This article was published in: Topology and its Applications, 73 (2) (1996) pp. 141-167

Surface Diffeomorphisms via Train-tracks*

Hessam Hamidi-Tehrani   Zong-He Chen

October 10, 2001

Abstract

In this paper we describe an algorithm to classify a given surface diffeomorphism in the sense of Thurston. We show that the algorithm can be put into a computer with exponential time in terms of the length of the word representing the diffeomorphism. The algorithm is based on the piecewise linear action of the mapping class group on the space of measured train-tracks (measured geodesic laminations), which was first noticed by Thurston. We use a new coordinate system for the space of measured train-tracks, which is based on the action of the mapping class group on the fundamental group of the surface, first used by Nielsen. We first discuss the algorithm for a punctured surface, and then describe the necessary modifications in the case of a closed surface. We will also indicate the relation with the Birman-Series Conditions for linear action of a diffeomorphism on a set of simple closed curves. We also present some open questions.

Key words: Mapping class group, Measured train-track, π_1-train-track, Dehn twist, Pseudo-Anosov, Birman-Series conditions.

AMS subject classification: 57S05, 57N05

§0. Introduction

In what follows we describe an algorithm to answer this question: Given a diffeomorphism f of a closed or punctured surface, is f reducible, pseudo-Anosov, or of finite order, up to isotopy? (See §3.) This algorithm also finds all the isotopy classes of simple closed curves fixed by f, if f is reducible. If f is pseudo-Anosov, it will find

*Partial support by NSF Grant DMS-94-02988 is acknowledged.

the invariant train-tracks (measured foliations) and the stretching factor. If f is of finite order, it will find the order.

There are other algorithms to carry out Thurston classification for surface diffeomorphisms. There is an algorithm described in [L] which decides whether a map of a punctured disc is pseudo-Anosov or not. A more general algorithm is given in [B-H], which also gives an algorithmic proof of the Thurston's Theorem, using Perron-Ferobenius theory of non-negative matrices.

Our main technique is using the action of the space of diffeomorphisms (resp. the mapping class group) on the space of measured train-tracks on the surface. The new coordinates we use for the space of measured train-tracks is based upon the action of the mapping class group on the fundamental group of the surface, an approach first used by Nielsen and then put in modern setting by [B-S]. Our Proposition 2.1 first appeared in [T]. Also, Mosher's Thesis [M1] is relevant.

We will be considering once punctured surfaces (which have free fundamental groups). The same techniques can be applied for surfaces with more punctures. We will show how to deal with a closed surface in §6.

To fix the notation, S will always be an oriented once punctured surface of genus $g \geq 2$, unless otherwise indicated. All the diffeomorphisms are assumed to be orientation preserving. We always fix a polygon R as a fundamental domain in the universal covering $\mathbf{H}^2$ of S and work within R. Notice that since S is once punctured the vertices of R lie in $\partial\mathbf{H}^2$, and all get identified to the puncture, under the projection (Figure 1). Let $MT(S)$ be the space of measured train-tracks on S, modulo isotopy and a set of moves, called splits, collapses and shifts. For definition of these moves see §1. For more details see [P-H]. The mapping class group $M(S)$ of S acts on $MT(S)$. If S is of genus 1, the action is piecewise linear, and this action is described in [B-S]. We want to investigate this action in the case of higher genus, and as we'll see in §1, it's a piecewise-linear action. We will be able to compute with this action. This is given in §3. In §4 we discuss how to work out our algorithm for understanding the type of a diffeomorphism of a surface, in the sense of Thurston.

Acknowledgments. We hereby thank Professor Joan Birman for being so kind as to introducing us to this subject, reading this manuscript more than once, and suggesting many useful points. Without her help this paper would never have existed. We also thank the referee for the useful comments and corrections.

§1. Basic definitions.

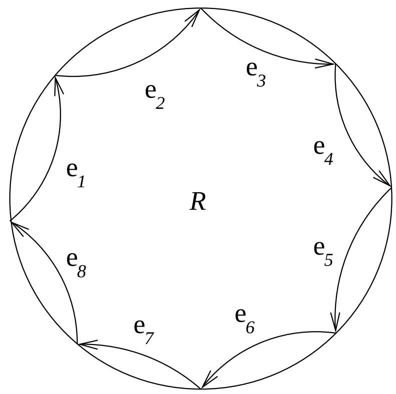


Figure 1.

1.1 Definition. (Train-track) (See [P-H].) Let S be a surface. A compact, connected subset τ of S is called a train-track if τ is a smooth branched 1-manifold embedded smoothly in S. At each branch point v (also called switch points) there is a well-defined tangent space. Every connected component of $\tau - \{\text{branch points}\}$ is called a branch. There is a natural partition for the branches b coming to a switch v (i.e., $v \in \bar{b}$) depending on which direction they become tangent at the switch point. We call these two sets *incoming* and *outgoing*. The particular choice does not matter.

1.2 Definition. (Measured train-track) A measured train-track (τ, μ) consists of a train-track τ, and an assignment of a non-negative number $\mu(b)$ for each branch b of τ, so that the following condition holds: For any switch v of τ,

$$\sum\{\mu(b)|b \text{ an incoming branch to } v\} = \sum\{\mu(b)|b \text{ an outgoing branch to } v\}.$$

The above condition is called the switch condition. We also use the term switch condition for a particular switch v.

1.3 Definition. (π_1-train-track) Suppose S is a once-punctured surface of genus $g \geq 1$. Fix a polygon R in the hyperbolic plane $\mathbf{H}^2$ as a fundamental domain for the action of $\pi_1(S)$ on $\mathbf{H}^2$. Notice that R is naturally identified with S cut open along a number of arcs. Let τ be a train-track in S. We call τ a π_1-train-track (with respect to the choice of R) if the following conditions hold: If we look at τ in the cut-open surface R, there is at most one switch point on each edge of R, no switch points in the interior of R, and all the branches are properly embedded in R, joining distinct vertices in R. (not necessarily distinct in S.)

Figure 2.

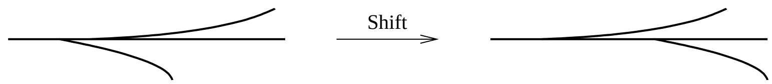


Figure 3.

1.4 We denote by $MT(S)$ the space of all measured train-tracks on a surface S, modulo an equivalence relation which is generated by the following three moves:

(i) Isotopy.

(ii) Right or left split.(Figure 2.)

(iii) Shift.(Figure 3.)

We have only shown the relevant piece of the train-track in Figures 2,3. Notice that the converse of a split is called a *collapse*.

§2. The Piecewise Linear Action

There is a rather nice parameterization of $MT(S)$, given in [B-S], described as follows: Any measured train-track τ can be put in a unique normal form, called π_1-train-track. We will prove this in §5. But let's explain here what it says about train-tracks. A π_1-train-track has at most one vertex on each edge of R, and all branches join vertices on distinct edges of R. So $MT(S)$ can be interpreted as the space of measured π_1-train-tracks. For any π_1-train-track τ, the space of (positive) measures on τ, denoted $V(\tau)$, is a (possibly trivial) set in Euclidean space defined by a set of linear inequalities and equalities. Also, $MT(S)$ can be considered as the cone over the Thurston boundary for the Teichmüller space T_S. This interpretation in fact gives a piecewise integral linear chart for ∂T_S.

Unfortunately when a diffeomorphism acts on a π_1-train-track it need not map it to a π_1-train-track. But there is some algebraic linearity, if we consider the action of a diffeomorphism f on $MT(S)$. See [B-S]. Let's explain what we mean by linearity. Every non-trivial (i.e., non-homotopic to a point or non-parallel to the puncture) isotopy class of a simple closed curve can be represented by an *integral measure* on a π_1-train-track. By an integral measure we just mean assignment of a non-negative integer to each branch of the train-track, so that they satisfy the switch conditions

Figure 4.

[P-H]. In the space of positive measures (or just measures) on a train-track τ, denoted $V(\tau)$, we can add, or multiply by non-negative scalars. Using this structure, for a positive integer n and isotopy classes of multiple simple closed curves C_1 and C_2 carried on some train-track τ, we can define multiple simple closed curves $C_1 + C_2$ and nC_1, and they'll be carried on τ. Let V be a subset of $V(\tau)$. We call V a cone if V is closed under positive linear combinations. Say f acts linearly on a cone V if $f(\lambda_1\mu_1 + \lambda_2\mu_2) = \lambda_1 f(\mu_1) + \lambda_2 f(\mu_2)$, for any non-negative real numbers λ_1, λ_2 and any $\mu_1, \mu_2 \in V$. Say f acts linearly on a set of simple closed curves $C = \{C_1, ..., C_n\}$ if f is linear on the *cone generated* by C(denoted $Span^+(C)$).

In [B-S], they give sufficient conditions for f to act linearly on $Span^+(C)$, for a C as above. For that they look at geodesics representing the curves in C in the universal cover. The mapping $\widetilde{f}$ (a lift of f to the universal cover) must have a kind of orientation-compatibility in mapping these geodesics, to guarantee that f will act linearly on C. For the precise statement see §7 and [B-S].

In Proposition 2.1 is is proved that $MT(S)$ can be 'partitioned' into finitely many cones so that f acts linearly on each of them.

2.1 Proposition.(Thurston,...) The action of $M(S)$ on $MT(S)$ is piecewise integral linear, i.e., any diffeomorphism f induces a piecewise integral linear map $f^* : MT(S) \to MT(S)$.

