ON COMPLEXITY OF THE WORD PROBLEM IN BRAID GROUPS AND MAPPING CLASS GROUPS

Hessam Hamidi-Tehrani

September 25, 1998

Abstract. We prove that the word problem in the mapping class group of the once-punctured surface of genus \( g \) has complexity \( O(|w|^2 g) \) for \( |w| \geq \log(g) \) where \( |w| \) is the length of the word in a (standard) set of generators. The corresponding bound in the case of the closed surface is \( O(|w|^2 g^2) \). We also carry out the same methods for the braid groups, and show that this gives a bound which improves the best known bound in this case; namely, the complexity of the word problem in the \( n \)–braid group is \( O(|w|^2 n) \), for \( |w| \geq \log n \). We state a similar result for mapping class groups of surfaces with several punctures.

Key words: Mapping class group, measured train-track, \( \pi_1 \)-train-track, braid group, word problem, complexity.

AMS subject classification: 20F10, 68Q25, 57S05

§0. Introduction

A group \( G \) is said to have a solvable word problem if there is a finite generating set \( S \) for \( G \) such that there is an algorithm to decide if a given word \( w \) in \( S \) represents the identity element in \( G \). The word problem is said to have complexity \( O(f(|w|)) \) if there exist such an algorithm which takes \( \leq kf(|w|) \) steps on a Turing Machine (TM) to produce a “yes” or a “no”, for a word \( w \) of length \( |w| \) where \( k \) is a constant (see Appendix for more on complexity and Turing Machine). The conjugacy problem is defined similarly, but the objective is to decide if two given words are conjugate in the group \( G \).

Sometimes one has to deal with sequence of groups \( G_n \) depending on an integer parameter \( n \) (say mapping class groups of closed surfaces which is parameterized by genus), and one

\[^1\text{Department of Mathematics, University of California at Santa Barbara, CA 93106 E-mail: hes sam@math.ucsb.edu. http://www.math.ucsb.edu/~hessam.}\]

Typeset by \texttt{AMSTeX}
can pose the question of how the complexity of a problem grows as \( n \) becomes larger. This is a crucial issue in implementation of a uniform algorithm, because the parameter becomes an input. In this case we say the word problem has uniform complexity \( O(f(|w|, n)) \) for the groups \( G_n \) if there exists some finite set of generators for each \( G_n \) such that for a word \( w \) in generators of \( G_n \) of length \(|w|\), it takes a Turing Machine \( \leq kf(|w|, n) \) steps to determine if \( w = 1 \).

The word problem and conjugacy problem in the mapping class group have been known to be solvable for a long time (see \([G],[He],[P],[Mo1]\)). In recent years, with development of the theory of automatic groups, some new ideas in this direction have been discovered. In \([E]\), the authors discuss an automatic structure derived from Garside’s algorithm \([G]\) for the braid groups. This results in an algorithm which is of uniform complexity \( O(|w|^2n \log n) \), where \( n \) is the number of strands, and \(|w|\) is the length of the braid, which is given as a word \( w \) in the standard set of Artin generators (see (3.1)). Mosher \([Mo2]\) proved that mapping class groups are automatic, giving an algorithm for the word problem which is quadratic in the word length \([Mo3]\), with no implication on uniform complexity. As the authors of \([E]\) mention, it is important to have a bound on the uniform complexity; i.e., in terms of the genus and the number of punctures. Here we prove that the word problem in the mapping class group of the closed surface of genus \( g \) has complexity \( O(|w|^2g^2 + |w|g^2 \log g) \). The corresponding bound for a once-punctured surface of genus \( g \) is \( O(|w|^2g + |w|g \log g) \).

In a sense we answer the Open Question 9.3.10 in \([E]\), but we do not use the automatic theory. Our methods rely on the action of the mapping class group on the space of curves, or measured train-tracks. This could be related to the Open Question 9.4.5 in there as well, although we do not speak about conjugacy problem at all. It is an interesting question to try to use the methods here to solve and analyze the complexity of the conjugacy problem in the mapping class groups. In this respect the work of Kleinberg and Menasco \([KM]\), Masur and Minsky \([MM1]\), \([MM2]\) is of interest. In particular, the authors of the latter prove that if two pseudo-Anosov maps are conjugate, then there is a conjugating element whose word length is linearly bounded by the larger of the word lengths of those elements.

Our methods apply to the braid groups \( B_n \) and give the complexity \( O(|w|^2n + |w|n \log n) \), which is the best known bound to date. In \([BKL]\) the authors give a fast and practical al-
algorithm for the word problem in $B_n$, which works well with a “Random Access Memory” (RAM) machine, and has “complexity” $O(|w|^2n)$. But RAM is usually much faster than TM (In particular, they assume that the braid index $n$ can be encoded in one unit of memory; see Appendix), and their algorithm gives the same complexity as in [E], namely $O(|w|^2n \log n)$ if practiced on a TM.

Here is an outline of the rest of this paper: In §1 we develop the necessary notation for measured $\pi_1$-train-tracks and the mapping class groups. In §2 we prove the bound on the complexity of the word problem in once-punctured surfaces. In §3 we apply our methods to deduce a bound on the complexity of the word problem in the braid groups. In §4 we develop the theory for closed surfaces; we prove the analog to Theorem 1.5 for closed surfaces. §5 is devoted to analyze the complexity of the word problem in closed surfaces. Finally in the appendix we briefly address some issues about our definition of complexity.

Acknowledgments. Some of this work was part of my Ph.D. thesis in Columbia University. I thank everybody in the Mathematics Department, especially my advisor Joan Birman for her graceful support and guidance. Also, I found the referee’s comments most valuable.

§1. Some notation and background on train-tracks

Let $S = S^p_g$ be an oriented surface of genus $g$ with $p$ fixed points, called punctures. Let $M = M_S = M^p_g$ the mapping class group of $S$, i.e., the group $\mathcal{H}(S)/\mathcal{H}_0(S)$, where $\mathcal{H}(S)$ is the group of homeomorphisms of $S$ fixing the punctures pointwise, and $\mathcal{H}_0(S) \subseteq \mathcal{H}(S)$ is the (normal) subgroup of the ones homotopic to identity within $\mathcal{H}(S)$. We denote the elements of $M$ by $f, g$, etc. An element of $M$ can be thought of as an isotopy class of a homeomorphism (or diffeomorphism) of $S$. Sometimes we pick a representative of the class $f$ and call it $f$ too. We assume $S$ has a given smooth or piecewise linear structure, depending on what suits the situation the best.

Notice that if $S'$ is a surface with $b$ boundary components, one can define the mapping class group $M_{S'}$ of $S'$ by the group of isotopy classes of diffeomorphisms which fix the boundary components pointwise. Let $S$ be obtained by shrinking the boundary components
of $S'$ to punctures. Then we have the short exact sequence

$$1 \to \mathbb{Z}^b \to M_{S'} \to M_S \to 1.$$  

In the following we only study the surfaces $S_g^p$. The corresponding information about surfaces with boundary can be obtained using (1.1).

**Definition 1.1** (Train-track). (See [PH].) A compact, connected subset $\tau$ of $S$ is called a train-track if $\tau$ is a smooth branched 1-manifold embedded smoothly in $S$. At each branch point $v$ (also called a switch point) there is a well-defined tangent space. Every connected component of $\tau - \{\text{branch points}\}$ is called a branch. There is a natural partition into two subsets for the set of 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. Also, there is a “hyperbolicity condition” on the complement $S - \tau$: The doubles of components of $S - \tau$ must have negative Euler characteristic. Notice that the double of “corners” give rise to punctures. In computing the Euler characteristic, every puncture contributes a -1.

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

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

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

**Definition 1.3** ($\pi_1$-train-track). (see [BS]) Suppose $S = S_g^p$ is a surface with $\chi(S) = 2 - 2g - p < 0$. The universal cover of $S$ then can be identified with hyperbolic plane $\mathbb{H}^2$. Fix a polygon $R$ in $\mathbb{H}^2$ as a fundamental domain for the action of $\pi_1(S)$ on $\mathbb{H}^2$. Notice that $R$ is naturally identified with $S$ cut open along a number of arcs. Let $\tau$ be a train-track in $S$. We call $\tau$ a $\pi_1$-train-track (with respect to the choice of $R$) if the following conditions hold: If we look at $\tau$ 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 $\partial R$. (not necessarily distinct in $S$.)
1.4. The Moves. (see [PH]) We denote by $\mathcal{MT}(S)$ 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 1.1).

