Metric distortion of subgroups of mapping class groups

Sebastian Wolfgang Hensel

January 4, 2012

Contents

Chapter 1

Introduction and Summary

The *mapping class group* $\mathrm{Map}(M)$ of a manifold M is the group of diffeomorphisms of M up to isotopy. The most prominent example is the mapping class group of a closed surface S_g of genus $g \geq 2$. This group (often called the *Teichmüller modular group*) appears in many branches of modern mathematics: for example Teichmüller theory, low-dimensional geometry, algebraic geometry and complex analysis. Consequently, $\mathrm{Map}(S_g)$ is studied from many different points of view (see [FM11] or [I02] for good introductions to the subject).

Here, we are interested in groups as geometric objects: a finitely generated group can be equipped with a (coarsely) unique left-invariant metric (the so called *word metric*). In the last decade, tools have been developed to effectively investigate the geometry of $\mathrm{Map}(S_g)$ and some of its natural subgroups (see e.g. [MM00], [Ha09b] and [MMS10]).

Mapping class groups of other manifolds are less well understood – in particular from a geometric point of view. In this thesis we consider two examples which show rich geometric behavior: the mapping class groups of handlebodies and of doubled handlebodies.

1.1 Handlebody groups

A handlebody V_g of genus g is the three-manifold obtained by attaching g three-dimensional one-handles to a three-dimensional ball. The mapping class group of V_g is commonly called *handlebody group*. The handlebody group has been studied from algebraic and dynamical points of view (see e.g.

[Wa98] and [M86]). However, hardly anything is known about the large-scale geometry of $\mathrm{Map}(V_g)$.

One approach to the geometry of $\mathrm{Map}(V_g)$ is via the geometry of surface mapping class groups. Namely, the boundary ∂V_g of V_g is a closed surface of genus g. Restriction of diffeomorphisms of V_g to the boundary surface ∂V_g defines a homomorphism $\iota : \mathrm{Map}(V_g) \to \mathrm{Map}(\partial V_g)$ which is injective. Hence, the handlebody group can be identified with a subgroup of the mapping class group of the closed surface ∂V_g. In Chapter 2 we study the extrinsic geometry of $\mathrm{Map}(V_g)$ as a subgroup of the surface mapping class group $\mathrm{Map}(\partial V_g)$.

By extrinsic geometry we mean the following. Let $H < G$ be a finitely generated subgroup of a finitely generated group G. The inclusion map $H \to G$ is coarsely Lipschitz with respect to the word metrics d_H and d_G on H and G, respectively. In general, however, this inclusion map may distort distances by an arbitrary factor. If the inclusion is in fact a *quasi-isometric embedding*, i.e. if there is a number $K > 0$ with

$$\frac{1}{K} d_G(h, h') - K \leq d_H(h, h') \leq K d_G(h, h') + K \qquad \forall h, h' \in H$$

then we say that the subgroup H is *undistorted* in G. Undistorted subgroups inherit their geometry from the ambient group G. Hence, if the handlebody group was undistorted in the surface mapping class group, its geometry could be easily understood.

However, we show in Chapter 2 that the intrinsic geometry of the handlebody group differs considerably from the geometry as a subgroup of $\mathrm{Map}(\partial V_g)$. The results in this chapter are joint work with Ursula Hamenstädt.

Theorem 1.1.1 ([HH11a])**.** *Let V be a handlebody of genus at least 2. Then the handlebody group of V is exponentially distorted in the mapping class group of ∂V.*

Recall that a finitely generated subgroup $H < G$ of a finitely generated group G is called *exponentially distorted* if the following holds. The word norm of H can be bounded by an exponential function of the word norm of G. On the other hand, there is no such bound of sub-exponential growth type.

The proof of Theorem 1.1.1 consists of two parts — a lower and an upper bound on distortion. To show the upper bound we define *racks*, which are an analog of train tracks for handlebodies. A *large rack* is a system of disks in

V and arcs on ∂V satisfying certain completeness conditions (see Section 2.6 for the formal definition). We define an equivalence relation (*rigid isotopy*) on the set of large racks, such that the handlebody group acts on the set of rigid isotopy classes of racks with finite quotient and finite stabilizers. Thus, one can define a *graph of rigid racks* which serves as a geometric model for the handlebody group. This construction is inspired by the train track graph for surface mapping class groups (see [Ha09b]). Using a surgery procedure, we then define distinguished paths in the graph of rigid racks (*splitting paths*) whose lengths can be bounded in terms of intersection numbers. This yields the upper bound on distortion.

To show the lower distortion bound, we use the geometry of outer automorphism groups $\mathrm{Out}(F_g)$ of free groups. Namely, the handlebody group projects onto $\mathrm{Out}(F_g)$ via its action on the fundamental group of V_g. Using point-pushing homeomorphisms as the main tool we explicitly define a sequence of handlebody group elements φ_n with the following properties. On the one hand, the word norm of φ_n in $\mathrm{Map}(\partial V_g)$ grows linearly in n. On the other hand, the images of φ_n in $\mathrm{Out}(F_g)$ have exponentially growing word norm. As a consequence, the word norm of φ_n in $\mathrm{Map}(V_g)$ also grows exponentially in n, showing the desired lower bound on distortion.

As a consequence of Theorem 1.1.1, the intrinsic large-scale geometry of the handlebody group cannot simply be inferred from the geometry of surface mapping class groups. In fact, the large-scale geometry of handlebody groups shows features that distinguish it from $\mathrm{Map}(\partial V_g)$.

One such difference can be found in the geometric structure of curve stabilizers. Consider a nonseparating disk D in V_g and a simple closed curve α on ∂V_g which intersects ∂D in a single point.

We consider the stabilizer of the free homotopy class of α both in $\mathrm{Map}(\partial V_g)$ and $\mathrm{Map}(V_g)$. In the surface mapping class group, stabilizers of free homotopy classes of simple closed curves are undistorted (see [MM00], [HM10] or [Ha09b]). The main theorem of Chapter 3 states that this is not the case for the handlebody group.

Theorem 1.1.2. *Suppose that the genus of V is at least 3. Then the stabilizer of a curve α as above is at least exponentially distorted in the handlebody group.*

To put this theorem into context, we again consider outer automorphism groups of free groups. Since the kernel of the projection $\mathrm{Map}(V_g) \to \mathrm{Out}(F_g)$ is infinitely generated [McC85], it is in general not possible to transfer geometric properties from $\mathrm{Out}(F_g)$ to the handlebody group.