The proof we give is as follows: We use the fact that $M(S)$ is finitely generated. In fact we use a basic set of curves, consisting of $2g+1$ curves $\{b_2, b_4, ..., b_{2g}, c_1, ..., c_{g-1}, b_1, b_{2g+1}\}$ (Figure 4). It is known that Dehn twists along these curves generate the whole mapping class group. [H].

It's enough to show that each Dehn twist corresponding to one of b_j or c_j and its inverse acts as a piecewise linear map on $MT(S)$. So let f be such a map, and suppose τ is a π_1-train-track on S. Let $V(\tau)$ be the space of measures on τ. Then the map $f^* : V(\tau) \to V(f(\tau))$ is a linear isomorphism. (For each branch b with measure $\mu(b)$, the map f^* assigns the same measure $\mu(b)$ on the image $f(b)$.) Now the idea is as follows. By Theorem 4.1, there is a sequence of splits and isotopies which change $f(\tau)$ to a train-track which can be collapsed to a π_1-train-track. Suppose we perform a single split move on $f(\tau)$. Then we claim we can write $V(\tau) = V_1(\tau) \cup V_2(\tau)$, each $V_i(\tau)$ being closed under positive linear combinations (i.e., they are subcones of $MT(S)$), such that the restriction of f^* to each $V_i(\tau)$ is linear. That is, the image of $f^*|_{V_i(\tau)}$ is contained in a chart, such that in this chart, $f^*|_{V_i(\tau)}$ is linear. To prove this we just note that each $V_i(\tau)$ is defined by a linear inequality in $V(\tau)$ making sure all the measures on the branches are non-negative, and depending on the fact that we make right or left split, we get different linear integral maps. Arguing in the same way we get that f^* is a piecewise linear map. ♠

In §4 we are going to investigate these maps in much more detail, in fact we're able to write down all the linear maps together with a finite decomposition

$$V(\tau) = \bigcup_{i=1}^{n} V_i$$

into subcones V_i each defined by a finite set of linear inequalities, so that $f^*|_{V_i} : V_i \to MT(S)$ is linear. Putting these all together for all π_1-train-tracks, gives the piecewise linear homeomorphism $f^* : MT(S) \to MT(S)$.

§3. The Algorithm

In what follows we discuss an algorithm for how to find the subcones. In particular we will be able to find exactly the subsets of $MT(S)$ on which f^* acts linearly (see §2).

3.1 Some notation.

(i) Let n be the number of edges in the polygon R, and call the edges $e_1, e_2, ..., e_n$ in clockwise order. Give each e_i the orientation induced by the clockwise orientation on ∂R. If e_i is identified with e_j in S (obviously with the opposite orientation, since S is orientable), we denote that by $e_i = e_j^{-1}$.

Notice that the notation $e_i = e_j^{-1}$ is logical, because e_i can be used to denote the covering transformation which pairs e_j with e_i. It's easily seen that now we have $e_i = e_j^{-1}$, as covering transformations. In §5 we will look at elements of $\pi_1(S)$ as covering transformations, and the set of $e_j^{\pm 1}$ is used as a set of generators for $\pi_1(S)$.

(ii) For each edge $e = e_i$ of R fix two disjoint closed subintervals e^+ and e^- in $int(e)$ so that e^- comes before e^+ in the orientation induced by the clockwise orientation of ∂R. We assume all the π_1-train-tracks τ we consider have the following property: If a vertex v of τ lies on an edge e then v is between e^- and e^+ and $v \notin e^- \cup e^+$. Also, clearly we assume $e_i^+ = (e_j^-)^{-1}$ and $e_i^- = (e_j^+)^{-1}$, so that the intervals match.

(iii) A simple closed curve C can be given by a cyclic word $e_{\alpha_1}^{\epsilon_1}...e_{\alpha_k}^{\epsilon_k}$ where $1 \leq \alpha_i \leq n$ and $\epsilon_i = \pm$ according as C crosses the edge e_{α_i} in the subinterval $e_{\alpha_i}^{\pm}$. To draw the curve in R from the given word, just start on the interval $e_{\alpha_1}^{\epsilon_1}$ and connect it to $(e_{\alpha_2}^{-1})^{-\epsilon_2}$, so that it'll come out of $e_{\alpha_2}^{\epsilon_2}$, etc. All the curves that we consider are assumed to be tight, i.e., $\alpha_i \neq \alpha_{i+1}$ for all i (Consider i to be a cyclic index). Although changing $+$ or $-$ signs doesn't change the isotopy class of C, we'll need the ϵ_i's to keep track of the intersection points of C with a given π_1-train-track.

Now assume that a π_1-train-track τ is given. We would like to represent τ as a matrix. Let $A(\tau)$ be an $n \times n$ matrix where there is an indeterminate variable a_{ij} on position (i,j) if there is a branch of τ from e_i to e_j. If there is no branch from e_i to e_j then the (i,j) position of $A(\tau)$ is zero (see Figure 16 and Example 3.9). Also, we make the following natural assumptions:
1. $A(\tau)$ is symmetric.
2. All the equations corresponding to the switch conditions hold.

3.2 Note. In fact, $MT(S)$ can be identified with the subset of all $n \times n$ matrices $A = [a_{ij}]$ with non-negative entries defined by the following conditions: (i) A is symmetric. Its diagonal entries are zero.
(ii) If $a_{ij} > 0$, with $j < i$, then $a_{i'j'} = 0$ if $j' < j < i' < i$ or $j < j' < i < i'$.
(This corresponds to the fact that the branches of τ do not cross each other.)
(iii) If $e_i = e_j^{-1}$ then $\sum_{k=1}^{n} a_{ki} = \sum_{k=1}^{n} a_{kj}$. (Switch conditions) If we projectivize this set by adding the condition $\sum_{1 \leq i,j \leq n} a_{ij} = 1$, then $PMT(S)$ (the space of projectivized measured train-tracks) will be identified with a polyhedron in $\mathbf{R}^{n^2}$, and therefore gives a PL-atlas for ∂T_S, the Thurston boundary of the Teichmüller space. Notice that f has a piecewise linear action on this polyhedron.

3.3 Remark. In practice when we are working with $A(\tau)$ we always solve the switch equations in terms of a bunch of independent variables and replace each entry of $A(\tau)$ in terms of those. That'll make it easier to work with the linear maps coming up in §4.

Now suppose we are given a π_1-train-track τ as a matrix $A(\tau)$, and a simple closed

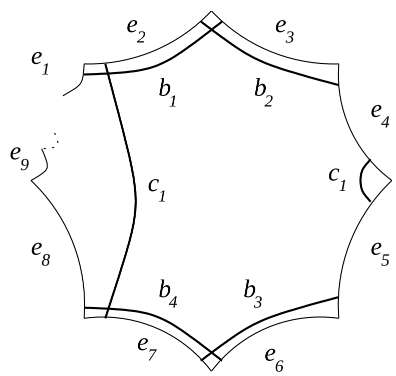


Figure 5.

curve C as a cyclic word $e_{\alpha_1}^{\epsilon_1}...e_{\alpha_k}^{\epsilon_k}$. Let $f = D_C$. We want to be able to investigate the map $f^* : V(\tau) \to MT(S)$. So we have to describe $f(\tau)$, which can be represented by different π_1-train-tracks, depending on the measure on it. We have to consider all possibilities. For example, when we do a split on $f(\tau)$ we have to do both right and left splits (the trivial split is a special case of both), and correspondingly we get a partition of $V(\tau)$ into 2 subcones, simply by the inequalities that should hold so that we're able to do certain splits. We will only deal with the Dehn twists along curves in a basic set of generators (Figure 4). If we open the surface up we get a $4g$-gon R with edges $e_1, ..., e_{4g}$ where $e_{4k-2} = e_{4k}^{-1}$ and $e_{4k-3} = e_{4k-1}^{-1}$, $k = 1, ..., g$. (Figure 5)

Our approach is to express f as a product of Dehn twists on the curves of Figure 4 and compute the action of each elementary Dehn twist separately.

3.4 The action of a single Dehn twist $f = D_{b_i}$ on a train-track $A(\tau)$, $i = 1, ..., 2g$.

As Figure 5 shows, b_i is opposite to an edge $e_{k(i)}$ where $k(i) = 2i$ for i odd and $k(i) = 2i - 1$ for i even. If we apply f, then $f(\tau)$ is collapsible to a π_1-train-track unless $a_{k(i),k(i)+1} \neq 0$. (Figure 6)

In this case, we have to push the little curve going from $e_{k(i)+1}$ to itself (we call such a curve a *bad curve*), through $e_{k(i)+1}$. It'll come out near $e_{k(i)+1}^{-1} = e_{k(i)-1}$, and we get a local picture like Figure 7 and 8.

near $e_{k(i)-1}$. Let n be the number of branches coming out of $e_{k(i)-1}$, i.e., the number of non-zero entries of the $(k(i)-1)$st row in $A(\tau)$. The claim is that $f^* : V(\tau) \to MT(S)$ can be broken up into n linear maps. In other words, there are exactly n ways to split the train-track in Figure 9, so that the resulting train-track is collapsible to a π_1-train-track.

To prove the claim, we argue as follows: Since a split is equivalent to choosing a

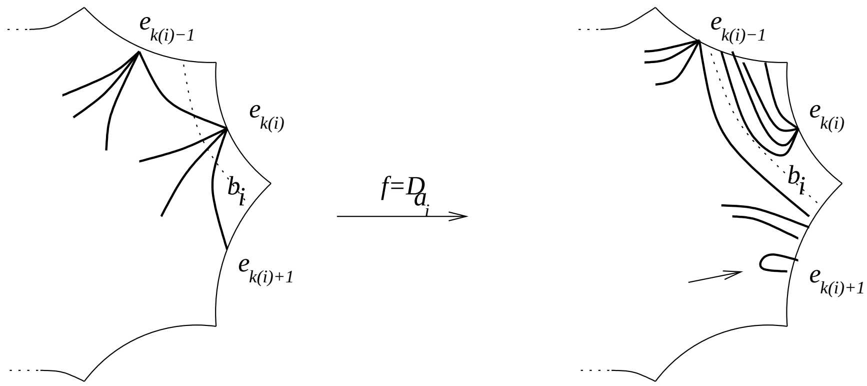


Figure 6.