(iii) Shift (Figure 1.2).

We have only shown the relevant piece of the train-track in Figures 1.1, 1.2. Notice that the inverse of a split is called a collapse.

The set of measures on a train-track $\tau$ is denoted by $V(\tau)$, and can be identified with a subset of some Euclidean space defined by a finite set of equalities and inequalities. The set $V(\tau)$ is closed under (positive) scalar multiplication and addition. In particular, it is a convex cone.

The following theorem, which is probably due to Thurston, gives a coordinate system for $\mathcal{MT}(S)$, in the case which $S$ has negative Euler characteristic and is not closed.

**Theorem 1.5.** Let $S$ be a non-closed surface (i.e., $p > 0$) with $\chi(S) < 0$, and let $R$ be a polygon representing a fundamental domain for the action of $\pi_1(S)$ on the hyperbolic plane. Then any measured train-track on $S$ is equivalent to a unique $\pi_1$-train-track with respect to
In particular, every non-trivial multiple closed curve corresponds to a unique (integral) measured $\pi_1$-train-track.

This theorem is proved in [HC] (see Theorem 5.1 there) in the case of a surface with 1 puncture. The general proof is completely similar. The following direct corollary gives a piecewise linear structure on $\mathcal{MT}(S)$.

**Corollary 1.6.** For a surface $S$ and polygon $R$ as above, $\mathcal{MT}(S)$ is the finite union of the cones $V(\tau)$ where $\tau$ ranges over the finite set of $\pi_1$-train-tracks with respect to $R$.

For any surface $S$ the mapping class group $M_S$ acts on $\mathcal{MT}(S)$, since if one changes a train-track $\tau$ by any of the moves (i)-(iii) or change a homeomorphism $f : S \to S$ by isotopy, then $f(\tau)$ changes by a sequence of the moves (i)-(iii). When a homeomorphism $f$ acts on a $\pi_1$-train-track $\tau$ it need not map it to a $\pi_1$-train-track. Using Theorem 1.5 one can put the image $f(\tau)$ in the $\pi_1$-train-track by a sequence of the moves (i)-(iii). We will study how these moves must be performed, and what the corresponding action of $f$ on $V(\tau)$ is.

Let $S = S^g_\chi$ be a surface with $\chi(S) < 0$, and the polygon $R$ be a fundamental domain for the action of $\pi_1(S)$ on $\mathbb{H}^2$.

**1.7.** 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 $\partial R$. If $e_i$ is identified with $e_j$ in $S$ (obviously with the opposite orientation, since $S$ is orientable), we denote that by $e_i = e_j^{-1}$.

Pick a base point $x_0$ in the interior of $R$. We want to specify a set of generators for $\Gamma = \pi_1(S, x_0)$. Let $\gamma_i$ be a simple closed curve based at $x_0$ defined as follows: It starts at $x_0$, it crosses $e_i$ (it naturally comes out of $e_j = e_i^{-1}$) and then it goes back to $x_0$, without crossing $\partial R$ any further. The curve $\gamma_i$ gives rise to an element in $\Gamma$, which by abuse of notation we call $e_i$ too. Notice that the equation $e_j = e_i^{-1}$ holds in $\Gamma$ as well. It is easy to see that $e_1, ..., e_n$ generate $\Gamma$.

**1.8.** A simple closed curve $C$ can be given by a cyclic word $e_{\alpha_1}...e_{\alpha_k}$ where $1 \leq \alpha_i \leq n$. To draw the curve in $R$ from the given word, just start on the base point $x_0$, go to $e_{\alpha_1}$, come out of interval $e_{\alpha_1}^{-1}$ and connect it to $e_{\alpha_2}$, so that it’ll come out of $e_{\alpha_2}^{-1}$, etc. All the
curves that we consider are assumed to be tight, i.e., \( \alpha_i^{-1} \neq \alpha_{i+1} \) for all \( i \) (consider \( i \) to be a cyclic index modulo \( k \)).

1.9. Let’s set up some notation for the case when \( S = S_g^1 \) is a surface of genus \( g \) with one puncture \( P \), since this case is the simplest case. We use the standard fundamental domain \( R \) for the surface \( S \), which is a \( 4g \)-gon with edges labeled as \( E = E(R) = (a_1, b_1, a_1^{-1}, b_1^{-1}...a_g, b_g, a_g^{-1}, b_g^{-1}) \), in clockwise order. We call \( E \) the edge set.

When we draw curves in \( R \), if we are only interested in their free isotopy class, we draw them off the base point. It is important to notice, for example, that the curve given by the sequence \( a_1 \) is different from edge \( a_1 \). It is actually parallel to the edge \( b_1 \), but in different orientation. Also, the curve \( b_1 \) is parallel to edge \( a_1^{-1} \), with the same orientation. Let’s introduce the curves \( x_1, ..., x_g \). For \( 1 \leq i \leq g \), the curve \( x_i \) is given by the sequence \( b_i a_{i+1} \) (take the indices mod \( g \), for example, in the case \( i = g \) in the definition of \( x_i \)). Let \( D_c \) denote the (right-handed) Dehn twist about the simple closed curve \( c \). By [Hu] or [B] we have

\[
M_S = M_g^1 = \langle D_{a_1}, D_{b_1}, ..., D_{a_g}, D_{b_g}, D_{x_1}, ..., D_{x_{g-1}} \rangle.
\]

One has to notice that, the same set generates \( M_g^0 \) if the curves are considered in the closed surface.

Any mapping class \( f \) on \( S_g^1 \) ( \( g > 2 \) ) is specified with its action on the simple closed curves (with base point)

\[
a_1, b_1, ..., a_g, b_g.
\]

If so, then for any simple closed curve \( c = e_1...e_N \), \( f(c) = f(e_1)...f(e_N) \). This is simply because \( f \) induces a homomorphism on the fundamental group, and if \( f \) induces the identity on \( \pi_1(S) \), \( f \) is the identity mapping class. (If \( g = 2 \) then \( f \) also could be hyperelliptic involution.)

We know that \( M_S \) is generated by finitely many Dehn twists. Therefore, it is enough to study the action of a single Dehn twist on a measured \( \pi_1 \)-train-track \( \nu = (\tau, \mu) \).

For a \( \pi_1 \)-train-track \( \tau \) on \( R \), we call a branch \( b \) of \( \tau \) outer if it connects two consecutive edges of the polygon \( R \). Otherwise we call \( b \) inner. By \( \text{out}(\tau) \) (resp. \( \text{inn}(\tau) \)) we mean the set
of outer (resp. inner) branches of $\tau$. The train-track $\tau$ is identified with the set of branches of $\tau$. So $\tau = \text{inn}(\tau) \cup \text{out}(\tau)$. We say a measured train-track $\nu$ is precisely carried on a $\pi_1$-train-track $\tau$, if $\nu$ is carried on $\tau$ and is not carried on any sub-train-track of $\tau$.

For a measured $\pi_1$-train-track $\nu = (\tau, \mu)$, the total measure of $\nu$ is defined by

$$T(\nu) = \sum_{b \in \tau} \mu(b).$$

Notice that $T(a\nu) = aT(\nu)$ for $a > 0$. The space of projective measured train tracks can then be defined by

$$\mathcal{PMT}(S) = \{\nu \in \mathcal{MT}(S) \mid T(\nu) = 1\}.$$

Also the canonical projection $\mathcal{MT}(S) \setminus \{0\} \xrightarrow{\cdot} \mathcal{PMT}(S)$ can be defined by $[\nu] = \nu/T(\nu)$.

§2. Complexity of the word problem in the mapping class groups of once-punctured surfaces

Let $S = S^1_g$. As we saw before, a generating set for $M_S$ is given by (1.2). In this section we consider the following problem: What is the complexity of computing (i) $D_{a_i}(\nu)$ or $D_{b_i}(\nu)$, (ii) $D_{x_i}(\nu)$ and (iii) $D_{\tilde{a}_i}(\nu)$ or $D_{\tilde{b}_i}(\nu)$ for a given integral measured $\pi_1$-train-track $\nu = (\tau, \mu)$. Let $T(\nu) = \ell$. Unfortunately the notation in [HC] is different from our notation. There $E(R) = (e_1, \cdots, e_{4g})$ while here $E(R) = (a_1, b_1, a_1^{-1}, b_1^{-1}, \cdots)$. Also, in [HC], for $1 \leq t \leq 2g$, the curve $b_t$ is defined to be $e_{2t-1}e_{2t+1}$ for odd $t$ and $e_{2t-2}e_{2t}$ for even $t$. 
In other words, our collection of simple closed curves \( \{a_1, b_1, \cdots, a_{2g}, b_{2g}\} \) is the same as \( \{b_1, \cdots, b_{2g}\} \) in [HC]. To make the notation clear, let \( b'_t \) denote the \( b_t \) in [HC]. We will only use this notation in 2.1 below. Let’s look at the complexity of the computation of \( D_{b'_t}(\nu) \).