Nevertheless, the analog of Theorem 1.1.2 holds in the case of $\mathrm{Out}(F_g)$. Namely, a simple closed curve α as above defines a primitive element in $\pi_1(V_g)$. The stabilizer of α in the handlebody group projects onto the stabilizer of the conjugacy class of the free factor generated by this primitive element. By a theorem of Handel and Mosher [HM10], this stabilizer is exponentially distorted in $\mathrm{Out}(F_g)$.

In fact, this distortion is the central ingredient to the proof of Theorem 1.1.2. Namely, we explicitly construct a sequence in the handlebody group such that on the one hand, the word norms in $\mathrm{Map}(V_g)$ grow linearly. On the other hand, we use the projections to $\mathrm{Out}(F_g)$ and the result of Handel and Mosher to conclude that the word norms in the stabilizer grow exponentially fast.

The upper bound of distortion is again obtained by a topological surgery argument using the graph of rigid racks as a geometric model for the handlebody group.

To further understand the intrinsic large-scale geometry of handlebody groups, one can consider quasi-isometry invariants from geometric group theory. One very useful of these invariants is the growth rate of the Dehn function of a finitely presented group. The Dehn function can be defined as the isoperimetric function of a presentation complex of the group. The growth rate of the Dehn function is a quasi-isometry invariant of finitely presented groups which has consequences for geometric and algorithmic properties of the group ([Gr87], [ECHLPT]).

Mapping class groups have quadratic Dehn functions (since they are automatic [Mo95]). Outer automorphism groups of free groups on the other hand are known to have exponential Dehn function ([HV96], [BV95], [HM10], [BV10]). Hence, estimating the Dehn function of the handlebody group may help in comparing it to both surface mapping class groups and outer automorphism groups of free groups. In Chapter 4 we prove an upper bound for the Dehn function of $\mathrm{Map}(V_g)$. The results in this chapter are again obtained as joint work with Ursula Hamenstädt (see [HH11c]). We show

Theorem 1.1.3 ([HH11c])**.** *The Dehn function of the handlebody group has at most exponential growth rate.*

The proof of this theorem uses a generalization of the graph of rigid racks. Using this new geometric model, we then explicitly decompose a loop of length L in the handlebody group into a concatenation of at most $e^{k \cdot L}$ uniformly short loops. The main tool used in this construction are the splitting paths already used in the proof of the upper distortion bound of Theorem 1.1.1. The main additional step in the proof of Theorem 1.1.3 is a careful combinatorial control of splitting paths issuing in adjacent vertices of the graph of rigid racks.

1.2 Outer automorphisms of free groups

The second type of mapping class groups we study in this thesis are those of doubled handlebodies. The closed three-manifold M_g obtained by doubling a handlebody V_g across its boundary surface is homeomorphic to a g-fold connected sum of $S^1 \times S^2$ (here, S^i denotes the i-dimensional sphere). By a theorem of Laudenbach [L74], the natural map from the mapping class group of M_g to $\mathrm{Out}(F_g)$ has finite kernel.

Both geometric and algebraic properties of outer automorphism groups of free groups have been studied intensively over the last years, (see e.g. [CV86], [BH92], [BFH97] or [V02] for an excellent survey).

The description of $\mathrm{Out}(F_g)$ as a cofinite quotient of the mapping class group of M_g allows to study it using topological methods. Although this point of view was already taken by Hatcher [Ha95] to study homological properties of $\mathrm{Out}(F_g)$, it has not been used to systematically investigate the large-scale geometry of $\mathrm{Out}(F_g)$.

In Chapter 5 we begin such a study. We use techniques inspired by recent results on surface mapping class groups to investigate the extrinsic geometry of two natural families of subgroups in $\mathrm{Out}(F_g)$. The results in this chapter are joint work with Ursula Hamenstädt (see [HH11b]).

The basic tool is a geometric model for $\mathrm{Out}(F_g)$ using simple sphere systems in M_g, which was defined by Hatcher [Ha95]. A *simple sphere system* is a collection of disjointly embedded, pairwise nonhomotopic 2–spheres in M_g

with simply connected complement. Using a topological surgery procedure for simple sphere systems in M_g we show that stabilizers of essential spheres in M_g are undistorted in $\mathrm{Map}(M_g)$. By rephrasing this result in group-theoretic terms we obtain a new short proof of the following recent result by Handel and Mosher [HM10].

Theorem 1.2.1 ([HM10, Theorem 1 (1) and Theorem 7] and [HH11b, Theorem 2.1])**.**

i) The stabilizer of the conjugacy class of a free splitting of F_g is undistorted in $\mathrm{Out}(F_g)$.

ii) The stabilizer of the conjugacy class of a corank 1 free factor is undistorted in $\mathrm{Out}(F_g)$.

As another application of this strategy we can connect the geometry of surface mapping class groups to the geometry of $\mathrm{Out}(F_n)$. Consider a surface $S_{g,1}$ of genus g with one puncture. The fundamental group of $S_{g,1}$ is a free group on $2g$ generators. Using a variant of the Dehn-Nielsen-Baer theorem, the mapping class group of $S_{g,1}$ can be identified with a subgroup of $\mathrm{Out}(F_{2g})$. We show

Theorem 1.2.2 ([HH11b, Theorem 3.2])**.** *The mapping class group of $S_{g,1}$ is undistorted in $\mathrm{Out}(F_{2g})$.*

The proof of this theorem again uses a topological argument in M_g. Namely, we consider embeddings of a surface S_g^1 of genus g with one boundary component into M_{2g} which induce an isomorphism of fundamental groups. By intersecting such an embedded surface with a simple sphere system for M_{2g} one obtains a binding arc system for the surface S_g^1. Such binding arc systems can be used to construct a geometric model for the mapping class group of $S_{g,1}$. We use this intersection procedure to produce paths in $\mathrm{Map}(S_{g,1})$ out of paths in $\mathrm{Out}(F_g)$. Using a careful control of homotopy classes of both the embedded surface and sphere systems we then use these paths to show the desired undistortion statement.

Acknowledgements. First and foremost, I would like to thank my supervisor Ursula Hamenstädt for her guidance and advice. I would also like to thank Piotr Przytycki for many interesting discussions. Finally, I would like to thank the Max Planck Institute for Mathematics in Bonn for the generous financial support during my PhD studies.

Chapter 2

Extrinsic geometry of handlebody groups[1]

2.1 Introduction

A handlebody V_g of genus g is a 3-manifold bounded by a closed orientable surface $\partial V_g = S_g$ of genus g. Explicitly, V_g can be constructed by attaching g one-handles to a 3-ball. Handlebodies are basic building blocks for closed 3-manifolds, since any such manifold can be obtained by gluing two handlebodies along their boundaries.