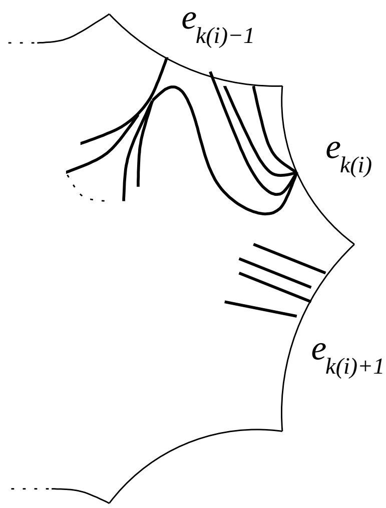


Figure 7.

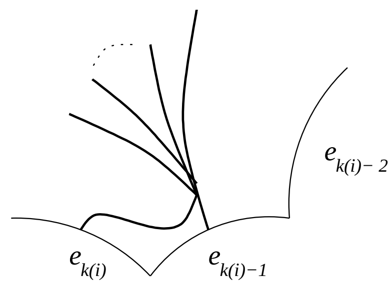


Figure 8. Near $e_{k(i)-1}$.

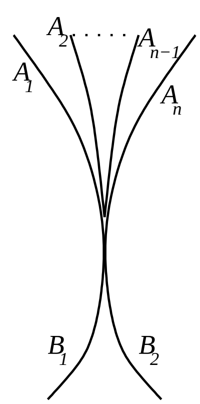


Figure 9.

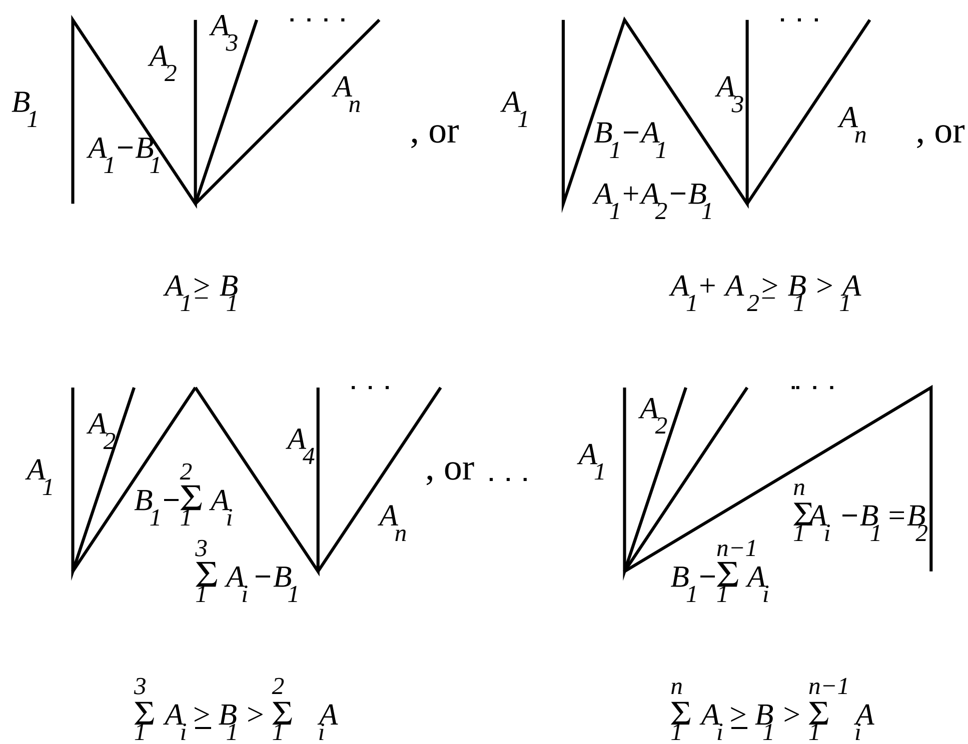


Figure 10.

number $1 \leq j \leq n$ and connecting the first j branches from the top to the first vertex in the bottom, there are exactly n ways of doing the split. In Figure 10 we see a few of these splits and corresponding conditions for measures.

After collapsing to π_1-train-tracks, this gives a partition $V(\tau) = \bigcup_{t=1}^{n} V_t(\tau)$, so that $f^*|_{V_i}$ is linear. We call this the $f-partition$ of $V(\tau)$.

3.5 Corollary. Let $f = D_{b_i}$, $i = 1, ..., 2g$. If μ_1, μ_2 are measured train-tracks carried by the same π_1-train-track τ, then $f(\mu_1 + \mu_2) = f(\mu_1) + f(\mu_2)$ if and only if $\mu_1, \mu_2 \in V_t(\tau)$ for some t, where V_t is one of the sets in the f-partition of $V(\tau)$.

Proof. The 'if' part follows from the previous discussion. To see the 'only if' part, Notice that if μ_1 and μ_2 are not in the same V_t, then by construction we have to use different splits to get $f(\mu_1)$ and $f(\mu_2)$ into the π_1-form. But two different splits as in Figure 8 makes it impossible for $f(\mu_1)$ and $f(\mu_2)$ to be carried by the same train-track. So $f(\mu_1) + f(\mu_2)$ doesn't even give a measured train-track, so it's distinct from $f(\mu_1 + \mu_2)$. ♠

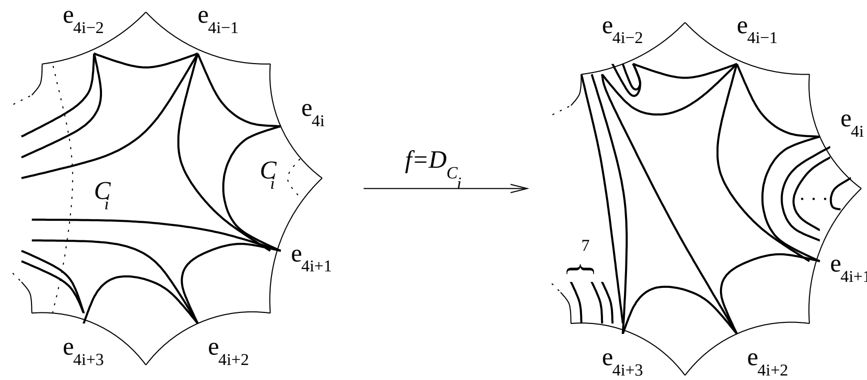


Figure 11.

3.6 Remark. Corollary 3.5 shows that there is an integer $N(g)$ only depending on g so that for a Dehn twist f along a non-separating curve b there is a PL-atlas for $MT(S)$ with $N(g)$ *connected* charts such that f^* acts linearly on two elements of $MT(S)$ if and only if they are in the same chart. It's not hard to compute $N(g)$.

3.7 Question. Can Remark 3.6 be carried out for a Dehn twist along a separating curve, or in general, for any diffeomorphism f? I.e., is there a PL-atlas for $MT(S)$ with connected charts, such that f acts linearly on two elements if and only if they belong to the same chart? If the answer is yes (which we doubt), the least number of charts necessary to do so is a good 'measure of complexity' of the diffeomorphism f.

3.8 The action of $f = D_{c_i}$ on a train-track $A(\tau)$, $i = 1, ..., g-1$.
The curve c_i can be given by the word $e_{4i}^+ e_{4i+3}^+$, this has the advantage that the π_1-train-tracks can only intersect the arc going from e_{4i+3}^+ to e_{4i-2}^-. The action of f on τ looks like Figure 11.

If there are edges from e_{4i+2} to $e_{4i+4}, ..., e_{4g}, e_1, ..., e_{4i+1}$, then there are going to be bad curves(at most $4g-6$). Let's say there are m bad curves. After pushing all of them to $e_{4i-2}^{-1} = e_{4i}$, we have the local picture as in Figure 12.

We consider two cases:

Case 1. $a_{4i,4i+1} = 0$.

In this case we can do the splits in the same way we did before. The only difference is that, if there are n branches coming out of e_{4i+1}, we have to split (Figure 13).

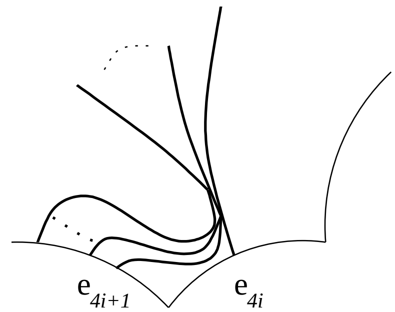


Figure 12.

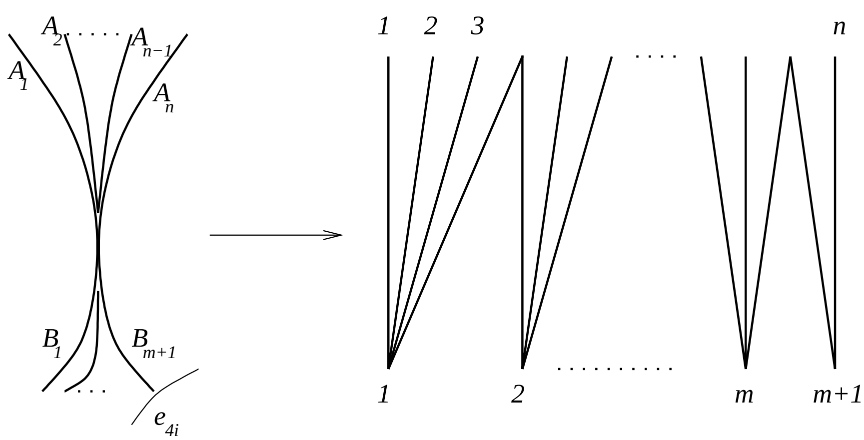


Figure 13.