### 2.1. Complexity of computing \( D_{b'_t}(\nu) \).

1. Enter \( \nu \) in the machine in the following form: \( L(\nu) = \{(e_i, e_j, \mu(e_i, e_j))\}_{i,j} \), where \( e_i, e_j \) are edges of \( R \), and \( \mu_{ij} = \mu(e_i, e_j) > 0 \) is the corresponding measure. Since there can be at most \( 2|E(R)| - 3 \) branches in \( \tau \), \( L(\nu) \) has \( O(g) \) elements. Since \( 1 \leq e_i, e_j \leq 4g \) and \( 1 \leq \mu_{ij} \leq \ell \), this has complexity \( O(g (\log \ell + \log g)) = O(g \log(g\ell)) \). Notice that entering a number of size \( O(N) \) into the machine has complexity \( O(\log N) \).

2. Put \( k(i) = 2i \) for \( i \) odd and \( k(i) = 2i - 1 \) for \( i \) even. Check if \( \mu_{k(t),k(t+1)} = 0 \). Looking at \( L(\nu) \), this has complexity \( O(g \log(g\ell)) \).

3. If \( \mu_{k(t),k(t+1)} \neq 0 \), go to step 5. If \( \mu_{k(t),k(t+1)} = 0 \), the resulting train-track after applying \( D_{b'_t} \) is collapsible to a \( \pi_1 \)-train-track. One can obtain \( D_{b_t}(\nu) \) by changing all \( (e_i, e_{k(t)}, \mu(e_i, e_{k(t)})) \) to \( (e_i, e_{k(t)+1}, \mu(e_i, e_{k(t)})) \) and adding \( (e_{k(i)}, e_{k(i)-1}, \sum_i \mu(e_i, e_{k(t)})) \) to \( L(\nu) \). This results in a collection \( L_1 = L_1(D_{b'_t}(\nu)) \). Notice that obtaining \( L_1 \) has complexity \( O(g \log(g\ell)) \) as well. Also, \( |L_1| = O(g) \). Also notice that since we added only some of the terms of \( L(\nu) \) at most once, \( T(D_{b'_t}(\nu)) \leq 2T(\nu) = 2\ell \).

4. To obtain \( L(D_{b'_t}(\nu)) \) from \( L_1 \), sort \( L_1 \) Lexicographically in terms of the first two components. Then combine any string of consecutive terms of the form \( (e, e', m_1), \cdots, (e, e', m_s) \) to \( (e, e', \sum_i m_i) \). This gives \( L(D_{b'_t}(\nu)) \), as desired. The sorting and combining processes each have complexity \( O(g \log(g\ell)) \).

5. If \( \mu_{k(t),k(t+1)} \neq 0 \), the resulting train-track after applying \( D_{b'_t} \) is not collapsible to a \( \pi_1 \)-train-track. As in step 3, one can obtain \( D_{b'_t}(\nu) \) by changing all \( (e_i, e_{k(t)}, \mu(e_i, e_{k(t)})) \) to \( (e_i, e_{k(t)+1}, \mu(e_i, e_{k(t)})) \) and adding \( (e_{k(i)}, e_{k(i)-1}, \sum_i \mu(e_i, e_{k(t)})) \) to \( L(\nu) \). This results in a collection \( L_1 = L_1(D_{b'_t}(\nu)) \). Notice that obtaining \( L_1 \) has complexity \( O(g \log(g\ell)) \) as well. The list \( L_1 \) has an element of the form \( (e_{k(i)+1}, e_{k(i)+1}, \mu(k(i), k(i)+1)) \). Drop this from \( L_1 \). This is equivalent to reducing the bad curve. Following 3.2 in [HC], Now we have to do a split. Create two lists \( \{A_1, \cdots, A_n\} \) and \( \{B_1, B_2\} \), as instructed in Figures 8 and 9 there. Then decide which split to do as in Figure 10. All these steps can be implemented with
complexity $O(g \log(g\ell))$. Change $L_1$ accordingly, and then go to step 4 to obtain $L(D_{b_t^*}(\nu))$.

The estimate $T(D_{b_t^*}(\nu)) \leq 2T(\nu) = 2\ell$ still holds.

The steps 1-5 show that

**Theorem 2.2.** Let $\nu = (\tau, \mu)$ be an integral measured $\pi_1$-train track with respect to the standard fundamental domain $R$ for $S^1_g$ with $T(\nu) = \ell$. Then one can compute $D_{a_t}(\nu)$ and $D_{b_t}(\nu)$ with complexity $O(g \log(g\ell))$ and one has $T(D_{a_t}(\nu)) \leq 2\ell$ and $T(D_{b_t}(\nu)) \leq 2\ell$.

Similarly, but a more detailed argument one can obtain from 3.3 in [HC] the following:

**Theorem 2.3.** Let $\nu = (\tau, \mu)$ be an integral measured $\pi_1$-train track with respect to the standard fundamental domain $R$ for $S^1_g$ with $T(\nu) = \ell$. Let $x_t$ be the simple closed curve $b_t a_{t+1}$. Then one can compute $D_{x_t}(\nu)$ with complexity $O(g \log(g\ell))$ and one has $T(D_{x_t}(\nu)) \leq 3\ell$.

The case of $D_{\tilde{a}_t}$ and $D_{\tilde{b}_t}$ was not discussed in [HC]. However, similar arguments can be applied. Since $T(\tilde{a}_t) = T(\tilde{b}_t) = T(\tilde{b}_t) = 4g + 1$, one needs to do steps similar to step 5 in 2.1 $O(g)$ times, therefore giving:

**Theorem 2.4.** There are 4 integral measured $\pi_1$-train-tracks $\nu_i$, $i = 1, \cdots, 4$ on $S^1_g$, $g \geq 2$, such that for $f \in \mathcal{M}_{g,1}^1$, the following condition implies $f = \text{id}$.

\[ (*) \quad f(\nu_i) = \nu_i \text{ for } i = 1, \cdots, 4. \]

**Proof.** Figure 2.1 shows a “pair of pant” decomposition of $S^1_g$ by a set of simple closed curves $P = \{\alpha_i, \beta_i, \gamma_i\}_{i=1}^{g-1} \cup \{\delta\}$. For any curve $\rho \in P$ one can define the simple closed curve $\rho'$ by Figure 2.2. If a mapping class $f$ fixes all the curves in $P$, then it must be a product of $D_{\rho}^{\pm 1}$, $\rho \in P$. If, moreover, $f$ fixes all $\rho'$, $\rho \in P$, then $f = \text{id}$. Set

\begin{align*}
\nu_1 &= \{\alpha_i, \beta_i, \gamma_i\}_{i=1}^{g-1} \cup \{\delta'\}, \\
\nu_2 &= \{\gamma_1', \cdots \gamma_{g-1}'\} \cup \{\delta\}, \\
\nu_3 &= \{\alpha_1', \alpha_3', \cdots \} \cup \{\beta_2', \beta_4', \cdots \}, \\
\nu_3 &= \{\alpha_2', \alpha_4', \cdots \} \cup \{\beta_1', \beta_3', \cdots \}.
\end{align*}

It is easy to see that each collection $\nu_i$ consists of mutually disjoint curves, so can be made into a measured $\pi_1$-train-track. Moreover, by construction, if a mapping class fixes all $\nu_i$, it must be the identity. ♠
Theorem 2.5. The word problem in $M_1^g$ has complexity $O(|w|^2 g + |w| g \log g)$, for a word $w$ in the generators given in (1.2) of length $|w|$.

Proof. Let $K = \max\{T(\nu_1), \cdots, T(\nu_4)\}$. Notice that $K = O(g)$. Compute each $w(\nu_i)$, $i = 1, \cdots, 4$ by applying generators iteratively. At each step, the total measure grows by a factor of at most 3. Therefore the total complexity is

$$O(g \log(gK) + g \log(3gK) + \cdots + g \log(3^{|w|^{-1}gK})) = O(|w|^2 g + |w| g \log g).$$

Now check if $w(\nu_i) = \nu_i$. This takes $O(|w| g \log g)$. This shows that the word problem in $M_1^g$ has complexity $O(|w|^2 g + |w| g \log g)$. ♠
Conjecture 2.6. The bound given in Theorem 2.5 is in fact optimal.