The handlebody group $\mathrm{Map}(V_g)$ is the subgroup of the mapping class group $\mathrm{Map}(\partial V_g)$ of the boundary surface defined by isotopy classes of those orientation preserving homeomorphisms of ∂V_g which can be extended to homeomorphisms of V_g. It turns out that $\mathrm{Map}(V_g)$ can be identified with the group of orientation preserving homeomorphisms of V_g up to isotopy.

The handlebody group is a finitely presented subgroup of the mapping class group (compare [Wa98] and [S77]), and hence it can be equipped with a word norm. The goal of this chapter is to initiate an investigation of the coarse geometry of the handlebody group induced by this word norm.

The geometry of mapping class groups of surfaces is quite well understood. Therefore, understanding the geometry of the inclusion homomor-

[1] This chapter is identical with the preprint [HH11a]
Ursula Hamenstädt and Sebastian Hensel,
The geometry of the handlebody groups I: Distortion,
arXiv:1101.1838

phism $\mathrm{Map}(V_g) \to \mathrm{Map}(\partial V_g)$ may allow to deduce geometric properties of the handlebody group from geometric properties of the mapping class group. This task would be particularly easy if the handlebody group was undistorted in the ambient mapping class group (i.e. if the inclusion was a quasi-isometric embedding).

Many natural subgroups of the mapping class group are known to be undistorted. One example is given by groups generated by Dehn twists about disjoint curves (studied by Farb, Lubotzky and Minsky in [FLM01]) where undistortion can be proved by considering the subsurface projections onto annuli around the core curves of the Dehn twists.

Another example of undistorted subgroups are mapping class groups of subsurfaces (compare [MM00] or [Ha09a]). In this case, the proof of undistortion relies on the construction of quasi-geodesics in the mapping class group – either train track splitting sequences as in [Ha09a] or hierarchy paths defined by Masur and Minsky in [MM00].

Other important subgroups of the mapping class group are known to be distorted. As one example we mention the Torelli group, which is exponentially distorted by [BFP07]. A finitely generated subgroup H of a finitely generated group G is called exponentially distorted in G if the following holds. On the one hand, the word norm in H of every element $h \in H$ is coarsely bounded from above by an exponential of the word norm of h in G. On the other hand, there is a sequence of elements $h_i \in H$ such that the word norm of h_i in G grows linearly, while the word norm of h_i in H grows exponentially.

The argument from [BFP07] can be used to show exponential distortion for other normal subgroups of the mapping class group as well. Since the handlebody group is not normal, it cannot be used to analyze the handlebody group.

Answering a question raised in [BFP07], we show that nevertheless the same conclusion holds true for handlebody groups in almost all cases.

Theorem. *The handlebody group for genus $g \geq 2$ is exponentially distorted in the mapping class group.*

Our result is also valid for handlebodies with marked points or spots; allowing to lower the genus to 1 if there are at least two marked points or spots. In the case of genus 0 and the solid torus with one marked point the handlebody group is obviously undistorted and hence we obtain a complete classification of distorted handlebody groups.

The basic idea for the proof of the main theorem can be sketched in the special case of a solid torus $V_{1,2}$ with two marked points. The handlebody group of a solid torus with one marked point is infinite cyclic, generated by the Dehn twist T about the unique essential simple diskbounding curve. Since point-pushing maps are contained in the handlebody group, the Birman exact sequence yields that $\mathrm{Map}(V_{1,2})$ is equal to the fundamental group of the mapping torus of the once-punctured torus defined by T. The Dehn twist T acts on the fiber $\pi_1(T_{1,1}) = F_2$ of the Birman exact sequence as a Nielsen twist, therefore in particular as an element of linear growth type. This implies that the fiber is undistorted in the handlebody group. As this fiber is exponentially distorted in the mapping class group by [BFP07], the handlebody group of a torus with two marked points is at least exponentially distorted in the corresponding mapping class group.

In the general case, the argument is more involved since we have no explicit description of the handlebody group. However, the basic idea remains to show that parts of the fiber of some suitable Birman exact sequence are undistorted in the handlebody group.

The upper distortion bound uses a geometric model for the handlebody group. This model, the *graph of rigid racks*, is similar in spirit to the train track graph which was used in [Ha09a] to study the mapping class group. We construct a family of distinguished paths connecting any pair of points in this graph to each other. The length of these paths can be bounded using intersection numbers. The geometric control obtained this way allows to show the exponential upper bound on distortion.

The chapter is organized as follows. In Section 2.2 we recall basic facts about handlebody groups of genus 0 and 1. Section 2.3 contains the lower distortion bound for handlebodies with at least one marked point or spot. In Section 2.4 we show the lower distortion bound for closed surfaces. Section 2.5 introduces a surgery procedure for disk systems which is important for the construction of paths in the handlebody group. Section 2.6 is devoted to the construction of racks, and demonstrates some of their similarities (and differences) to train tracks on surfaces. Section 2.7 contains the construction of the geometric model for the handlebody group and a distinguished family of paths establishing the upper bound on distortion.

2.2 Low-complexity cases

As a first step, we analyze the cases of those genus 0 and 1 handlebody groups which turn out to be undistorted. The results in this section are easy and well-known, and we record them here for completeness.

To formulate the results in full generality, we need to introduce the notion of handlebodies with marked points and spots. A handlebody of genus g with k marked points and s spots $V^s_{g,k}$ is a handlebody of genus g, together with s pairwise disjoint disks $D_1, \dots, D_s$ on its boundary surface S_g, and k pairwise distinct points $p_1, \dots, p_k$ in $\partial V_g \setminus (D_1 \cup \dots \cup D_s)$.

The mapping class group $\mathrm{Map}(\partial V^s_{g,k}, p_1, \dots, p_k, D_1, \dots, D_s)$ of the boundary surface (with the same marked points and disks) consists of homeomorphisms of ∂V_g which fix the set $\{p_1, \dots, p_k\}$ and restrict to the identity on each of the D_i up to isotopy respecting the same data. Note that this group agrees with the mapping class group of the bordered surface obtained by removing the interior of the marked disks, as these mapping classes have to fix each boundary component (following the definition in [FM11, Section 2.1]). In the same way as for the case without marked points or spots, the handlebody group $\mathrm{Map}(V^s_{g,p}, p_1, \dots, p_k, D_1, \dots, D_s)$ is defined as the subgroup of those isotopy classes of homeomorphisms that extend to the interior of $V^s_{g,p}$.

All curves and disks are required not to meet any of the marked points. A simple closed curve on ∂V is *essential* if it is neither contractible nor freely homotopic to a marked point. A disk D in V is called *essential*, if $\partial D \subset \partial V$ is an essential simple closed curve.