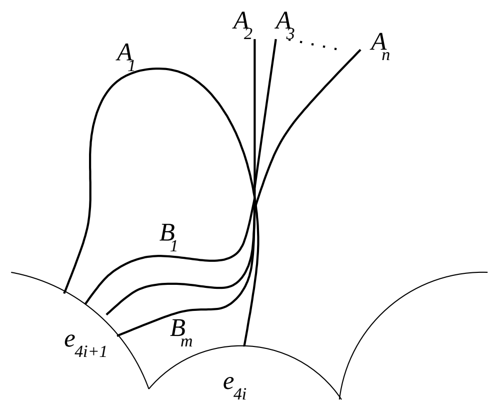


Figure 14.

And there are going to be $\begin{pmatrix} m+n-1 \\ m \end{pmatrix}$ ways, because there are $\begin{pmatrix} m+n-1 \\ m \end{pmatrix}$ ways to split n into sum of $m+1$ ordered non-negative integers, the first one being positive.

Case 2. (The case of a *bad pair*.) $a_{4i,4i+1} \neq 0$.

Here if we split as we did in case 1, the resulting train-track wouldn't be collapsible to a π_1-train-track. The local picture is as in Figure 14.

Let $A, B_1, ..., B_m$ be the measures for the corresponding branches as in Figure 14. Either there is a $1 \leq t \leq m-1$ such that

$$B_1 + ... + B_t \leq A < B_1 + ... + B_{t+1} \quad [subcase\ t]$$

or

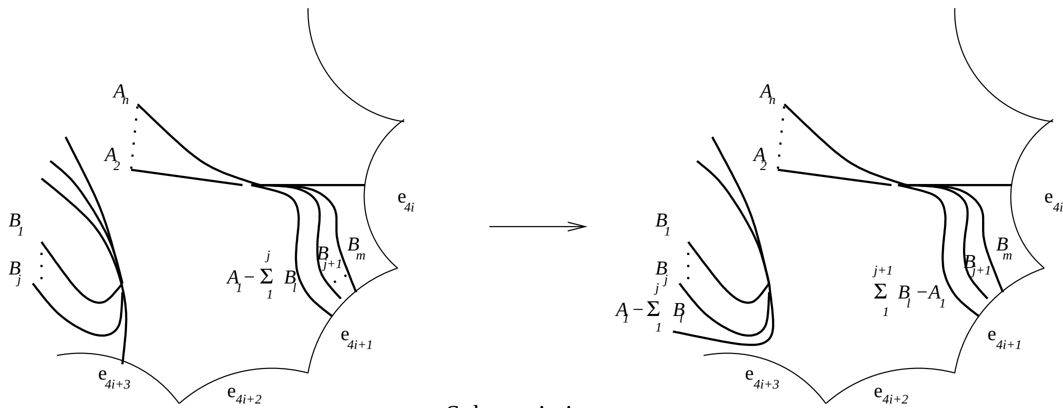
$$B_1 + ... + B_m \leq A \quad [subcase\ m]$$

or

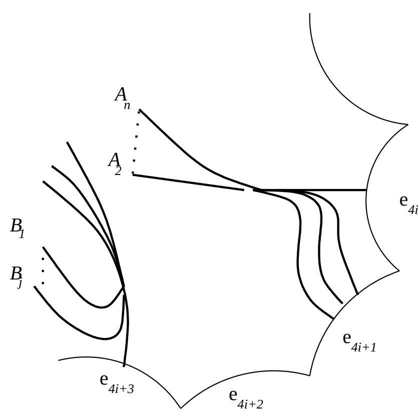
$$A < B_1 \quad [subcase\ 0].$$

In subcase j $(0 \leq j \leq m)$, we can push the branches labeled $B_1, ..., B_j$ along A to $e_{4i+1}^{-1} = e_{4i+3}$. The branch A will be reduced to $A - \sum_1^j B_l$. If $j < m$ then we can push $A - \sum_1^j B_l$ along B_{j+1} also.Then the picture would be like Figure 15.

Now it's easily seen that there are no more bad pairs, and we can split near e_{4i} and e_{4i+3}.



Subcase j , j<m.



Subcase m.

Figure 15.

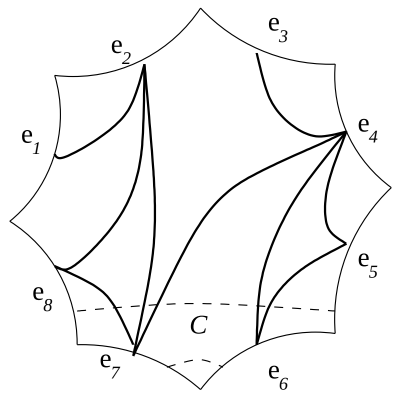


Figure 16.

In this case we are going to get many more linear maps, but this complication only arises because our coordinates have been set up to simplify the action of D_{b_i} the most.

We have written a computer program which computes the action of D_{b_i} and D_{c_j}, $i = 1,...,2g$, $j = 1,...,g-1$, on $MT(S)$. Writing a similar program to compute the action of a Dehn twist along a (presumably complicated) curve in the given coordinates seems hard, and might need arbitrarily large memory, since there is no bound on the number of bad pairs or bad curves which occur after applying the Dehn twist. However, we can extend the program to find the action of 'sufficiently simple' curves in the given coordinates. To show how this can be done, and in particular how the above algorithm works, let's look at an example.

3.9 Example. Suppose S is the once punctured surface of genus 2, and let R be an octagon with edges $e_1,...,e_8$ where $e_1 = e_3^{-1}, e_2 = e_4^{-1}, e_5 = e_7^{-1}$ and $e_6 = e_8^{-1}$. Let τ be the train-track in Figure 16.

which is given by the matrix

$$A(\tau) = [a_{ij}] = \begin{bmatrix} 0 & a_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{12} & 0 & 0 & a_{24} & 0 & 0 & a_{27} & a_{28} \\ 0 & 0 & 0 & a_{34} & 0 & 0 & 0 & 0 \\ 0 & a_{24} & a_{34} & 0 & a_{45} & a_{46} & a_{47} & 0 \\ 0 & 0 & 0 & a_{45} & 0 & a_{56} & 0 & 0 \\ 0 & 0 & 0 & a_{46} & a_{56} & 0 & 0 & 0 \\ 0 & a_{27} & 0 & a_{47} & 0 & 0 & 0 & a_{78} \\ 0 & a_{28} & 0 & 0 & 0 & 0 & a_{78} & 0 \end{bmatrix}$$

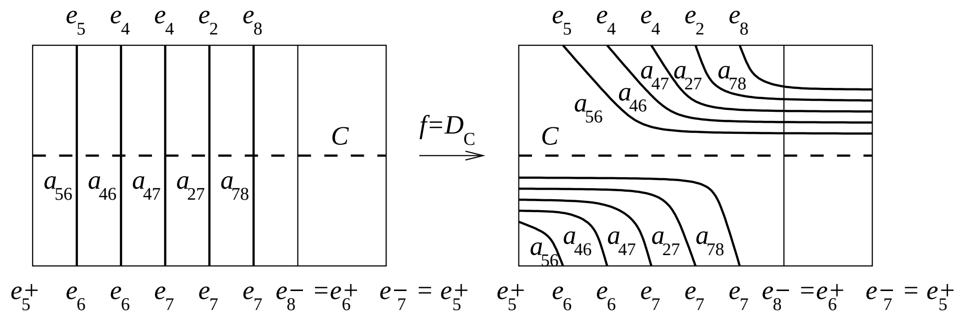


Figure 17.

After solving the switch-condition equations in terms of 6 variables $x_1, \ldots, x_6$ we see that $V(\tau)$ is the space of matrices of the form

$$A(x_1,\ldots,x_6) = [a_{ij}] = \begin{bmatrix} 0 & x_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ x_3 & 0 & 0 & x_1 & 0 & 0 & x_2 & x_4+x_5 \\ 0 & 0 & 0 & x_3 & 0 & 0 & 0 & 0 \\ 0 & x_1 & x_3 & 0 & x_2 & x_4 & x_5 & 0 \\ 0 & 0 & 0 & x_2 & 0 & x_5+x_6 & 0 & 0 \\ 0 & 0 & 0 & x_4 & x_5+x_6 & 0 & 0 & 0 \\ 0 & x_2 & 0 & x_5 & 0 & 0 & 0 & x_6 \\ 0 & x_4+x_5 & 0 & 0 & 0 & 0 & x_6 & 0 \end{bmatrix}$$

Where $x_1, \ldots, x_6 \geq 0$. We want to perform a Dehn twist along the curve C given by the cyclic word $e_5^+ e_6^+$ depicted in the Figure 16. Notice that this curve is not among the elementary curves in Figure 4, but we chose it to show how our methods can be easily worked out for all 'fairly simple' curves, besides the elementary curves in Figure 4. Let's show how to find the image $f(\tau)$. We only have to look at a neighborhood of C because everything else stays the same (Figure 17).