§3. THE COMPLEXITY OF THE WORD PROBLEM IN BRAID GROUPS

To study the complexity of the word problem in the Braid group \( B_n \), \( n \geq 3 \), we can use similar methods as before. First we study the mapping class group \( M_{n+1}^{n+1} \) of the \((n+1)\)-punctured sphere \( S_{n+1}^{n+1} \). Let’s call the punctures \( P_0, ..., P_n \). Because of the nature of braid groups, we have to allow mapping classes to permute the punctures \( P_1, ..., P_n \) but keep \( P_0 \) fixed. Let’s call this extended group \( \tilde{M}_{n+1}^{n+1} \). Then we have an exact sequence

\[
1 \to M_{n+1}^{n+1} \to \tilde{M}_{n+1}^{n+1} \to S_n \to 1,
\]

where \( S_n \) is the symmetric group on \( n \) elements.

We can use the fundamental polygon \( R = (a_1, a_1^{-1}, ..., a_n, a_n^{-1}) \) to represent \( S = S_{n+1}^{n+1} \). Let’s assume that \( P_i \) is the vertex shared by \( a_i, a_i^{-1} \). We can look at the space of measured train-tracks on \( S \). As in [HC], one can prove that any measured train-track can be represented uniquely as a measured \( \pi_1 \)-train-track.

To determine the action of \( f \in \tilde{M} = \tilde{M}_{n+1}^{n+1} \) on a measured \( \pi_1 \)-train-track \( \nu = (\tau, \mu) \) one has to also specify a permutation \( \sigma \in S_n \). The group \( \tilde{M} \) is generated by \( n - 1 \) half-twists \( H_i \) along the curves \( \gamma_i = a_i a_{i+1} \) for \( i = 1, ..., n - 1 \).

By a half-twist along \( \gamma_i \) we mean the following mapping class, which interchanges \( P_i \) and \( P_{i+1} \), and is obtained by cutting \( S \) along a strip parallel to \( \gamma_i \), rotating the component containing \( P_i, P_{i+1} \) by \( 180^\circ \), and then gluing to the rest of the surface continuously, twisting towards left (we could use twists to right as well, since the situation is completely symmetric).

We will use the set of generators \( H_1, \cdots, H_{n-1} \) as our basic set of generators for \( M \).

3.1 Computation of \( H_i \) on a measured \( \pi_1 \)-train-track.

Now let’s see how one can compute \( H_i(\nu) \) for a given measured \( \pi_1 \)-train-track \( \nu = (\tau, \mu) \) on \( R \). Look at Figure 3.1, where we have a “general” \( \pi_1 \)-train-track. We have shaded the region bounded by \( \gamma_i \) containing \( P_i \) and \( P_{i+1} \). The outcome of \( H_i(\tau) \) is shown in Figure 3.2. To put \( H_i(\nu) \) in \( \pi_1 \)-train-track form, we have to consider different cases, as follows:
Case 1. \( \tau \) and \( \gamma_i \) do not intersect. To get \( H_i(\nu) \), we just have to change the branches according to the rotation of the hexagon bounded by \( a_i^{\pm 1}, a_{i+1}^{\pm 1} \) and \( \gamma_i \) by \( 180^\circ \). Namely, \( a_i^{\pm 1} \to a_{i+1}^{\pm 1} \) and \( a_{i+1}^{\pm 1} \to a_i^{\pm 1} \). This can be done by searching through a list of length \( O(g) \) and replacing numbers of order \( T(\nu) \).

Case 2. \( \mu(a_{i+1}^{-1}, a_k^{\pm 1}) \neq 0 \) only possibly for \( k = i, i + 1 \). In this case \( H_i(\tau) \) is collapsible to a \( \pi_1 \)-train-track. Therefore \( H(\nu) \) can be computed by \( O(n) \) additions of numbers \( \leq T(\nu) \).

Case 3. Otherwise. In this case there are going to be bad curves, i.e., curves going from \( a_i^{-1} \) to \( a_i^{-1} \). By reducing the bad curves one can see that after a split the resulting train-track will be collapsible to a \( \pi_1 \)-train-track. Again the number of operations needed to obtain the answer is \( O(n) \), and the numbers involved are \( O(T(\nu)) \).

This finishes the computation. One can observe that this computation is much less detailed that the corresponding one in \( M_0^1 \). Let’s summarize the above discussions in the
following Theorem:

**Theorem 3.2.** Let $\nu = (\tau, \mu)$ be a measured $\pi_1$-train-track on the standard fundamental domain $R$ for $S_0^{n+1}$ with $T(\nu) = \ell$. Let $H_i$ be one of the standard generators of $\tilde{M}_0^{n+1}$. Then one can compute $H_i(\nu)$ as a measured $\pi_1$-train-track with complexity $O(n(\log(n\ell)))$. Moreover, $T(H_i(\nu)) \leq 2\ell$.

The following is similar to Theorem 2.4.

**Theorem 3.3.** There are 3 integral measured $\pi_1$-train-tracks $\nu_i$, $i = 1, 2, 3$ on $S_0^{n+1}$, $n \geq 3$, such that for $f \in \tilde{M}_0^{n+1}$, the following condition implies $f = id$.

\[ (*) \quad f(\nu_i) = \nu_i \text{ for } i = 1, 2, 3. \]

**Proof.** The $(n + 1)$-punctured sphere can be divided up into “pairs of pants” using the
simple closed curves $\alpha_1, \cdots, \alpha_{n-2}$. See Figure 3.3. If $f$ fixes $\gamma_1, \cdots, \gamma_{n-2}$ and $\alpha_1, \cdots, \alpha_{n-2}$ then it has to fix all the punctures. This is easy to see when $n \geq 4$. If $n = 3$, i.e., there are 4 punctures, then use the fact that $P_0$ is fixed by all mapping classes $f \in \tilde{M}_0^4$. It follows that $f$ must be a product of twists in $\alpha_i$. If $f$ fixes $\gamma_i, i = 1, \cdots, n-2$, then $f$ can not have a twist in $\alpha_i$, so $f = \text{id}$. Now let $\nu_1$ be the measured train-track obtained by $\{\alpha_1, \cdots, \alpha_{n-2}\}$; $\nu_2$ be obtained by $\{\gamma_1, \gamma_3, \cdots\}$ and $\nu_3$ be obtained by $\{\gamma_2, \gamma_4, \cdots\}$. Now if $f$ fixes $\nu_1, \nu_2$ and $\nu_3$ then it fixes all $P_i, \gamma_i, \alpha_i$. Therefore $f = \text{id}$. ♠

**Theorem 3.4.** The word problem in $\tilde{M}_0^{n+1}$ has complexity $O(n|w|^2 + |w|n \log n)$, for a word $w$ in $\{H_1, \cdots, H_{n-1}\}$ of length $|w|$.

**Proof.** Let $K = \max\{T(\nu_1), T(\nu_2), T(\nu_3)\}$. Notice that $K = O(n)$. Compute each $w(\nu_i), i = 1, 2, 3$. Each has complexity

$$O(n \log(nK) + n \log(n2K) + \cdots + n \log(n2^{|w|-1}K)) = O(n|w|^2 + |w|n \log n).$$

Now check if $w(\nu_i) = \nu_i$. This takes $O(n \log n|w|)$. This shows that the word problem in $\tilde{M}_0^{n+1}$ has complexity $O(n|w|^2 + n \log n|w|)$. ♠

Now we turn to the word problem in the braid groups. The $n$-braid group $B_n$ is given by the mapping class group of an $n$-punctured disk, with the possibility of permuting punctures.
Notice that
\[ 1 \to \mathbb{Z} \to B_n \to \tilde{M}_0^{n+1} \to 1. \]

Also, $B_n$ has the Artin presentation
\[ B_n = \langle \sigma_1, \ldots, \sigma_{n-1} \mid \sigma_i \sigma_j = \sigma_j \sigma_i, \ |i-j| \geq 2, \]
\[ \sigma_i \sigma_j \sigma_i = \sigma_j \sigma_i \sigma_j, \ |i-j| = 1 \rangle. \]

It is easily seen geometrically that $\sigma_i \to H_i$ in the natural projection $B_n \to \tilde{M}_0^{n+1}$. Therefore given a word $w$ in of length $|w|$ one can check if the image of $w$ is the identity in $\tilde{M}_0^{n+1}$ with complexity $O(n|w|^2 + |w|n \log n)$. To solve the word problem in $B_n$, we have to only check the following: For a word $w \in \ker(B_n \to \tilde{M}_0^{n+1})$, is $w = \text{id}$? Geometrically, this means that the given word is a twist around the boundary of the disk; i.e., a power of $\Delta$, where $\Delta$ is the generator of the center of $B_n$. We need to know if this power is 0. For this let’s take a look at the fundamental domain $R$. and the arc $\beta$ connecting $P_i$ to a point in the boundary of the disk as in Figure 3.4. We can find the action of $w$ on $\beta$. It’s natural to encode the arc $\beta$ as a measured $\pi_1$-train-track with dead-ends.