Proposition 2.2.1. *Let $V = V^s_{0,k}$ be a handlebody of genus 0, with any number of marked points and spots. Then the handlebody group of V is equal to the mapping class group of its boundary.*

Proof. Let $f : S^2 \to S^2$ be any homeomorphism of the standard 2-sphere $S^2 \subset \mathbb{R}^3$ onto itself. We can explicitly construct a radial extension $F : D^3 \to D^3$ to the standard 3-ball $D^3 \subset \mathbb{R}^3$ by setting $F(t \cdot x) = t \cdot f(x)$ for $x \in S^2, t \in [0, 1]$. Therefore every mapping class group element is contained in the handlebody group. □

In particular, the handlebody groups of genus 0 are undistorted in the corresponding mapping class groups. Similarly, for a solid torus with at most one marked point or spot, the handlebody group can be explicitly identified and turns out to be undistorted.

To this end, suppose V is a solid torus with at most one marked point ($V = V_{1,0}$ or $V = V_{1,1}$) or with one marked spot ($V = V_1^1$). Let δ be an essential simple closed curve on the boundary torus of V that bounds a disk in V. The curve δ is uniquely determined up to isotopy.

Proposition 2.2.2. *The handlebody group of V is the stabilizer of δ in the mapping class group. In particular, it is undistorted in the mapping class group.*

Thus, if $V = V_{1,0}$ or $V = V_{1,1}$, then the handlebody group is cyclic and generated by the Dehn twist about δ.

If $V = V_1^1$, then the handlebody group is the free abelian group of rank 2 *which is generated by the Dehn twist about δ and the Dehn twist about the spot.*

Proof. The handlebody group fixes the set of isotopy classes of essential disks in V. Since δ is the unique diskbounding curve up to isotopy, $\mathrm{Map}(V)$ therefore is contained in the stabilizer of δ. On the other hand, the disk bounded by δ cuts V into a spotted ball. Hence, by Proposition 2.2.1 the handlebody group $\mathrm{Map}(V)$ contains the stabilizer of δ.

If $V = V_{1,0}$ or $V_{1,1}$, the complement of δ in ∂V is an annulus (possibly with a puncture). From this, it is immediate that the handlebody group is generated by the Dehn twist about δ.

If $V = V_1^1$, the same argument shows that then the handlebody group is generated by the Dehn twist about δ and the spot. It is clear that these mapping classes commute.

Since stabilizers of simple closed curves are known to be undistorted subgroups of the mapping class group (compare [MM00] or [Ha09b]), the handlebody group of a solid torus with at most one spot or marked point is undistorted. ☐

2.3 Handlebodies with marked points

In this section we describe the lower bound for distortion of handlebody groups with marked points. We begin with the case of genus $g \geq 2$ with a single marked point. The case of several marked points or spots will be an easy consequence of this result. The case of a torus with several marked points requires a different argument which will be given at the end of this section.

Theorem 2.3.1. *Let $V = V_{g,1}$ be a handlebody of genus $g \geq 2$ with one marked point, and let $\partial V = S_{g,1}$ be its boundary surface. Then the handlebody group* $\mathrm{Map}(V) < \mathrm{Map}(\partial V)$ *is at least exponentially distorted.*

The proof is based on the relation between the mapping class group of a closed surface S_g and the mapping class group of a once-punctured surface $S_{g,1}$. We denote the marked point of $\partial V = S_{g,1}$ by p, and we will often denote the mapping class group of $S_{g,1}$ by $\mathrm{Map}(S_g, p)$.

Recall the definition of the *point-pushing map* $\mathcal{P} : \pi_1(S, p) \to \mathrm{Map}(S, p)$. Namely, let $\gamma : [0, 1] \to S$ be a loop in S based at p. Then there is an isotopy $f_t : S \to S$ supported in a small neighborhood of the loop $\gamma[0, 1]$ such that $f_0 = \mathrm{id}$, and $f_t(p) = \gamma(t)$. To see this, note that locally around $\gamma(t_0)$ such an isotopy certainly exists (for example, since any orientation preserving homeomorphism of the disk is isotopic to the identity). The image of γ is compact, and hence the desired isotopy can be pieced together from finitely many such local isotopies. The endpoint f_1 of such an isotopy is a homeomorphism of (S, p). We call its isotopy class the point pushing map $\mathcal{P}(\gamma)$ along γ. It depends only on the homotopy class of γ.

The image of the point pushing map is contained in the handlebody group $\mathrm{Map}(V, p)$ – to see this, simply define the local version by pushing a small half-ball instead of a disk.

By construction, the image of the point pushing map lies in the kernel of the forgetful homomorphism $\mathrm{Map}(S, p) \to \mathrm{Map}(S)$ induced by the puncture forgetting map $(S, p) \to (S, S)$. In fact this is all of the kernel, compare [Bi74].

Theorem 2.3.2 (Birman exact sequence)**.** *Let S be a closed oriented surface of genus $g \geq 2$ and $p \in S$ any point. The sequence*

$$1 \longrightarrow \pi_1(S, p) \xrightarrow{\mathcal{P}} \mathrm{Map}(S, p) \longrightarrow \mathrm{Map}(S) \longrightarrow 1$$

is exact.

The point pushing map is natural in the sense that

$$\mathcal{P}(f\alpha) = f \circ \mathcal{P}(\alpha) \circ f^{-1} \tag{2.1}$$

for each $f \in \mathrm{Map}(S, p)$ (see [Bi74] for a proof of this fact).

The Birman exact sequence corresponds to the relation between the inner and the outer automorphism group of $\pi_1(S,p)$:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & \pi_1(S,p) & \xrightarrow{\mathcal{P}} & \mathrm{Map}(S,p) & \longrightarrow & \mathrm{Map}(S) & \longrightarrow & 1 \\ & & \downarrow \cong & & \downarrow \cong & & \downarrow \cong & & \\ 1 & \longrightarrow & \mathrm{Inn}(\pi_1(S,p)) & \longrightarrow & \mathrm{Aut}(\pi_1(S,p)) & \longrightarrow & \mathrm{Out}(\pi_1(S,p)) & \longrightarrow & 1 \end{array}$$

where $\pi_1(S,p)$ can be identified with its inner automorphism group because it has trivial center, and the other two isomorphisms are given by the Dehn-Nielsen-Baer theorem. In other words, we have the following.

Lemma 2.3.3. *Let $[\gamma],[\alpha] \in \pi_1(S,p)$ be two loops at p. Then*

$$\mathcal{P}(\alpha)(\gamma) = [\alpha] * [\gamma] * [\alpha]^{-1}$$

where $*$ denotes concatenation of loops, and takes place left-to-right.

Now we are ready to give the proof of the main theorem of this section.

Proof of Theorem 2.3.1. Let δ be a separating simple closed curve on S such that one component of $S \setminus \delta$ is a bordered torus T with one boundary circle, and such that δ bounds a disk $\mathcal{D}$ in the handlebody V. Without loss of generality we assume that the base point p lies on δ.

