Sigma)$ through which the expansion map factors, as well as a map $\mathbb{Z}[\mathrm{SC}(\Sigma)] \to \mathrm{Sk}(\Sigma)$. These maps are inverses. □

There is a map from $\mathrm{Sk}(\Sigma)$ to the upper cluster algebra $\bar{A}(\Sigma)$ associated to Σ. It is not known whether this map is an injection in general. This has been proven under stronger assumptions on the coefficients.

Question from the audience: what does this map look like?

Answer: first we need the following.

Theorem 15.9. *$Sk(\Sigma)$ has a product structure given by union of diagrams.*

Proof. (Sketch) Let $C(\Sigma)$ denote curve diagrams up to regular isotopy (curve diagrams modulo RII, RIIb, RIII). There is a natural map $C(\Sigma) \times C(\Sigma) \to C(\Sigma)$ given by union of diagrams, and this descends to a product structure on $\mathrm{Sk}(\Sigma)$. □

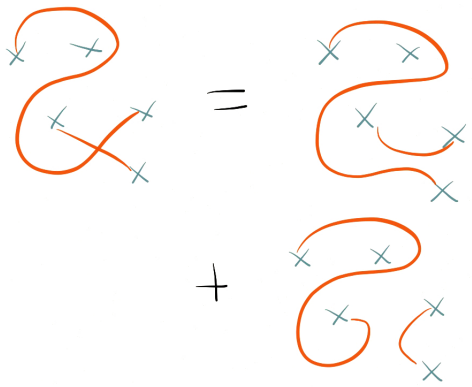


Figure 4: Multiplying by an edge in the triangulation.

Now consider the ring of Laurent polynomials with integer coefficients in the λ-lengths associated to some triangulation T. Given the class of a simple multicurve in $\mathrm{Sk}(\Sigma)$, we can multiply it by an edge in the triangulation and expand using the skein relation.

Repeatedly doing this eventually gives edges in the triangulation, which can be assigned their λ-lengths, and then we can divide by the λ-lengths of the edges we multiplied by.

Exercise 15.10. *Expand a loop in an annulus as a Laurent polynomial.*

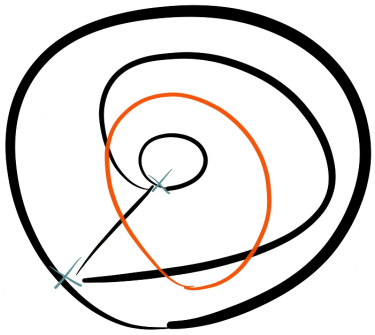


Figure 5: A loop in an annulus.

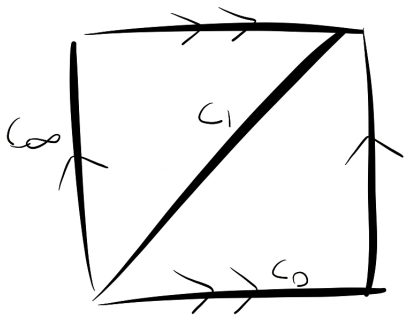


Figure 6: Three curves on a torus.

Exercise 15.11. *Find a relation between the three curves c_0, c_1, c_∞ in $Sk(T^2)$.*

Exercise 15.12. *Try to find skein relations for tagged arcs.*