![Diagram](image-url)
This goes as in the case of measured \( \pi_1 \)-train-tracks, but one has to keep a neighborhood of the both ends of \( \beta \) fixed while applying each generator. The details are similar to the case of \( \pi_1 \)-train-tracks. In particular, this has complexity \( O(n|w|^2 + n \log n|w|) \). This implies

**Theorem 3.5.** The word problem in the Braid group \( B_n \) has complexity \( O(n|w|^2 + |w|n \log n) \), where \( |w| \) is the length of the word \( w \) in the Artin generators \( \{\sigma_1, \ldots, \sigma_{n-1}\} \).

**Corollary 3.6.** If \( w \) is a word in the Artin generators of \( B_n \) of length \( |w| \), with \( |w| \geq \log n \), one can determine if \( w = id \) with complexity \( O(|w|^2 n) \) on a Turing Machine.

Using similar ideas with a standard fundamental domain for the surface \( S^p_g \) with \( p \geq 2 \) incorporating the cases of once-punctured surfaces and braid groups one can similarly prove:

**Theorem 3.7.** The word problem in \( M^p_g \) has complexity \( O(|w|^2(g+p) + |w|(g+p) \log(g+p)) \), for a word \( w \) in a set of “standard” generators of length \( |w| \).

§4. THE CASE OF A CLOSED SURFACE

Now let’s discuss the case of a closed surface, i.e., when \( S = S^0_g = S_g \). The basic group structure of \( M_g \) in terms of \( M^1_g \) is given by the short exact sequence (see [B])

\[
1 \to \pi_1(S,*) \to M^1_g \overset{\phi}{\to} M_g \to 1.
\]

Here the canonical map \( \phi \) is defined by just forgetting the puncture.

Notice that by this exact sequence the generators in (1.2) can be naturally considered as generators of \( M^0_g \).

Let’s introduce an artificial puncture on \( S \); i.e., let’s fix a point \( P \) on \( S \), and call the corresponding once-punctured surface \( S^P \). Also let \( R \) be the standard fundamental domain for \( S \), having all vertices equivalent to \( P \) on \( S \). Suppose a simple closed curve \( \alpha \) is given on the closed surface \( S \), and \( \alpha \) does not pass through \( P \). The curve \( \alpha \) can be considered as a curve on the punctured surface \( S^P \), and can be given by a cyclic word \( w = e_{s_1} \ldots e_{s_n} \) where \( e_i \in E(R) \), as in 1.8. Notice that by isotoping \( \alpha \) in \( S \), we might obtain a shorter cyclic word.

We want to discuss here a geometric analog of Dehn’s well-known algorithm (see [J], e.g.)
to get a shortest representative for \( \alpha \). The shortest representative is not unique, as we will see below.

Represent \( w \) by a measured \( \pi_1 \)-train-track \( \nu \) carried precisely on a \( \pi_1 \)-train-track \( \tau \). Let \( m = |E(R)|/2 = 2g \). If there is no path \( b = (b_1, ..., b_k) \) of outer branches in \( \tau \) such that \( k \geq m \) then we claim that \( w \) is a shortest representative for \( \alpha \). Recall that in this case \( \alpha \) is given by a word \( w = e_{s_1}...e_{s_n} \) in letters in \( E(R) \) representing a simple closed curve \( \alpha_0 \) on \( S^P \). We can assume \( w \) does not have back-tracking; i.e, \( e_{s_i} \neq e_{s_{i+1}}^{-1} \) for all \( i \) (mod \( n \)). If \( w \) is not a shortest representative for \( \alpha \), then there is another word \( w' = e_{s'_1}...e_{s'_{n'}} \) with \( n' < n \) representing a curve \( \alpha_1 \) in \( S^P \) which is isotopic to \( \alpha \) in \( S \). Take an isotopy \( \alpha_t \) between \( \alpha_0 \) and \( \alpha_1 \) on \( S \). By changing \( \alpha_t \) a little bit, one can subdivide this isotopy to subintervals in which, either (i) no part of \( \alpha \) is in a small neighborhood of \( P \), or (ii) only one segment of \( \alpha \) is passing through \( P \), and everything else is fixed. Notice that in intervals of type (i) the word representing the curve in \( S^P \) does not change since that part of the isotopy can be looked at as an isotopy of \( S \setminus \{P\} \) and \( \pi_1(S \setminus \{P\}) \) is free. Therefore, one can find a finite sequence \( \alpha_0 = \alpha_{t_0}, \alpha_{t_1}, ..., \alpha_{t_\ell} = \alpha_1 \) which give all the different simple closed curves that appear on \( S^P \). Every element in this sequence is obtained by the previous one by taking a piece of \( \alpha \) and passing it through \( P \). We can assume that \( \alpha_{t_i} \neq \alpha_{t_j} \) for \( i \neq j \), otherwise we can just drop the repeating part of the isotopy. Since by assumption \( \alpha_0 \) does not have a path of outer branches \( b = (b_1, ..., b_k) \) with \( k \geq m \), we must have \( T(\alpha_{t_1}) \geq T(\alpha_{t_0}) \), with equality only in the case in which \( \alpha_{t_0} \) has a path \( b \) as above with \( k = m - 1 \), and the move is to just push the path to the other side of \( P \) (Figure 4.1). Notice that Figure 4.1 is drawn in the universal cover of \( S \).

If the sequence \( \alpha_{t_0}, \alpha_{t_1}, ..., \alpha_{t_\ell} \) only consists of moves which push a path of length \( m - 1 \) across the puncture, then \( T(\alpha_0) = T(\alpha_1) \); i.e., \( n = n' \), which is a contradiction. Otherwise, let \( 1 \leq j \leq \ell \) be such that \( T(\alpha_{t_0}) = ... = T(\alpha_{t_{j-1}}) < T(\alpha_{t_j}) \). If \( j = \ell \) then \( n' > n \) which is a contradiction. Since \( \alpha_{t_{j+1}} \) can not be equal to any of the preceding \( \alpha_{t_i} \), it is easy to see that \( T \) keeps monotonically increasing on the sequence \( \alpha_{t_0}, \alpha_{t_1}, ..., \alpha_{t_\ell} \). This shows that \( n' > n \), which is again a contradiction.

Let’s summarize the above arguments in the following Theorem. A subword of a word \( e_1 \cdots e_n \) is any word of the form \( e_i e_{i+1} \cdots e_j \).
**Theorem 4.1.** Let $R$ be a standard fundamental domain for the closed surface $S = S_g$ with $2m$ edges ($m = 2g$), with the vertices of $R$ equivalent to a point $P$ on $S$. Let $S^p$ be a once-punctured surface obtained by fixing $P$ on $S$. Let $\alpha$ be a simple closed curve on $S$ not passing through $P$, and let $w = e_{s_1}...e_{s_n}$ be a cyclic word in letters in $E(R)$ representing $\alpha$ up to isotopy in $S$. Then $w$ is a shortest representative if and only if

1. $e_{s_i} \neq e_{s_{i+1}}^{-1}$ for all $i \mod n$, and
2. $w$ does not have a subword of length $\geq m$ consisting of outer branches.

Moreover, any two shortest length representatives of $\alpha$ are related to each other by pushing a finite number of identical subwords of length $m$ of the outer branches to the other side of $P$. ♠

With the same assumptions on the fundamental domain $R$, let $\nu$ be a measured train-track carried precisely on a $\pi_1$-train-track $\tau$ on $S$. As we know by now from simple closed curves, the $\pi_1$-train-track representative is not unique in $S$. We want to describe an algorithm to put $\nu$ in a $\pi_1$-train-track form which has the smallest $T$. We call $\tau$ a **reduced-length $\pi_1$-train-track** if it has no path of outer branches of length $\geq m$.