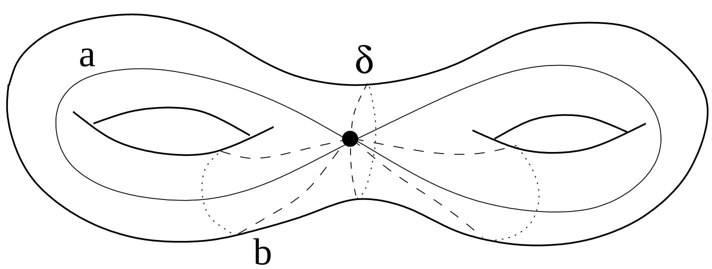


Figure 2.1: The setup in the proof of Theorem 2.3.1. Generators for the fundamental group of the handlebody are drawn solid, the loops extending these to a generating set of $\pi_1(S,p)$ are drawn dashed.

Choose loops a, b based at p which generate the fundamental group of T and such that b bounds a disk in V (and hence a does not). Extend a, b to a generating set of the fundamental group of $\pi_1(S,p)$ by adding loops in the complement of T (see Figure 2.1). Let $f \in \mathrm{Map}(S, p)$ be a mapping class such that $f(a) = a^2 * b$ and $f(b) = a * b$ which preserves δ and acts as the identity on $S \setminus T$. Such an f can for example be obtained as the composition of suitably oriented Dehn twists along a and b.

Define $\Phi_k = \mathcal{P}(f^k a)$. By Equation (2.1), in the mapping class group $\mathrm{Map}(S, p)$ we have $\Phi_k = f^k \mathcal{P}(a) f^{-k}$, and hence the word norm of Φ_k in the mapping class group with respect to any generating set grows linearly in k.

On the other hand, consider the map

$$\mathrm{Map}(V, p) \xrightarrow{\pi} \mathrm{Aut}(\pi_1(V, p)) = \mathrm{Aut}(F_g)$$

defined by the action on the fundamental group. Lemma 2.3.3 implies that Φ_k acts on $\pi_1(S, p)$ as conjugation by $f^k(a)$. To compute the action of $\pi(\Phi_k)$ on $\pi_1(V, p)$, denote the projection of the fundamental group of the surface S to the fundamental group of the handlebody by $P : \pi_1(S, p) \to \pi_1(V, p)$.

Since b bounds a disk in V, its projection vanishes: $P(b) = 0$. The generator a of $\pi_1(S, p)$ projects to a primitive element in $\pi_1(V, p)$, $P(a) = A$. Hence $P(f^k(a)) = A^{N_k}$ for some $N_k > 0$. The choice of f guarantees that we have $N_k \geq 2^k$. Since the point pushing map is natural with respect to the projection to the handlebody, $\pi(\Phi_k)$ acts on $\pi_1(V, p)$ as conjugation by A^{N_k}.

In other words, as an element of $\mathrm{Aut}(F_g)$ the projection $\pi(\Phi_k)$ is the N_k–fold power of the conjugation by A. Since conjugation by A is an infinite order element in $\mathrm{Aut}(F_g)$ and all infinite order elements have positive translation length (compare [A02, Theorem 1.1]) this implies that the word norm of $\pi(\Phi_k)$ grows exponentially in k. As $\pi : \mathrm{Map}(V, p) \to \mathrm{Aut}(F_g)$ is a surjective homomorphism between finitely generated groups, it is Lipschitz with respect to any choice of word metrics. Therefore, the word norm of Φ_k in $\mathrm{Map}(V, p)$ also grows exponentially in k. This shows the theorem. □

Remark 2.3.4. The proof we gave extends verbatim to the case of the pure handlebody group of a handlebody of genus $g \geq 2$ with several marked points and any number of spots (just move everything but one marked point into the complement of T). Here, the pure handlebody group is the subgroup of those mapping classes which send each marked point to itself. Since this group has finite index in the full handlebody group, the proof also shows that handlebody groups with several marked points and any number of spots are at least exponentially distorted if the genus is at least 2.

As a next case, we consider handlebody groups of handlebodies with spots instead of marked points.

Corollary 2.3.5. *Let $V = V_g$ be a genus $g \geq 2$ handlebody and let $D \subset \partial V$ be a spot. Then the handlebody group* $\mathrm{Map}(V, D) < \mathrm{Map}(\partial V, D)$ *of the spotted handlebody is at least exponentially distorted.*

Proof. Note that there is a commutative diagram with surjective projection homomorphisms

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \langle T \rangle & \longrightarrow & \mathrm{Map}(V, D) & \longrightarrow & \mathrm{Map}(V, p) & \longrightarrow & 0 \\ & & \downarrow{=} & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & \langle T \rangle & \longrightarrow & \mathrm{Map}(\partial V, D) & \longrightarrow & \mathrm{Map}(\partial V, p) & \longrightarrow & 0 \end{array}$$

induced by collapsing the marked spot to a point. The kernel of such a projection homomorphism is infinite cyclic and generated by the Dehn twist T about the spot. In particular, every element g in $\mathrm{Map}(\partial V, p)$ lifts to an element in $\mathrm{Map}(\partial V, D)$, and if $g \in \mathrm{Map}(V, p)$ then the lift is contained in the handlebody group $\mathrm{Map}(V, D)$. These lifts are well-defined up to the Dehn twist T which lies in the handlebody group and acts trivially on $\pi_1(V, p)$.

Choose any lift $\widetilde{f}$ of the element f used in the proof of Theorem 2.3.1. Let $\widetilde{\Phi}$ be a lift of the point pushing map Φ_0 defined in the proof of Theorem 2.3.1, and define $\widetilde{\Phi}_k = \widetilde{f}^k \widetilde{\Phi} \widetilde{f}^{-k}$. Note that these elements are lifts of the elements Φ_k and therefore contained in the handlebody group.

Now $\widetilde{\Phi}_k$ has word norm in $\mathrm{Map}(S, D)$ again bounded linearly in k. As elements of the spotted handlebody group the word norm of $\widetilde{\Phi}_k$ grows exponentially in k, as this is true for the Φ_k. □

Remark 2.3.6. Again, the same proof works for handlebodies with more than one spot and any number of marked points.

As a last case, we consider the handlebody of a torus with more than one marked point.