From this we can read off what $f(\tau)$ is going to be. The only thing that makes the situation complicated is that there is a bad curve going from e_8 to e_8^-. To get rid of that, we have to push it by isotopy through e_8 and by looking at Figure 16 we can see that it's going to be an edge from e_7^- to e_6 attaching 'backwards'. In fact, in the computer program, we record all the information given by Figure 17 and work with that. After splitting, we get Figure 18.

Therefore, if we write

$$V_1(\tau) = \{A(x_1,\ldots,x_6) \in V(\tau) : x_6 \geq x_4\} = \{A(\tau) : a_{78} \geq a_{46}\}$$

$$V_2(\tau) = \{A(x_1,\ldots,x_6) \in V(\tau) | x_6 \leq x_4\} = \{A(\tau) : a_{78} \leq a_{46}\}$$

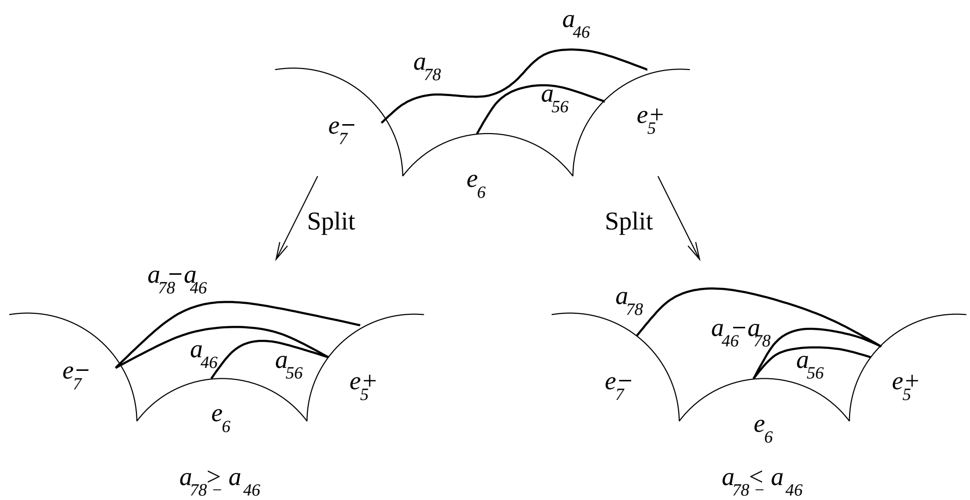


Figure 18.

$$f_1^* = f^*|_{V_1}, f_2^* = f^*|_{V_2},$$

then obviously $V(\tau) = V_1(\tau) \cup V_2(\tau)$ and f_1^*, f_2^* are linear functions given by:

$$f_1^*([a_{ij}]) =$$

$$\begin{bmatrix} 0 & a_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{12} & 0 & 0 & a_{24} & 0 & 0 & 0 & a_{27}+a_{28} \\ 0 & 0 & 0 & a_{34} & 0 & 0 & 0 & 0 \\ 0 & a_{24} & a_{34} & 0 & a_{45} & 0 & 0 & a_{46}+a_{47} \\ 0 & 0 & 0 & a_{45} & 0 & \gamma & \alpha & a_{56} \\ 0 & 0 & 0 & 0 & \gamma & 0 & \beta & 0 \\ 0 & 0 & 0 & 0 & \alpha & \beta & 0 & 0 \\ 0 & a_{27}+a_{28} & 0 & a_{46}+a_{47} & a_{56} & 0 & 0 & 0 \end{bmatrix},$$

where $\alpha = a_{46} + a_{47} + a_{27} + a_{78}, \beta = a_{46} + a_{47} + a_{27} + a_{56}$ and $\gamma = a_{56} + a_{46} - a_{78}$.

$$f_2^*([a_{ij}]) =$$

$$\begin{bmatrix} 0 & a_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{12} & 0 & 0 & a_{24} & 0 & 0 & 0 & a_{27}+a_{28} \\ 0 & 0 & 0 & a_{34} & 0 & 0 & 0 & 0 \\ 0 & a_{24} & a_{34} & 0 & a_{45} & 0 & 0 & a_{46}+a_{47} \\ 0 & 0 & 0 & a_{45} & 0 & \gamma & \alpha' & a_{56} \\ 0 & 0 & 0 & 0 & \gamma & 0 & \beta' & 0 \\ 0 & 0 & 0 & 0 & \alpha' & \beta' & 0 & 0 \\ 0 & a_{27}+a_{28} & 0 & a_{46}+a_{47} & a_{56} & 0 & 0 & 0 \end{bmatrix},$$

where $\alpha' = a_{47} + a_{27} + 2a_{78}, \beta' = a_{46} + a_{47} + a_{27} + a_{56}$. We can write the corresponding equations in terms of x_i, $i = 1, ..., 6$.

3.10 Remark. We can also consider other fundamental domains for the surface S. For example, as in [B-S], we can take a fundamental domain as a $4g$ -gon with identifications $a_1b_1a_2b_2...a_gb_ga_1^{-1}b_1^{-1}...a_g^{-1}b_g^{-1}$. We can also show the curves in the basic set of generators as cyclic words of length 2, but now different train-tracks may intersect both of the arcs in each curve simultaneously. This makes it harder to analyze the action in this setting.

On the other hand, it's been recently proved that $M(S)$ can be generated with 2 elements [W]. Namely, with the notation as in Figure 4, the elements

$$x = D_{b_1}D_{b_2}D_{c_1}D_{b_4}D_{c_2}...D_{c_{g-1}}D_{b_{2g}}$$

and

$$y = D_{b_{2g+1}}^{-1}D_{b_{2g-1}}$$

generate $M(S)$. This would not help in simplifying our algorithm, because we still have to analyze the action of these two diffeomorphisms (instead of two 'kinds' that we analyzed, see 3.4 and 3.8). Also, for a general diffeomorphism f, the length of f as a word in x, y is presumably much longer than its length in the set of generators the we considered, although the situation is not completely clear.

Now let's explain how we can find a decomposition of $MT(S)$ for a given diffeomorphism. As we said before, we can write every diffeomorphism f as a word $\varphi_k...\varphi_1$ in a basic set of generators, up to isotopy. By the above procedure construct the decomposition

$$MT(S) = \bigcup_{j=1}^{m_i} V_i^{(j)}$$

into subcones $V_i^{(j)}$ so that for each $1 \leq j \leq m_i$, φ_i induces a linear isomorphism

$$\varphi_i^{(j)} : V_i^{(j)} \to W_i^{(j)} (:= \varphi_i^{(j)}(V_i^{(j)}))$$

Then it's easily seen that the map f is the linear map $\varphi_k^{(j_k)}...\varphi_1^{(j_1)}$ on the sets $V^{j_k,...,j_1} := (\varphi_k^{(j_k)})^{-1}(W_k^{(j_k)}) \cap ...(\varphi_1^{(j_1)})^{-1}(W_1^{(j_1)})$, and in fact these sets give a decomposition of $MT(S)$ into $m_1...m_k$ subcones such that f acts linearly on each piece.

§4. **Thurston Theory**

Assume $f \in \mathrm{Diff}^+(S)$ is given by a word in $\{f_1^{\pm 1}, ..., f_{2g+1}^{\pm 1}\}$. According to Thurston classification of surface diffeomorphisms, exactly one of the following holds:

1. f is reducible, i.e., $f(C) \simeq C$, where C is a finite collection of disjoint essential simple closed curves on S.
2. f is (isotopic to) a pseudo-Anosov diffeomorphism.
3. f is isotopic to a finite order diffeomorphism.

Here we want to show that, using §3, we can decide which category the given diffeomorphism f belongs to. Furthermore, we will be able to find all simple closed curves that f fixes in case 1, the foliations and the stretching factor in case 2, and the order of f in case 3.

Suppose $f = \varphi_k ... \varphi_1$ where $\varphi_i \in \{f_1^{\pm 1}, ..., f_{2g+1}^{\pm 1}\}$. As we saw, $f^* : MT(S) \to MT(S)$ is piecewise linear. Every φ_i gives rise to at most N linear maps, where N only depends on g. Suppose φ_i acts by a linear map $A_i^{(j)}$ on $V_i^{(j)}$, where $MT(S) = \bigcup_j V_i^{(j)}$. These we can find using the program based on §3. Now look at all linear maps of the form $A_k^{(i_k)} ... A_1^{(i_1)}$. Picking any of those see if it has an eigenvalue 1. Find the space X of eigenvectors in $V_1^{(i_1)}$ for $\lambda = 1$. Check if $A_l^{(i_l)} ... A_1^{(i_1)} X \subset V_l^{(i_l)}$, for $l = 1, ..., k$. Take an appropriate subspace of X so that all these conditions hold. If $X \neq 0$ then you have found a multiple simple closed curve which is fixed componentwise by f. You can write down all the multiple curves fixed by f in this fashion. We have to check that the multiple curves that we find are not simple closed curves parallel to the puncture. Now suppose we didn't find any curve fixed by f in this procedure. Then we start looking at eigenspaces for eigenvalues $\lambda > 1$. In this way we find out if there is any measured train-track μ mapped to $\lambda\mu$ by f^*. If so, that will show that f is pseudo-Anosov and μ obviously gives rise to a measured foliation. Note that this implies λ has to be an eigenvalue of an integral matrix, i.e., an algebraic integer.

We can also find the other foliation μ' with property $f^*\mu' = 1/\lambda\ \mu'$, by looking at eigenspaces of $1/\lambda$.

If we don't find any fixed curve or foliation, then f must be of finite order. To find the order, we need a little lemma.

4.1 Lemma. For a surface S of genus g, there are $6g - 6$ simple closed curves on S such that any diffeomorphism fixing these curves (up to isotopy) is itself isotopic to identity.