**Lemma 4.2.** If $\nu = (\tau, \mu)$ is a measured $\pi_1$-train-track on $S$, there exists a measured $\pi_1$-train-track $\nu' = (\tau', \mu')$ which represents $\nu$ and it has the smallest possible $T$. 
Proof. Let \( \{c_n\} \) be a sequence of simple closed curves on \( S \) and \( \lambda_n > 0 \) be such that \( \lambda_n c_n \to \nu \) as \( n \to \infty \). Put each \( c_n \) in a reduced form \( \tilde{c}_n \). By passing to a subsequence we can assume all the \( \tilde{c}_n \) are carried on a reduced-length \( \pi_1 \)-train-track \( \tau' \). Now one can look at the sequence \( \{[\tilde{c}_n]\} \) in \( \mathcal{PMT}(S^P) \). By compactness, this sequence has a convergent subsequence. Without loss of generality, let’s assume \( [\tilde{c}_n] \to \nu' \in \mathcal{PMT}(S^P) \). Notice that \( \nu' \) is of reduced length since it is carried on \( \tau' \). Using the surjection \( \mathcal{PMT}(S^P) \to \mathcal{PMT}(S) \), one gets a corresponding convergent sequence \( [c_n] \to \nu' \) in \( \mathcal{PMT}(S) \). We denote the limit point with the same notation since it is given by the same measured \( \pi_1 \)-train-track. This shows that \( \nu = \nu' \) i.e., \( \nu \) is equivalent to a reduced-length measured \( \pi_1 \)-train-track.

Now we have to prove that \( T(\nu') \) is minimal among all \( T(\nu'') \), where \( \nu'' \) is a measured \( \pi_1 \)-train-track representative for \( \nu \). Suppose \( T(\nu'') < T(\nu') \), for such a \( \nu'' \). Then by definition of the space of measured train-tracks, there is a finite sequence

\[
\nu' = \nu_1 \to \nu_2 \to \ldots \to \nu_k \to \ldots \to \nu_n = \nu''
\]

where each \( \nu_j \) is obtained by performing one of the following moves on \( \nu_{j-1} \): (i) Split, (ii) Shift, (iii) Isotopy without crossing \( P \), (iv) Pulling a branch from one side to the other side of \( P \), and (v) Collapse. It is easily seen that one can arrange the sequence (4.1) so that \( \nu_1, \ldots, \nu_k \) are obtained by performing the moves of type (i)-(iv), and the rest of \( \nu_j \) are obtained only using the collapse move. Choose a simple closed curve \( c' \) and \( \lambda > 0 \) such that \( c' \) is carried on \( \tau' \), and it stays \( \epsilon \)-close to \( \nu_i \) at each step along the sequence \( \nu_1 \to \ldots \to \nu_k \), as we perform the corresponding move on \( \lambda c' \), where \( \epsilon > 0 \) is an arbitrary pre-chosen number. Here \( \epsilon \)-close is used in the sense that at each stage, the sum of the differences the measures in corresponding branches is bounded above by \( \epsilon \). In particular, \( |T(\nu') - T(\lambda c')| < \epsilon \). After collapsing to \( \nu'' \), we get a (measured) simple closed curve \( \lambda c'' \) which is \( \epsilon \)-close to \( \nu'' \). In particular, \( |T(\nu'') - T(\lambda c'')| < \epsilon \). If we choose \( 2\epsilon < T(\nu'') - T(\nu') \), we get \( T(\lambda c'') < T(\lambda c') \), which contradicts Theorem 4.1, since \( c' \) is carried on a reduced-length measured \( \pi_1 \)-train-track. This finishes the proof of the lemma. ♠

Corollary 4.3. If \( \nu = (\tau, \mu) \) is a measured \( \pi_1 \)-train-track on the closed surface \( S \) carried on a reduced-length measured \( \pi_1 \)-train-track \( \tau \), then \( T(\nu) \) is minimal among \( T \) of all other measured \( \pi_1 \)-train-track representatives of \( \nu \).
A similar limit argument as in the proof of the lemma shows that:

**Corollary 4.4.** Any two reduced-length measured $\pi_1$ train-track representatives of the same measured train-track on the closed surface $S$ are related by the following move: Pulling some measure off a path of outer branches of length $m - 1$, where $|E(R)| = 2m$, to the other side of the puncture $P$.

Here is an algorithm to put a given measured $\pi_1$-train-track in the reduced (shortest) form. Let’s start with a measured $\pi_1$-train-track $\nu_1 = (\tau_1, \mu_1)$ which is not reduced-length. So there is a unique maximal path $b = (b_1, ..., b_k)$ of outer branches in $\tau_1$ where $k \geq m$. 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 4.2 to put $\tau_1$ in a position with smaller $T$. To be able to do that move, we have to assume $x_i = \min\{x_1, ..., x_k\}$.

We claim that, after doing the move finitely many times, we will get a sequence of measured $\pi_1$-train-tracks $\nu_1, ..., \nu_t$ where $\nu_i = (\tau_i, \mu_i)$, and $\tau_t$ is of reduced-length. The reason is that first of all we know that there is a sequence of moves of type (i)-(v) putting $\nu_1$ is reduced form. Now notice that as in the case of simple closed curves, If you make a move and increase $T$, to reduce $T$ later on you have to undo the move. This proves that there is a sequence to monotonically decrease $T$, which proves our assertion, since at any
given stage, there is only one way to reduce the $T$, if the train-track is not already in the reduced-length position.

The analog of Theorem 1.5 is

**Theorem 4.5.** Let $S = S_g$ where $g \geq 2$, and let $R$ be an standard fundamental domain for the action of $\pi_1(S, \ast)$ on $\mathbb{H}^2$. Then every measured train-track is equivalent to some measured $\pi_1$-train-track $\nu = (\tau, \mu)$ with respect to $R$ having the smallest possible $T$. This representative is unique if and only if $\tau$ has no path of outer branches of length $|E(R)|/2 - 1$. Otherwise any representative is obtained from any other representative by pulling some measure from a path of outer branches of length $|E(R)|/2 - 1$ to the other side of the puncture. ♠

§5. **The complexity of the word problem in the mapping class groups of closed surfaces**

Since the $\pi_1$-train-track representation is not unique for closed surfaces, the main issue here is the following problem:

**5.1 Problem.** Find the complexity of the following computation: Given an integral measured $\pi_1$-train-track $\nu = (\tau, \mu)$ on the standard fundamental domain $R$ for the surface $S = S_g$ with $T(\nu) = \ell$, compute a $\nu' = (\tau', \mu')$ of reduced form such that $\nu'$ is equivalent to $\nu$ on $S$.

Recall that $m = 2g = |E(R)|/2$. It is easy to check if $\nu$ is not of reduced length with complexity $O(g)$. One has to check if there is a path of outer branches of length $\geq m$. Therefore suppose $\nu$ is not of reduced length, to start with. Let $b = (b_1, \ldots, b_{n(\nu)})$ be the unique maximal path of outer branches in $\tau$ of length $n(\nu) \geq m$, and let $\psi(\nu) = \min\{\mu(b_1), \cdot \cdot \cdot, \mu(b_n)\}$.

We use $n(\nu)$ as a measure of complexity. Notice that $n(\nu) \leq |\text{out}(\tau)|$ (Recall that $\text{out}(\tau)$ is the set of outer branches of $\tau$). We will put $\nu$ in the reduced-length form by a sequence of moves each of which reduces the complexity function $n(.)$. Notice that $\nu$ is of reduced form if $n(\nu) \leq 2g - 1$. Moreover, it is always possible to reduce $\nu$ such that $|\text{out}(\tau)| \leq 4g - 3$, as we will see below.
Case 1. \( n(\nu) < 4g - 1 \). Let \( x = \mu(b_i) = \psi(\nu) \). We can pull a measure of \( x \) to the other side of the puncture. This may involve changing some inner branches which connect to the both ends of the path \( b \) to outer ones. In particular, this may add a measure of \( x \) to at most two of the branches in \( b \). If none of these branches are \( b_i \), then we have reduced the complexity function, because one can easily see that the added outer branches can not extend \( b \) from either side. Now let’s consider the case which pulling the measure adds to \( b_i \), so that after the pulling, we still have \( \mu(b_i) = x \). This subtracts \( x \) from all the branches of \( b \) except for \( b_i \) and possibly another branch \( b_j \). By examining the size of the measures \( \mu(b_k), 1 \leq k \leq n(\nu) \) and the ones connecting to the endpoints of \( b \), we can see how many times this move is possible, and we can do them all at once. After we do that, there is a \( k \neq i, j \) such that \( \mu(\beta_k) \) has become \( < x \), which means \( \psi(\nu) \) is now \( < x \). Now pull this measure across the puncture, and this will reduce the complexity function. This shows that one can put \( \nu \) in reduced-length form after \( O(g) \) steps. Each step involves \( O(g) \) operations on numbers which are \( O(T(\nu)) \). Therefore, the complexity of putting \( \nu \) in reduced-form in this case is \( O(g^2 \log T(\nu)) \). If at the end the final \( \nu \) satisfies \( n(\nu) = 2g - 1 \), then one can easily force \( \lvert \text{out}(\tau) \rvert \leq 4g - 3 \): If \( n(\nu) = 2g - 1 \) and \( \lvert \text{out}(\tau) \rvert = 4g - 2 \), by finding the outer branch with smallest measure and pulling that measure through the puncture in a similar fashion as above, we get \( \lvert \text{out}(\tau) \rvert \leq 4g - 3 \).