Theorem 2.3.7. *Let $V = V_{1,n}$ be a solid torus with $n \geq 2$ marked points. Then the handlebody group* $\mathrm{Map}(V)$ *is at least exponentially distorted in* $\mathrm{Map}(\partial V)$*.*

Proof. The strategy of this proof is similar to the preceding ones. We consider the Birman exact sequence for pure mapping class groups and pure handlebody groups.

$$\begin{array}{ccccccccc} 1 & \longrightarrow & \pi_1\mathcal{C}_n & \xrightarrow{\mathcal{P}} & P\mathrm{Map}(\partial V, p_0, p_1, \ldots, p_n) & \longrightarrow & \mathrm{Map}(\partial V, p_0) & \longrightarrow & 1 \\ & & \uparrow{=} & & \uparrow & & \uparrow & & \\ 1 & \longrightarrow & \pi_1\mathcal{C}_n & \longrightarrow & P\mathrm{Map}(V, p_0, p_1, \ldots, p_n) & \longrightarrow & \mathbb{Z} = \langle T \rangle & \longrightarrow & 1 \end{array}$$

where $\mathcal{C}_n$ denotes the configuration space of n points in $\partial V \setminus \{p_0\}$, and T the Dehn twist along the (unique) disk δ on $\partial V \setminus \{p_0\}$. An element of $\pi_1\mathcal{C}_n$ can be viewed as an n-tuple of parametrized loops γ_i, where γ_i is based at p_i (subject to the condition that at each point in time, the values of all these loops are distinct). Note that the pure mapping class group $P\mathrm{Map}(\partial V, p_0, p_1, \ldots, p_n)$ acts on $\mathcal{C}_n$ by acting on all component loops. The map $\mathcal{P}$ is the generalized point pushing map, pushing all marked points simultaneously along the loops γ_i. The map $\mathcal{P}$ is natural with respect to the action of $P\mathrm{Map}(\partial V, p_0, p_1, \ldots, p_n)$ in the sense that $\mathcal{P}(f\gamma) = f \circ \mathcal{P}(\gamma) \circ f^{-1}$.

Every element of $P\mathrm{Map}(V, p_0, p_1, \ldots, p_n)$ can be written in the form $\mathcal{P}(\gamma) \cdot \widetilde{T}^l$, where γ denotes an n-tuple of loops, and $\widetilde{T}$ is some (fixed) lift of the Dehn twist T. In this description, the multiplicity l and the homotopy class of the n-tuple of loops γ is well-defined. Now note that

$$\begin{aligned} \left(\mathcal{P}(\gamma) \cdot \widetilde{T}^l\right) \cdot \left(\mathcal{P}(\gamma') \cdot \widetilde{T}^{l'}\right) &= \mathcal{P}(\gamma) \cdot \mathcal{P}\left(\widetilde{T}^{l'}(\gamma')\right) \widetilde{T}^{l+l'} \qquad (2.2) \\ &= \mathcal{P}\left(\widetilde{T}^{l'}(\gamma') * \gamma\right) \widetilde{T}^{l+l'} \end{aligned}$$

by the naturality of $\mathcal{P}$ and the fact that $\mathcal{P}$ is a homomorphism (note that concatenation of loops is executed left-to-right, while composition of maps is done right-to-left).

Choose an element $\beta \in \pi_1(\partial V, p_0) = F_2$ which extends δ to a basis of $\pi_1(\partial V, p_0)$. Note that then β is a generator of the fundamental group $\pi_1(V, p_0) = \mathbb{Z}$ of the solid torus V_1. We also choose loops $\beta_i \in \pi_1(\partial V, p_i)$ for all $i = 1, \ldots, n$ which are freely homotopic to β. These loops give an identification of $\pi_1(V, p_i)$ with $\mathbb{Z}$.

Define a map $b : P\mathrm{Map}(V, p_0, p_1, \ldots, p_n) \to \mathbb{Z}$ as follows. Let $\varphi = \mathcal{P}(\gamma) \cdot \widetilde{T}^l$ be any element of the pure handlebody group. Each component loop γ_i of γ defines a loop in $\pi_1(V, p_i)$ (which might be trivial). This loop is homotopic to the k_i-th power of β_i for some number k_i. Associate to φ the sum of the k_i.

Now choose any generating set $\gamma^1, \ldots, \gamma^N$ of $\pi_1 \mathcal{C}_n$. Then the pure handlebody group $P\mathrm{Map}(V, p_0, p_1, \ldots, p_n)$ is generated by $\mathcal{P}(\gamma^j)$ and $\widetilde{T}$. We claim that there is a constant k_0, such that

$$b\left(\varphi \cdot \mathcal{P}(\gamma^i)\right) \geq b(\varphi) - k_0 \tag{2.3}$$

Namely, by equation (2.2), we have to compare the projections of the components of

$$\gamma \qquad \text{and} \qquad \widetilde{T}^l(\gamma^j) * \gamma$$

to each of the $\pi_1(V, p_i)$. However, applying $\widetilde{T}$ does not change this projection. Since γ^j is one of finitely many generators, there is a maximal number of occurrences of the projection of β_i which can be canceled by adding the projection of γ^j. This shows inequality (2.3).

Now we can finish the proof using a similar argument as in the proof of Theorem 2.3.1. Namely, choose again f a pseudo-Anosov element with the property that applying f multiplies the number of occurrences of β_i by 2 in all $\pi_1(\partial V, p_i)$. Then $\mathcal{P}(f^k\beta)$ has length growing linearly in the mapping class group, while $b(f^k\beta)$ grows exponentially. By inequality (2.3) this implies that the word norm in the pure handlebody group also grows exponentially. Since the pure handlebody group has finite index in the full handlebody group the theorem follows. □

Remark 2.3.8. The same argument that extends Theorem 2.3.1 to Corollary 2.3.5 applies in this case and shows that also all torus handlebody groups with at least two spots or marked points are exponentially distorted.

2.4 Handlebodies without marked points

In this section we complete the proof of the exponential lower bound on the distortion of the handlebody groups by showing that the handlebody group of a handlebody of genus $g \geq 2$ without marked points or spots is distorted in the mapping class group.

For genus $g \geq 3$, the idea is to replace the point pushing used in the proofs above by pushing a subsurface around the handlebody. The resulting handlebody group element does not induce a conjugation on $\pi_1(V, p)$, but instead induces a partial conjugation on the fundamental group of the complement of the pushed subsurface. Since $g \geq 3$, such an element projects to a nontrivial element in the outer automorphism group of F_g. Then a similar reasoning as in Section 2.3 applies. The case of genus 2 requires a different argument and will be given at the end of this section.

Theorem 2.4.1. *For a handlebody $V = V_g$ of genus $g \geq 3$, the handlebody group* $\mathrm{Map}(V)$ *is at least exponentially distorted in the mapping class group* $\mathrm{Map}(\partial V)$.