Proof. As in [FLP], take a permissible pair of pants decomposition for S, and

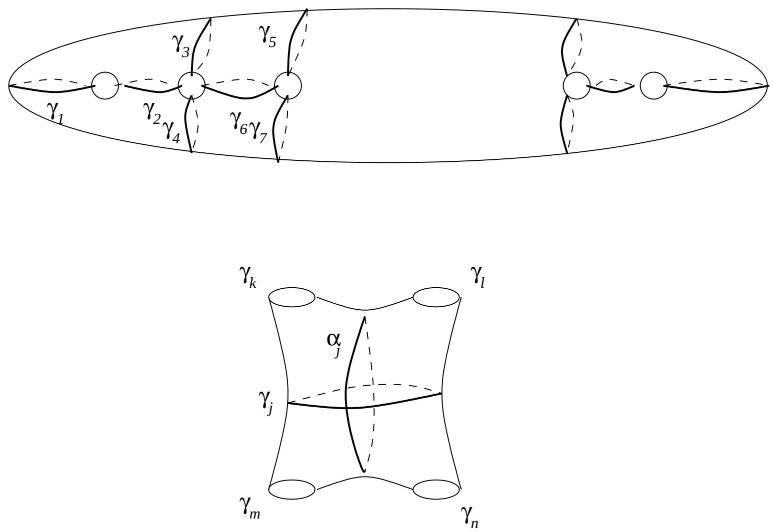


Figure 19.

take all $3g-3$ curves $\gamma_1, ..., \gamma_{3g-3}$ appearing as boundaries of the pairs of pants, and also a ‘yellow curve’ α_i for each γ_i as shown in Figure 19. Now if a diffeomorphism f fixes all γ_i, f must fix all the pairs of pants. Therefore f is a composition of Dehn twists in γ_i. But if f fixes α_i too, f can not have any twist along γ_i either, so f must be isotopic to identity. ♠

Now keeping the notations as in the Lemma 4.2, we look at the images of the curves $\gamma_1, ..., \gamma_{3g-3}$ and $\alpha_1, ..., \alpha_{3g-3}$ under the iterates of f. The first n so that f^n acts by identity on these curves is the order of f. Now concerning the time for this algorithm, we have the following Theorem.

4.2 Theorem. Keeping the genus fixed, the time this algorithm takes to identify the diffeomorphism f is exponential in the length of f as a word in the set of basic generators. I.e., there is a constant M (depending on the genus of the surface), such that for any $f = \varphi_k ... \varphi_1$ where $\varphi_i \in \{f_1^{\pm 1}, ..., f_{2g+1}^{\pm 1}\}$, the time is $\leq M^k$.

4.3 Note. By time of the algorithm we mean the following: Suppose each basic operation (addition, multiplication, etc.) of two numbers takes 1 unit of time. Then what 4.2 is claiming is that the total time needed for the algorithm to come up with the answer is $\leq M^k$ for some constant M.

Proof of 4.2. Assume every Dehn twist in a basic set of generators splits into at most N linear maps. Then we have to deal with the eigenvalues of $\leq N^k$ matrices, which obviously takes exponential time. The only problem is when f is of finite order, because we don't know how much we have to wait until a power of f acts as identity on the set of curves introduced in the lemma. But here a famous upper bound comes to rescue, as follows. First, every finite order diffeomorphism can be realized as an isometry for a suitable hyperbolic metric [K]. On the other hand, a finite group of isometries of a Riemann surface has order $\leq 84(g-1)$ (see, for example, [ACGH], pp 45-46). Therefore, f has order at most $84(g-1)$, and the time will still stay exponential. ♠

§5. New coordinate system for the space of measured train-tracks

In this section we prove

5.1 Theorem. Any measured train-track on a once-punctured surface S of genus $g \geq 2$ is equivalent to a unique π_1-train-track.

To prove this theorem, we need some definitions. We denote measured train-tracks by (τ, μ) etc., where τ is train-track and μ is a measure on τ.

Let

$$(\tau, \mu) \to (\tau_1, \mu_1) \to \ldots \to (\tau_n, \mu_n)$$

be a sequence of moves. Then we say that μ is carried by the chain $\tau \to \tau_1 \to \ldots \to \tau_n$. Call μ rational if $\mu(b)$ is rational for every branch b of τ. If (τ, μ) and (τ_n, μ_n) are measured train-tracks, say $\lim_n \mu_n = \mu$ if $\lim_{n\to\infty} \mu_n(b) = \mu(b)$ for any branch b of of τ.

5.2 Lemma. Let μ be an irrational measure on τ. Then there is a sequence of rational measures on τ, such that $\lim_n \mu_n = \mu$.

Proof. Recall that the space of measures $V(\tau)$ is a set defined by linear inequalities with rational coefficients. Therefore, if $\mu \in V(\tau)$ is irrational, the line $t\mu$ ($t \in \mathbf{R}$) can not be isolated in $V(\tau)$, hence there must be a sequence $\mu_n \in V(\tau)$ of rational measures such that $\lim_n \mu_n = \mu$. ♠

5.3 Lemma. Let $(\tau_1, \mu_1) \to \ldots \to (\tau_n, \mu_n)$ be a sequence of measured train-tracks and moves. Then there is a sequence

$$\{(\tau_1, \mu_1^{(k)}) \to \ldots \to (\tau_n, \mu_n^{(k)})\}_k$$

of rational measured train-tracks such that $\lim_k \mu_i^{(k)} = \mu_i$.

Proof. Without loss of generality, we can assume $n = 2$. The space of measures μ_1 at which we can perform the move, forms a set defined by linear inequalities with rational coefficients. Again as in the proof of Lemma 1, if μ_1 is not rational, one can find a sequence of rational measures converging to μ_1 so that we can still perform the same move. ♠

Proof of 5.1. Let (τ, μ) be a measured train-track.

Step 1. The measured train-track (τ, μ) is equivalent to a π_1-train-track.

Assume τ is transversal to ∂R. As in §3, we will try to eliminate bad pairs and bad arcs. We can eliminate a bad pair by splitting. This will not increase the number of bad curves. Eliminating a bad curve might not decrease the number of the bad curves, but it definitely reduces $\#(\partial R \cap \tau)$, which is a finite number by transversality. Therefore using $\#(\text{bad pairs}) + \#(\partial R \cap \tau)$ as a measure of complexity, we see that the procedure of simplifying the measured train-track ends, and therefore after collapsing, we arrive at a π_1-train-track.

Step 2. Uniqueness. Suppose (τ, μ), (τ', μ') are two π_1-train-tracks, and there is a sequence of moves

$$(\tau, \mu) \to (\tau_1, \mu_1) \to \ldots \to (\tau_n, \mu_n) = (\tau', \mu') \qquad (5.4)$$

We have to show $(\tau, \mu) = (\tau', \mu')$.

Let's prove this first in case where μ is an integral measure. In this case all $\mu_1, \ldots, \mu_n$ are integral as well. Using μ we can construct a multiple simple closed curve, by just replacing each branch of measure N by N parallel strands. We denote this curve by $c(\mu)$. Then The sequence (5.4) will correspond to

$$c(\mu) \to c(\mu_1) \to \ldots \to c(\mu_n) = c(\mu').$$

and all the moves will just correspond to isotopy. So $c(\mu)$ is isotopic to $c(\mu')$. Without loss of generality, we can assume that $c(\mu)$ has only one component. Hence $c(\mu)$ and $c(\mu')$ can be represented by cyclic words in a set of generators of $\pi_1(S)$ (namely, the generators obtained by the edge pairing isometries of the fundamental domain R. See §7). But two different cyclic words can not represent the same simple closed curve, because $\pi_1(S)$ is free. This shows $\mu = \mu'$, and the uniqueness is proved in the case of integral (or rational) measures.

Now suppose that μ is no longer rational. By Lemma 5.3, there is a sequence of rational measured train-tracks

$$\{(\tau, \mu^{(k)}) \to (\tau_1, \mu_1^{(k)}) \to \ldots \to (\tau_n, \mu_n^{(k)}) = (\tau', \mu'^{(k)})\}_k$$

such that $\lim_k \mu_i^{(k)} = \mu_i$. By the proof in the special case, $\mu^{(k)} = \mu'^{(k)}$ for all k. Therefore, $\mu = \lim_k \mu^{(k)} = \lim_k \mu'^{(k)} = \mu'$. This finishes the proof of the theorem. ♠

§6 The case of a closed surface.

In this section we want to show how to carry out our algorithm in the case of a closed surface. The section in [B-S] on closed surfaces is relevant. Also look at [T].

Let M be a closed surface of genus $g > 1$ (the case $g = 1$ being trivial). Let $f \in \mathrm{Diff}^+(M)$ be given as a word w in the basic set of generators (Dehn twists along the curves in Figure 4, except the surface is closed now). Since none of the curves in Figure 4 pass through the puncture, the word w can be considered as a diffeomorphism of the punctured surface S. If the algorithm (see §§3, 4) says w is isotopic to a reducible or finite order diffeomorphism of S, then the same can be said in M. So let's concentrate on the case when w is a pseudo-Anosov diffeomorphism of S. We can construct a measured foliation S using the invariant train-track (see [T] or [M2]). This measured foliation is allowed to have a finite number of singularities of order ≥ 3 except for the puncture point, which is allowed to have a singularity of order ≥ 1. If the order of singularity at the puncture is ≥ 3, the foliation can be considered as a measured foliation on the closed surface M, and w gives a pseudo-Anosov diffeomorphism on M.

If the order of singularity at the puncture is 2, then by changing the puncture to another singular point (which always exists if $g > 1$) and treating the singularity of order 2 as a regular point we get the same result that w is pseudo-Anosov on M.