Case 2. \( n(\nu) = 4g - 1 \). (Equivalently, \( \lvert \text{out}(\tau) \rvert = 4g - 1 \).) In this case the complexity of the problem can be much higher, in fact it will be of linear order with respect to \( T(\nu) \). The problem is that one can pull a small piece of the curve \( \nu \) around arbitrarily long and then hook it up with the puncture. Then to simplify the curve one has to undo that, which has complexity \( O(T(\nu)) \). See Figure 5.1.

5.2. Solution to the word problem. In the solution to the word problem in \( M^0_g \) we have to avoid Case 2 in 5.1, because it will have an effect of making it exponential, since our polynomial algorithms are all based on the fact that the computations with a curve are of order \( \log N \), if the size of the curve at hand is \( N \).

Here is our strategy for the solution of the word problem in \( M^0_g \): Let \( w = h_1 \cdots h_n \) be a word in the basic set of generators of \( M^0_g \) (see (1.2) and the note below it). similar to
Theorem 2.4. We know that there are 4 measured $\pi_1$-train-tracks $\nu_1, \ldots, \nu_4$ with $T(\nu_i) = O(g)$ on $S_g$ such that if $w(\nu_i) = \nu_i$ for all $1 \leq i \leq 4$ then $w = \text{id}$. (This holds only for $g \geq 3$; in $M_2$ there is a mapping class of order 2 fixing all simple closed curves). Put $\nu_i^{(0)} = \nu_i$ and $\nu_i^{(j+1)} = h_{n-j}(\nu_i^{(j)})$. Notice that $\nu_i^{(n)} = w(\nu_i)$. For $j = 0, \ldots, n$, we compute $\nu_i^{(j)}$. After each computation, we put $\nu_i^{(j)}$ in the reduced-length form. What we would like to show is that, if $h$ is a generator and $\nu$ is of reduced length, $h(\nu)$ can be put into reduced-length form with complexity $O(\log T(\nu))$ with respect to $T(\nu)$. For that we have to again look closely how each of the generators act on a reduced-length measured-train-track $\nu = (\tau, \mu)$. By the above argument in Case 1, it is enough to show that $n(h(\nu)) < 4g - 1$.

**Lemma 5.3.** Suppose $h^{\pm 1}$ is one of the generators in (1.2), and $\nu = (\tau, \mu)$ is an integral measured $\pi_1$-train-track on the standard fundamental domain for $S_g$, $g \geq 2$ of reduced-length. Put $h(\nu) = (\tau_1, \mu_1)$. Then $n(\tau_1) < 4g - 1$, or equivalently $|\text{out}(\tau_1)| < 4g - 1$.

**Proof.** We will only discuss the cases which $h$ is a generator in (1.2). The cases where $h^{-1}$ is a generator are done by symmetry.

**Case 1.** $h = D_{a_i}$. Let $\nu = (\tau, \mu)$ be a reduced-length measured $\pi_1$-train-track with $n(\nu) \leq 2g - 1$ and $|\text{out}(\tau)| \leq 4g - 3$ (see the argument in Case 1 in 5.1 above). We claim that $n(\nu_1) < 4g - 1$. The proof has many steps.

(i) $\mu(a_i^{-1}, b_i) = 0$. (No bad curves) Notice that $\mu_1(a_i^{-1}, b_i) = 0$. If $g > 2$, at least one of
\[\mu(a_{t+1}, b_{t+1}), \ldots, \mu(a_{t-1}, b_{t-1})\] must be 0, which stays 0 with \(\mu_1\) instead of \(\mu\). This shows that \(n(\nu_1) < 4g - 1\). Suppose \(g = 2\) and, say \(t = 1\). Since \(\nu\) is of reduced-length, one of the values

\[\mu(b_1^{-1}, a_2), \mu(a_2, b_2), \mu(b_2, a_2^{-1}), \mu(a_2^{-1}, b_2^{-1})\]

must be 0 and stays 0 if we replace \(\mu\) by \(\mu_1\). Therefore the estimate \(n(\nu_1) < 4g - 1\) holds in this case too.

(ii) \(\mu(a_t^{-1}, b_t) \neq 0\) but \(\mu(\alpha_t, b_t) = 0\). Then again \(\mu_1(a_t^{-1}, b_t) = 0\) and the argument is similar to (i).

(iii) \(\mu(a_t^{-1}, b_t) \neq 0\) and \(\mu(\alpha_t, b_t) \neq 0\). In this case \(\text{out}(\tau) = \text{out}(\tau_1)\) unless \(\mu(b_t, b_t^{-1}) \neq 0\) and \(\mu(a_t^{-1}, b_t^{-1}) = 0\), in which case \(\text{out}(\tau_1) = \text{out}(\tau) \cup \{(a_t^{-1}, b_t^{-1})\}\). Since \(n(\tau) < 4g - 2\), \(n(\tau_1) < 4g - 1\).

**Case 2.** \(h = D_{b_t}\). This case is similar to case 1.

**Case 3.** \(h = D_{x_t}\).

(i) No bad curves. This means that \(\mu(b_t, e) = 0\) for \(e \in E(R) \setminus \{a_t^{-1}, b_t^{-1}, a_{t+1}, b_{t+1}, a_{t+1}^{-1}\}\). If \(x_t\) and \(\tau\) do not intersect, then \(\tau_1 = \tau\), and we are done. If \(\mu(b_t, a_{t+1}^{-1}) \neq 0\), then \(x_t\) and \(\tau\) intersect only when \(\mu(a_{t+1}^{-1}, e) \neq 0\) for some \(e \in E \setminus \{b_t, a_t^{-1}, b_t^{-1}, a_{t+1}, b_{t+1}, a_{t+1}^{-1}\}\). In that case, \(\text{out}(\tau_1) = \text{out}(\tau) \cup \{(b_t^{-1}, a_{t+1})\}\), therefore \(|\text{out}(\tau_1)| \leq 4g - 2\). So suppose \(\mu(b_t, a_{t+1}^{-1}) = 0\) as well. Applying \(h\) may create new outer branches only of one of the following types:

\[(b_t, a_t^{-1}), (b_t^{-1}, a_{t+1}), (a_{t+1}^{-1}, b_{t+1}^{-1})\]

Since \(\mu(a_t, b_t) = 0\), we have \(\mu_1(a_t, b_t) = 0\). If any of \(\mu(a_t^{-1}, b_t^{-1}), \mu(a_{t+1}, b_{t+1}), \mu(b_{t+1}, a_{t+1}^{-1})\) are 0, then they will be 0 with \(\mu_1\) instead of \(\mu\) and we are done. So let’s assume they are all non-zero. Let’s look at the case \(g \geq 4\), since the argument is easiest in this case. Because \(\tau\) is of reduced-length, one of the outer branches which does not intersect any of the simple closed curves \(x_t, a_t, b_t, a_{t+1}, b_{t+1}\) must have zero measure, and this is going to stay zero in \(\mu_1\). This gives \(n(\tau_1) < 4g - 1\). Now let’s look at the case \(g = 3\), and without loss of generality assume \(t = 1\). If \(\mu(a_2^{-1}, b_2^{-1}) \neq 0\), then again one of the same type of outer branches must have 0 measure, and again we are done. Therefore assume \(\mu(a_2^{-1}, b_2^{-1}) = 0\).