Proof. Choose a curve δ which bounds a disk $\mathcal{D}$, such that $V \setminus \mathcal{D}$ is the union of a once-spotted genus 2 handlebody V_1 and a once-spotted genus $g-2$ handlebody V_2. Denote the boundary of V_i by S_i, and choose a basepoint $p \in \delta$. This defines a free decomposition of the fundamental group of the handlebody

$$F_g = \pi_1(V, p) = \pi_1(V_1, p) * \pi_1(V_2, p) = F_2 * F_{g-2}.$$

We denote by $\mathrm{Map}(S_i, \delta)$ the mapping class group of the bordered surface S_i, emphasizing that each such mapping class has to fix δ pointwise. The stabilizer of δ in the mapping class group of S is of the form

$$G_S = \mathrm{Map}(S_1, \delta) \times \mathrm{Map}(S_2, \delta)/\sim$$

where the equivalence relation $\sim$ identifies the Dehn twist about δ in the groups $\mathrm{Map}(S_1, \delta)$ and $\mathrm{Map}(S_2, \delta)$. Note that the Dehn twist about δ lies in the handlebody group and acts trivially on $\pi_1(V, p)$. Therefore, the stabilizer of δ in the handlebody group is of the form

$$G_V = \mathrm{Map}(V_1, \mathcal{D}) \times \mathrm{Map}(V_2, \mathcal{D})/\sim.$$

In particular, the handlebody group $\mathrm{Map}(V_1, \mathcal{D})$ injects into G_V. There is a homomorphism $G_V \to \mathrm{Aut}(F_2) \times \mathrm{Aut}(F_{g-2})$ induced by the actions of $\mathrm{Map}(V_i, p)$ on $\pi_1(V_i, p)$. This homomorphism is natural with respect to the inclusion $\mathrm{Aut}(F_2) \times \mathrm{Aut}(F_{g-2}) \to \mathrm{Aut}(F_g)$ defined by the free decomposition of $\pi_1(V, p)$ given above. It is also natural with respect to the inclusion $\mathrm{Aut}(F_2) \to \mathrm{Aut}(F_2) \times \mathrm{Aut}(F_{g-2})$ defined by $\mathrm{Map}(V_1, \mathcal{D}) \to G_V$. Summarizing, we have the following commutative diagram.

$$\begin{array}{ccccccc}
\mathrm{Map}(S_1,\delta) & \longrightarrow & G_S & \longrightarrow & \mathrm{Map}(S,p) & \longrightarrow & \mathrm{Map}(S) \\
\uparrow & & \uparrow & & \uparrow & & \uparrow \\
\mathrm{Map}(V_1,\mathcal{D}) & \longrightarrow & G_V & \longrightarrow & \mathrm{Map}(V,p) & \longrightarrow & \mathrm{Map}(V) \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
\mathrm{Aut}(F_2) & \longrightarrow & \mathrm{Aut}(F_2)\times\mathrm{Aut}(F_{g-2}) & \longrightarrow & \mathrm{Aut}(F_g) & \longrightarrow & \mathrm{Out}(F_g)
\end{array}$$

Let $\widetilde{\Phi}_k \in \mathrm{Map}(V_1, \mathcal{D})$ be the elements constructed in the proof of Corollary 2.3.5. The image of $\widetilde{\Phi}_k$ in $\mathrm{Aut}(F_2) \times \mathrm{Aut}(F_{g-2})$ is the N_k–th power of a conjugation in the free factor F_2 defined by V_1, and the identity on the free factor F_{g-2} defined by V_2, where $N_k \geq 2^k$. In other words, this projection is a N_k–th iterate of a partial conjugation. Therefore, it projects to a nontrivial element of infinite order in $\mathrm{Out}(F_g)$. From there, one can finish the proof using the argument in the proof of Theorem 2.3.1. □

The last case is that of a genus 2 handlebody without marked points or spots. In this case, the strategy is to use the distortion of the handlebody group of a solid torus with two spots to produce distorted elements in the stabilizer of a nonseparating disk in the genus 2 handlebody.

To make this precise, we use the following construction. Let V be a genus 2 handlebody and S its boundary surface. Choose a nonseparating essential simple closed curve δ that bounds a disk $\mathcal{D}$ in V. Cutting S at δ yields a torus S_1^2 with two boundary components δ_1 and δ_2. Choose once and for all a continuous map $S_1^2 \to S$ which maps both δ_1 and δ_2 to δ and which restricts to a homeomorphism

$$S_1^2 \setminus (\delta_1 \cup \delta_2) \to S \setminus \delta.$$

The isotopy class of such a map depends on choices, but we fix one such map for the rest of this section. This map induces induces a homomorphism

$$\mathrm{Map}(S_1^2) \to \mathrm{Stab}_{\mathrm{Map}(S)}(\delta)$$

since the homeomorphisms and isotopies used to define the mapping class group $\mathrm{Map}(S_1^2)$ of the torus S_1^2 have to fix δ_1 and δ_2 pointwise and therefore extend to S.

Since δ bounds a disk, an analogous construction works for the handlebody groups, and we obtain

$$\mathrm{Map}(V_1^2) \to \mathrm{Stab}_{\mathrm{Map}(V)}(\mathcal{D}).$$

Let $p \in \delta$ be a base point, and let a, b be smooth embedded loops in S with the following properties (compare Figure 2.2).

i) The projections A and B of a and b to $\pi_1(V, p)$ form a free basis of $\pi_1(V, p) = F_2$.

ii) The loops a and b intersect δ exactly in the basepoint p.

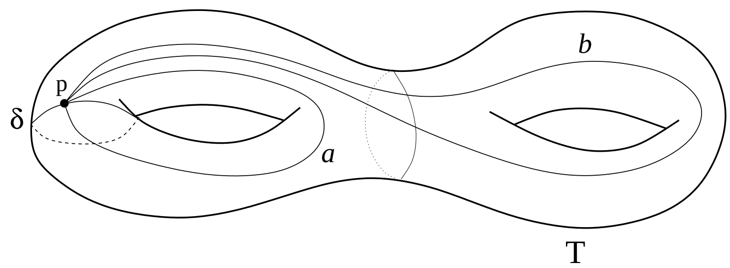


Figure 2.2: The setting for a genus 2 handlebody.

iii) The loop a hits δ from different sides at its endpoints, while b returns to the same side.

On the surface S_1^2 obtained by cutting S at δ, the loop a defines an arc from one boundary component to the other, while b defines a loop. By slight abuse of notation we will denote these objects by the same symbols. We choose the initial point of the loop b as base point of this cut-open surface, and call it again p. Then the projection B of b to the spotted solid torus V_1^2 is a generator of its fundamental group $\pi_1(V_1^2, p) = \mathbb{Z}$.