If the order of the singularity at the puncture is 1, then by the uniqueness of stable and unstable invariant measured foliations [FLP], f can not be a pseudo-Anosov diffeomorphism of M , because then f will keep invariant (up to isotopy) two measured foliations of different combinatorial type. So by Thurston classification (§4), $f \in \mathrm{Diff}^+(M)$ is either reducible or isotopic to a finite order diffeomorphism.

Before we indicate how to find the isotopy classes of simple closed curves fixed by f (in case f is reducible), we have to discuss how to represent simple closed curves on M.

6.1 Shortest length representative for a simple closed curve

Suppose a simple closed curve α is given on the closed surface M. Also assume that α does not pass through the puncture. The curve α can be considered as a

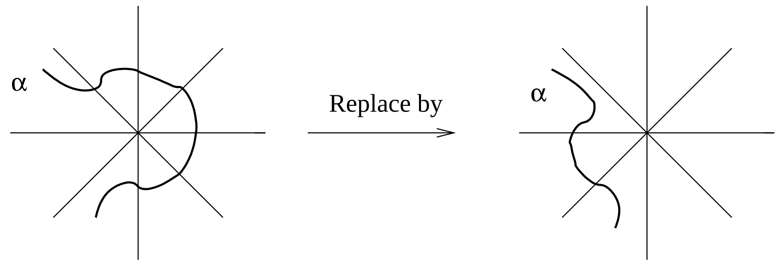


Figure 20. A neighborhood of the puncture.
Straight lines are parts of the boundry of R
coming together.

curve on the punctured surface S, and can be given by a word as in 3.1(iii). Notice that by isotoping α in M, we might obtain a shorter word. We want to discuss here a geometric analogue of Dehn's well-known algorithm (see [J], e.g.) for getting a shortest representative. The shortest representative is not unique, as we will see below.

First isotope α so that the word representing α in the punctured surface does not have back-tracking. During the procedure, always reduce back-trackings as they come along. By following the path of the curve, start replacing arcs which go around the puncture more than half way by arcs going the other way. (Figure 20)

Every move like this reduces the length of the word, so in the end we get a word which can not be reduced any further. The only non-uniqueness comes from the fact that there is no preferred choice between the arcs that go half-way around the puncture. This is because any such α is carried by a π_1-train-track τ, and for a π_1-train-track, there is a canonical sequence to put it in a simple form, as we will see in 6.4.

6.2 How to detect when a diffeomorphism is reducible.

By arguments as in §4, we only have to answer this question:

Suppose $f \in \mathrm{Diff}^+(M)$. Let τ_1, τ_2 be π_1-train-tracks, and $V_1 \subset V(\tau_1)$, $V_2 \subset V(\tau_2)$ be linear subspaces such that f induces a linear map

$$f^* : V_1 \to V_2$$

How can we find all the simple closed curves α carried on τ_1 such that the corresponding measure is $\mu(\alpha) \in V_1$ and $f^*(\mu(\alpha)) = \mu(\alpha)$?

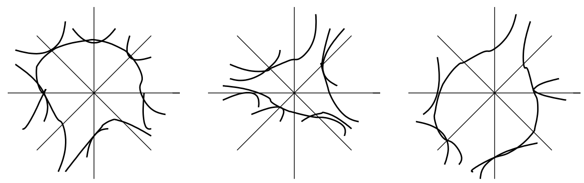


Figure 21. Left to right: Non−simple, simple, simple.

Notice that this problem is more subtle than just looking at the eigenspace of f^* corresponding to $\lambda = 1$, since a curve on the closed surface M can be represented by integral measures on different π_1-train-tracks.

To solve this problem, first we have to put both τ_1 and τ_2 in a simple form. Let's see what we mean by a simple form.

6.3 Definition. (Simple π_1-train-track) A π_1-train-track τ on a once punctured surface S is called simple if there is no smooth path in τ going around the puncture more than half-way.

For examples, look at Figure 21.

6.4 The move. Now we want to describe a move to put a measured π_1-train-track in a simple form, so that the closed curves on the closed surface can be carried along with the move.

So let's start with a non-simple measured π_1-train-track τ_1. So there is a unique maximal path in τ_1 going around the puncture more than half-way. The path is made of a finite number of branches $b_1, ..., b_k$. Let $x_i = \mu(b_i)$ be the measure on each branch b_i, $i = 1, ..., k$. We have to use a move as illustrated in Figure 22 to put τ_1 in the most simplified form. To be able to do that move, we have to assume $x_i = \min\{x_1, ..., x_n\}$.

This means that we have to consider n different cases, and continue in the same manner until there is no smooth path of τ_1 going around the puncture more than half-way. The important thing is the number of cases is bounded by a function $N(g)$ of the genus g. Thus we get a sequence

$$(f^*)_j : (V_1)_j \to (V_2)_j$$

where $(V_1)_j$, $(V_2)_j$ are subspaces of the space of measures on simple π_1-train-tracks $(\tau_1)_j$, $(\tau_2)_j$. The index j varies in an index set J with $|J| \leq N(g)$. The train-tracks $(\tau_1)_j, (\tau_2)_j$ are derived from τ_1, τ_2 by a sequence of moves described in Figure 22,

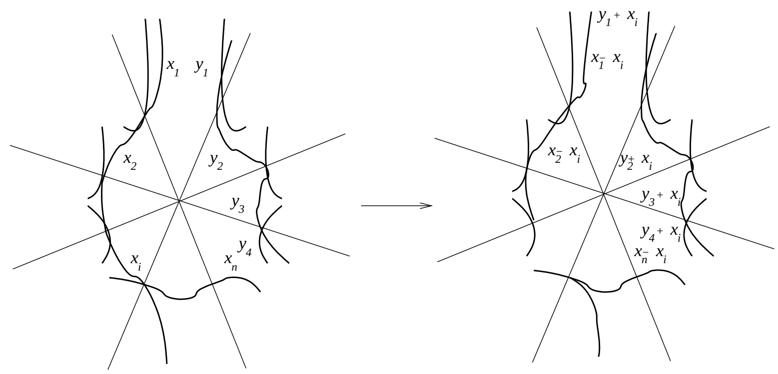


Figure 22.

and the linear map $(f^*)_j$ is the map derived from f^*. There are two possibilities for $(\tau_1)_j$. The first possibility is that around the puncture, $(\tau_1)_j$ has 2 smooth paths, each one going half-way around the puncture. This is the case that non-uniqueness comes in, so let's discuss this case later. If this not the case, the problem is easy. just look at the cases which $(\tau_1)_j = (\tau_2)_j$, and find the eigenspace of $(f^*)_j$ corresponding to $\lambda = 1$, and then intersect it with $(V_1)_j$ and $(V_2)_j$. All the integral non-negative measures gotten this way will give a family of simple closed curves fixed by f up to isotopy.

Now let's discuss the case where non-uniqueness occurs. The following example will illustrate the general solution.

6.5 Example. Suppose τ is a π_1-train-track and $F : U_1 \to U_2$ is a linear map, where U_1, U_2 are linear subspaces of $V(\tau)$. (Figure 23)

Any simple closed curve α carried on τ can be represented by many measures on τ. Namely, if you replace x_i by $x_i + t$ and y_i by $y_i - t$ for all $i = 1, ..., 4$, you get the same curve. By this it is easily seen that finding fixed systems of curves for F reduces to solving the system of linear equations

$$F(..., x_i + t, ..., y_i - t, ...) = (..., x_i + s, ..., y_i - s, ...)$$

The non-zero solutions of the above system with all measures non-negative integers (s, t are allowed to be any integers) give all collections of curves fixed by F carried on τ.

6.6 Remark. Notice that if a diffeomorphism f fixes a collection of simple closed curves, it has to occur as one of the cases discussed, since any collection of simple

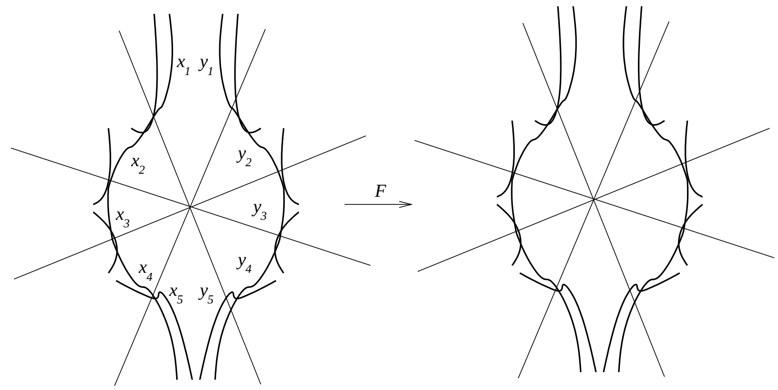


Figure 23.

closed curves can be put in a shortest form and hence can be carried by a simple π_1-train-track. This provides us with an algorithm to find out if a diffeomorphism f of a closed surface is reducible. The algorithm is still of exponential time because of the existence of the function $N(g)$.

6.7 How to detect when a diffeomorphism is isotopic to a finite order diffeomorphism

This is now very easy. Lemma 4.1 is still true for a closed surface. As in the case of punctured surface, check if f^n, $n = 1, ..., 84(g-1)$ (see the proof of 4.2) fixes any of the $6g-6$ closed curves mentioned 4.1 and its proof. Have in mind that to find out if a curve α is fixed under f, you have to put both α and $f(\alpha)$ in the shortest form, and see if they are the same curve.