The assumptions force

\[\text{out}(\tau) = E(R) \setminus \{(a_1, b_1), (b_1^{-1}, a_2), (a_2^{-1}, b_2^{-1})\}\]
but this is not a reduced-length train-track. This takes care of the case \( g = 3 \). Now look at the case \( g = 2 \). Similar to the case of \( g = 3 \), it follows that \( \mu(a_1, b_2^{-1}) = \mu_1(a_1, b_2^{-1}) = 0 \), and we are done.

(ii) There are bad curves but \( \mu(a_t, b_t) = 0 \). The existence of bad curves means that \( \mu(b_t, e) \neq 0 \) for some \( e \in E(R) \setminus \{a_t^{-1}, b_t^{-1}, a_{t+1}, b_{t+1}, a_{t+1}^{-1}\} \). In this case the train-track obtained by pushing the bad curves across \( \partial R \) is collapsible to \( \tau_1 \). Notice that \( (a_t, b_t) \notin \text{out}(\tau_1) \) and \( \text{out}(\tau_1) \setminus \text{out}(\tau) \) may only contain \( (b_t^{-1}, a_{t+1}), (a_{t+1}, b_{t+1}), (a_{t+1}^{-1}, b_{t+1}^{-1}) \). Therefore as in (i), if \( g \geq 4 \) we are done. If \( g = 2 \) or \( 3 \) and say \( t = 1 \), then one of the branches in

\[
E(R) \setminus \{(a_1, b_1), (b_1^{-1}, a_2), (a_2, b_2), (a_2^{-1}, b_2^{-1})\}
\]

must be missed by \( \text{out}(\tau) \) (since \( \tau \) is of reduced-length) and it will be missed by \( \text{out}(\tau_1) \) as well.

(iii) There are bad curves and \( \mu(a_t, b_t) \neq 0 \). In this case after pushing the bad curves, we still have to push some “bad pairs” which come out near the edge \( a_{t+1}^{-1} \). In this case

\[
\text{out}(\tau_1) \setminus \text{out}(\tau) \subseteq \{b_t^{-1}, a_{t+1}^{-1}\}.
\]

Since \( |\text{out}(\tau)| \leq 4g - 3 \), \( |\text{out}(\tau_1)| \leq 4g - 2 \) and we are done. ♠

**Theorem 5.4.** The complexity of the word problem in \( M_g \) is \( O(|w|^2g^2g^2 + |w|g^2 \log g) \), where \( |w| \) is the word length in the set of generators (1.2). In particular, for \( |w| \geq \log g \), the word problem has complexity \( O(|w|^2g^2) \).

**Proof.** Since the word problem in \( M_2 \) is quadratic in the word length, we need to prove the theorem for \( g \geq 3 \). (This is because there are mapping classes in \( M_2 \) which fix all simple closed curves but are not the identity element, so our methods purely do not solve the word problem). Put the analog of each \( \nu_i, i = 1, \ldots, 4 \) given in Theorem 2.4 in a reduced-length form. This takes \( O(g) \) since \( T(\nu_i) = O(g) \). Given the word \( w = h_1 \cdots h_{|w|} \), apply each generator on the \( \nu_i, i = 1, \cdots, 4 \). After each application put the resulting measured train-track in a reduced-length form. This takes \( O(g^2 \log(\text{size})) \). But the size grows by at most a factor of 3, therefore the total complexity is

\[
O(g^2 \log(g) + g^2 \log(3g) + \cdots + g^2 \log(3|w|^{-1}g))
\]
which is \( O(|w|^2 g^2 + |w| g^2 \log g) \). ♠

APPENDIX: TURING MACHINE AND COMPUTATIONAL COMPLEXITY

A Turing Machine (see [Br] or [S], for example) is a hypothetical machine consisting of an infinitely long tape, a read/write head connected to a control mechanism. The tape is divided into infinitely many cells, each of which contains a symbol from a finite alphabet (the alphabet contains a special symbol for blank cell). The cells are scanned one at a time using the read/write head, which can write a new symbol on the cell just read, move in either direction or not move at all. At any given time, the machine is in one of the finitely many internal states. The behavior of

the machine and a possible change of state depends on the current state, and the symbol read from the tape.

Formally, let \( X \subset Y \) be finite alphabets. A Turing Machine is a quadruple \( (Q, \delta, q_0, q_F) \) where \( Q \) is a finite set of states, \( \delta \) is a function defined on a subset of \( Q \times Y \) to \( Q \times Y \times \{L, R, 0\} \) which is the state transition function, \( q_0 \in Q \) is the start state, and \( q_F \in Q \) is the halt state. The symbols \( L, R, 0 \) should be interpreted as moving the head to the left, right, or no move at all, respectively. The set \( X \) is the input alphabet.

Intuitively, any problem which is solvable by a finite instruction set is solvable by a Turing Machine (see Church’s Thesis say in [Br]). Therefore, we only describe a ”program” for our solutions.

To define the complexity of an algorithm, there isn’t a unique way. We have chosen the complexity to be the number of steps the Turing Machine takes to come up with the answer.

To compute an upper bound for the complexity of a problem, we add up the number of steps needed for each sub-problem. They are all computed according to the following idea: To input a number of size \( N \) into the machine takes \( \log N \) steps. The reason is one can write it in base 2, with \( O(\log_2 N) \) digits. Also, to add two numbers of size \( \leq N \) takes \( \log N \) steps as well. Now one can devise a Turing machine to add the numbers in \( O(\log N) \) steps which we leave as an exercise.
From a theoretical point of view this definition (or any equivalent one with respect to complexity) seems appropriate since a Turing Machine is in a sense the most basic computer. In a Random Access Memory machine (say a typical PC), one assumes that it takes a constant time to add any two numbers. This assumption seems reasonable only when using machine-size numbers.

REFERENCES

[B] J. S. Birman, Braids, Links, and mapping class groups, Princeton Univ. Press, 1974.

[BKL] J. S. Birman, K. H. Ko, and J. S. Lee, A new approach to the word problem and conjugacy problems in the braid groups, To appear in Adv. Math. Available at xxx.lanl.gov as math.GT/9712211.

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

[Br] D. S. Bridges, Computability, a mathematical sketchbook, GTM 146, Springer-Verlag, 1994.

[E] Epstein et al., Word processing in groups, Jones and Bartlett Publishers, 1992.

[FLP] A. Fathi, F. Laundenbach and V. Poenaru, Travaux de Thurston sur les surfaces, Astérisque 66-67, Soc. Mathématique de France, 1979.

[G] F. A. Garside, The braid group and other groups, Oxford Quart. J. of Math., 20 (1969), 235-254.

[HC] H. Hamidi-Tehrani and Z.-H. Chen, Surface diffeomorphisms via train-tracks, Topology and its Applications, 73 (1996), 141-167.

[He] G. Hemion, On classification of knots and 3-dimensional spaces, Oxford University Press, New York, 1992.

[Hu] S. Humphries, Generators for the mapping class group, in Lecture notes in Mathematics 722 (Springer, Berlin) 44-47.

[J] D. L. Johnson, Topics in the theory of group presentations, London Math. Soc. Lecture note Ser. 42 (Cambridge Univ. Press, 1980).

[KM] R. Kleinberg and W. Menasco, Train tracks and zipping sequences for pseudo-Anosov
braids, preprint. Available at http://math.Cornell.edu/~rdk.

[MM1] H. Masur and Y. Minsky, Geometry of the complex of curves I: Hyperbolicity. Preprint. Available at http://www.math.sunysb.edu/~yair.

[MM2] H. Masur and Y. Minsky, Geometry of the complex of curves II: hierarchal Structure. Preprint. Available at http://www.math.sunysb.edu/~yair.

[Mo1] L. Mosher, Classification of pseudo-Anosovs, in Low dimensional topology and Kleinian groups, (1986) 13-75.

[Mo2] L. Mosher, Mapping class groups are automatic, Ann. of Math. 142 (1995) No. 2, 303-384.

[Mo3] L. Mosher, A user’s guide to the mapping class group: Once punctured surfaces, In Geometric and computational perspectives on infinite groups, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., 25 (1996) 101-174.

[P] R.C. Penner, The action of the mapping class group on curves in surfaces, L’Enseignement Mathematique 30 (1984), 39-55.

[PH] R. C. Penner with J. L. Harer, Combinatorics of train tracks, Ann. of Math. Stud., vol. 125, Princeton Univ. Press, Princeton, NJ, 1995.

[S] A. Salomaa, Computation and automata, Cambridge University Press, 1985.