§7. Relation with Birman-Series conditions

In this section we prove

7.1 Theorem. Let $C = \{c_1, ..., c_n\}$ be a set of simple closed curves in some cone produced as in Proposition 2.1 for a diffeomorphism f. Then the action of f on C satisfies the Birman-Series Conditions. In particular, f acts linearly on C.

First let's explain briefly what Birman-Series Conditions are. For further details and proofs see [B-S]. Choose a fixed hyperbolic metric on S, and a fundamental domain R in $\mathbf{H}^2$. Elements of $\Gamma = \pi_1(S)$ will be thought of as isometries of $\mathbf{H}^2$. Use the set of generators Γ_R formed by edge-pairing transformations of R. Make Γ_R

symmetric, i.e., so that $g \in \Gamma_R$ implies $g^{-1} \in \Gamma_R$. Let N be the image of ∂R under Γ. Each edge in N is labeled by an element in Γ_R. If α is an oriented simple closed curve on S, let $\tilde{\alpha}$ be any lift of α to $\mathbf{H}^2$. Then we can associate to α a cutting sequence $\sigma(\alpha) = ...e_{-1}e_0e_1...$, $e_i \in \Gamma_R$. This sequence does not depend on the particular $\tilde{\alpha}$ chosen, up to shifting. Call $\tilde{\alpha}$ *tight* if each finite subsequence of $\sigma(\alpha)$ is freely reduced. A family $\tilde{C}$ of tight curves is called tight if each two curves in $\tilde{C}$ intersect in at most one point, and no more that two curves pass through each point. Let $\tilde{f}$ be a lift of f to $\bar{\mathbf{H}}^2 = \mathbf{H}^2 \cup \partial\mathbf{H}^2$. Let α_1 and α_2 be two simple loops supported on a π_1-train-track τ. Assume $f(\alpha_1)$ and $f(\alpha_2)$ are supported on a π_1-train-track τ' as well. Let $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ be any lifts of α_1 and α_2, such that $\tilde{\alpha}_1 \cap \tilde{\alpha}_2 = \{p\}$. If $p \in gR$, then $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ hit a common side of gR (Otherwise they can't be supported on the same train-track). So $\sigma(\alpha_1)$ and $\sigma(\alpha_2)$ have a common block. We say $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ have coherent orientation if they have the same orientation on the common block. Notice that $\tilde{f}(\tilde{\alpha}_1)$ and $\tilde{f}(\tilde{\alpha}_2)$ intersect in one point too. They are supported on τ' by assumption, so we can see if the orientation induced by f and a coherent orientation on $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ on $\tilde{f}(\tilde{\alpha}_1)$ and $\tilde{f}(\tilde{\alpha}_2)$ is coherent or not. In the case this orientation is coherent, we say that f preserves τ-orientation on $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ (or α_1 and α_2). If C is a family of simple closed curves supported on τ such that $f(C)$ is supported on τ', We say f preserves τ-orientation on C if f preserves τ-orientation for every pair $\alpha_1, \alpha_2 \in C$ and any lifts $\tilde{\alpha}_1, \tilde{\alpha}_2$ such that $\#(\tilde{\alpha}_1 \cap \tilde{\alpha}_2) = 1$.

Suppose now that $\alpha_1, \alpha_2, \alpha_3$ are simple closed curves supported on τ, and $f(\alpha_1)$, $f(\alpha_2), f(\alpha_3)$ are supported on τ'. Assume $\tilde{\alpha}_i$ is a lift of α_i such that they are tight, and each two of them intersect at one point. Moreover, assume that all the intersection points lie in the same polygon gR (We can always make this happen by using finger moves; see [B-S]). Give $\tilde{\alpha}_i$ orientations such that the triangle Δ formed by the intersection points and segments of $\tilde{\alpha}_i$ has clockwise orientation. We can look at the corresponding triangle for $\tilde{f}(\tilde{\alpha}_i)$. If the orientation induced by f on this triangle is clockwise, we say that f preserves τ-orientation on the trigon Δ. If C is a collection of simple closed curves on S, we say that f preserves τ-orientation on C-trigons if for any $\alpha_1, \alpha_2, \alpha_3 \in C$ and any trigon Δ formed by some lifts of the α_i lying entirely in a polygon gR ($g \in \pi_1(S)$), the diffeomorphism f preserves τ-orientation on Δ.

Now we are ready to state the Birman-Series Conditions. Suppose C is a collection of simple closed curves supported on a π_1-train-track τ. Also assume that all the curves in $f(C)$ are supported on a π_1-train-track τ'. Then we say that C satisfies the Birman-Series Conditions for f if

(i) The diffeomorphism f preserves τ-orientation on C.

(ii) The diffeomorphism f preserves τ-orientation on C-trigons.

Proof of Theorem 7.1. Take $C = \{c_1, ..., c_n\}$ to be a family of simple closed curves on S and choose the representatives c_i in their isotopy class so that the intersection number of each pair is minimal (e.g., take all of them be geodesics in a fixed hyperbolic metric). Let $\tilde{c}_i$ be a (tight) lift of c_i to the universal cover $\mathbf{H}^2$. The curves $\tilde{c}_i$ are supported on $\tilde{\tau}$, the lift of τ. Suppose $\tilde{c}_i \cap \tilde{c}_j = \{p\}$. Orient $\tilde{c}_i, \tilde{c}_j$ coherently. The image under $\tilde{f}$ of $\tilde{\tau}$, $\tilde{f}(\tilde{\tau})$, might not be a π_1-train-track, but by Theorem 5.1, There are (possibly many) sequences of moves to make it a π_1-train-track. Since c_i and c_j are in the same cone (see Proposition 2.1), this means that they both satisfy the same types of inequalities. For example, when we shift or split, we can do it for both curves in the same way. This means that the angle between $\tilde{f}(\tilde{c}_i), \tilde{f}(\tilde{c}_j)$ will not be altered much in the process of isotoping them onto τ'. In particular, we can keep the angle always acute. This implies that f preserves τ-orientation on C.

To prove part (ii) of the Birman-Series Conditions for C, let $c_i, c_j, c_k \in C$. Assume $\tilde{c}_i, \tilde{c}_j, \tilde{c}_k$ specify a triangle Δ in some gR. Orient the curves so that Δ has clockwise orientation. The trigon Δ corresponds to a trigon in $\tilde{\tau}$, so $\tilde{f}(\Delta)$ will correspond to a trigon in $\tilde{f}(\tilde{\tau})$. But again as before, since c_i, c_j, c_k are in the same cone, they all can be isotoped using exactly the same moves onto $\tilde{\tau}'$ equivariantly. This proves part (ii) of the Birman-Series Conditions. ♠

Let's now look at the relation between the Birman-Series Conditions and decomposition into cones (Proposition 2.1).

7.2 Theorem. Birman-Series conditions are necessary and sufficient in case when f is a Dehn twist along a non-separating simple closed curve α. Namely, if f is a Dehn twist along α, and $C = \{c_1, ..., c_k\}$ is a set of non-trivial isotopy classes of simple closed curves on S then f acts linearly on $Span^+(C)$ if and only if the diffeomorphism f satisfies Birman-Series conditions on C.

Proof. This a direct consequence of Theorem 2.2 and Corollary 3.5. ♠

7.3 Remark. Theorem 7.2 will not be correct in case of a general diffeomorphism. To give a counterexample, all we have to do is to find a diffeomorphism f, a π_1-train-track τ_1, such that $f(\tau_1)$ can be changed to some π_1-train-track τ_2 using the moves in two different ways. Then find appropriate curves carried on τ_1.

7.4 Question. Is there a theorem corresponding to 7.2 for a general diffeomorphism f? I.e., What is a necessary and sufficient condition for a set of simple closed curves C so that f acts linearly on C?

References

[ACGH] E. Arbarello, M. Cornalba, P. A. Griffiths, J. Harris, 'Geometry of algebraic curves', Vol. I, Springer-Verlag, (1985).

[B-H] M. Bestvina and M. Handel, 'Train-tracks for surface homeomorphisms', Topology 34 (1995), No.1, 109-140.

[B-S] J. S. Birman and C. Series, 'Algebraic linearity for an automorphism of a surface group', J. Pure. App. Alg. 52 (1988), 227-275.

[FLP] A. Fathi, F. Landerbach, V. Poenaru, 'Travaux de Thurston sur les surfaces', Seminaire Orsay, (1975), Asterisque No. 66/67.

[H] S. Humphries, 'Generators for the mapping class group', Springer-Verlag lecture notes, # 772, 44-47.

[J] D. L. Johnson, 'Topics in the theory of group presentations', London Math. Soc. Lecture Note Ser., 42, cambridge Univ. Press, 1980.

[K] S. Kerckhoff, 'The Nielsen realization problem', Annals of math 117, (1983), 235-265.

[L] J. E. Los, 'Pseudo-Anosov maps and invariant train-tracks in the disk: a finite algorithm', Proc. London Math. Soc. 66(2), (1993), 400-430.

[M1] L. Mosher, 'Pseudo-Anosovs on punctured surfaces', 1983. Princeton dissertation.

[M2] L. Mosher, 'The classification of pseudo-Anosovs', Low dimensional topology and Kleinian groups, London Math. Soc. Lecture Note Ser., 112, cambridge Univ. Press, 1986.

[P-H] R. Penner and J. Harer, 'Combinatorics of train-tracks', Princeton Univ. Press, (1992).

[T] W.P. Thurston, Geometry and Topology of 3-manifolds, Princeton University math department, 1979.

[W] B. Wajnryb, 'Two generators for the mapping class group of a surface', preprint.

Department of Mathematics,
Columbia University,
New York, NY 10027.

E-mail: hessam@math.columbia.edu, zc@math.columbia.edu