Now consider the torus $T' \subset S$ with one boundary component obtained as the tubular neighborhood of $a \cup \delta$ in S (compare Figure 2.2 for the situation). The complement of T' in S again is a torus with one boundary component which we denote by T. Choose a reducible homeomorphism f of V_1^2 which preserves T and restricts to a pseudo-Anosov homeomorphism f on the torus $T \subset S$ with the property that the projection of the loop $f^k(b)$ to $\pi_1(V_1^2)$ is B^{N_k}, for $N_k \geq 2^k$. Such an element can be constructed explicitly as in the proof of Theorem 2.3.5. In particular, we may assume that f fixes the arc a pointwise.

Consider now as in the proof of Theorem 2.3.5 the map that collapses the boundary components of V_1^2 to marked points. On this solid torus $V_{1,2}$ with two marked points, a defines an arc from marked point to marked point, and b defines a based loop at one of the marked points which we again use as base point for this surface. Let $P = \mathcal{P}(b)$ be the point pushing map on $V_{1,2}$ defined by b, and let $\widetilde{P}$ be any lift of this point-pushing map to the surface S_1^2 with boundary. As before, $\widetilde{P}$ is an element of the handlebody group. We define

$$\Phi_k = f^k \circ \widetilde{P} \circ f^{-k}.$$

Lemma 2.4.2. *Φ_k is an element of the handlebody group of V_1^2. $\Phi_k(B)$ is* *homotopic to B as a loop based at p in the handlebody V_1^2, and $\Phi_k(A)$ is homotopic, as an arc relative to its endpoints, to $A * B^{N_k}$ in V_1^2.*

Proof. Φ_k projects to the point-pushing map along $f^k(b)$ on the solid torus with two marked points $V_{1,2}$ obtained by collapsing the boundary components of V_1^2. Hence, Φ_k is the lift of a handlebody group element and therefore lies in the handlebody group itself (see the discussion in the proof of Theorem 2.3.5). This yields the first claim.

To see the other claims, we can work in the solid torus $V_{1,2}$ with two marked points, as the projection from V_1^2 to $V_{1,2}$ that collapses the spots to marked points induces a isomorphism on fundamental groups.

Here by construction Φ_k projects to the point-pushing map along $f^k(b)$. Lemma 2.3.3 now implies that this projection acts as conjugation by B^{N_k} on the fundamental group, giving the second claim.

By construction of f, the arc a and the loop $b_k = f^k(b)$ only intersect at the base point. The loop b_k is a simple curve and thus there is an embedded tubular neighborhood of b_k on $V_{1,2}$ which is orientation preserving homeomorphic to $[0,1]/(0 \sim 1) \times [-1,1] = S^1 \times [-1,1]$ and such that $S^1 \times \{0\}$ is the loop b_k. After perhaps reversing the orientation of b_k and performing an isotopy, we may assume that the intersection of a with this tubular neighborhood equals $\{0\} \times [-1,0]$.

Since b_k is simple, the point pushing map along b_k is isotopic to the map supported on the tubular neighborhood which is defined by

$$\begin{aligned}(x,t) &\mapsto (x + (t+1), t) \quad \text{for} \quad t \in [-1,0] \\ (x,t) &\mapsto (x - t, t) \quad \text{for} \quad t \in [0,1]\end{aligned}$$

This implies that the point pushing map acts on the homotopy class of a by concatenating a with the loop $f^k(b)$ (up to possibly changing the orientation of a). Since $f^k(b)$ projects to B^{N_k} in the handlebody, this implies the last claim of the lemma. □

Theorem 2.4.3. *The handlebody group of a genus 2 handlebody is at least exponentially distorted.*

Proof. We use the notation from the construction described above. Consider the image Ψ_k of Φ_k in the stabilizer of $\mathcal{D}$ in the handlebody group $\mathrm{Map}(V_2)$. By construction, Ψ_k fixes the curve δ pointwise and therefore acts on $\pi_1(V, p)$.

By the preceding lemma, this action is given by

$$\begin{array}{rcl} A & \mapsto & A * B^{N_k} \\ B & \mapsto & B \end{array}$$

Therefore, Ψ_k acts as the N_k-th power of a simple Nielsen twist on F_2. In particular, it projects to the N_k-th power of a nontrivial element in $\mathrm{Out}(F_2)$. From here, one can finish the proof as for the preceding distortion theorems. □

2.5 Disk exchanges and surgery paths

In this section we study disk systems in handlebodies and introduce certain types of surgery operations for disk systems. These surgery operations form the basis for the construction of distinguished paths in the handlebody group (see Lemma 2.7.8).

In the sequel we always consider a handlebody V of genus $g \geq 2$ with a finite number m of marked points on its boundary ∂V. The discussion remains valid if some of the marked points are replaced by spots.

Definition 2.5.1. A *disk system* for V is a set of essential disks in V which are pairwise disjoint and non-homotopic. A disk system is called *simple* if all of its complementary components are simply connected. It is called *reduced* if it is simple and has a single complementary component.

We usually consider disk systems only up to isotopy. For a handlebody of genus g, a reduced disk system consists of precisely g non-separating disks. The complement of a reduced disk system in V is a ball with $2g$ spots (and possibly some marked points). The boundary of a reduced disk system is a multicurve in ∂V with g components which cuts ∂V into a $2g$-holed sphere (with some number of marked points). The handlebody group acts transitively on the set of isotopy classes of reduced disk systems.

We say that two disk systems $\mathcal{D}_1, \mathcal{D}_2$ are in *minimal position* if their boundary multicurves intersect in the minimal number of points and if every component of $\mathcal{D}_1 \cap \mathcal{D}_2$ is an embedded arc in $\mathcal{D}_1 \cap \mathcal{D}_2$ with endpoints in $\partial\mathcal{D}_1 \cap \partial\mathcal{D}_2$. Disk systems can always be put in minimal position by applying suitable isotopies. In the sequel we always assume that disk systems are in minimal position.

Note that the minimal position of disks behaves differently than the normal position of sphere systems as defined in [Ha95]. Explicitly, let Σ be a reduced disk system and D an arbitrary disk. Suppose D is in minimal position with respect to Σ. Then a component of $D \setminus \Sigma$ may have several boundary components on the same side of a disk in Σ. In addition, the collection of components of $D \setminus \Sigma$ does not determine the disk D uniquely.

Let $\mathcal{D}$ be a disk system. An *arc relative to* $\mathcal{D}$ is a continuous embedding $\rho : [0,1] \to \partial V$ such that its endpoints $\rho(0)$ and $\rho(1)$ are contained in $\partial\mathcal{D}$. An arc ρ is called *essential* if it cannot be homotoped into $\partial\mathcal{D}$ with fixed endpoints and if the number of intersections of ρ with $\partial\mathcal{D}$ is minimal in its isotopy class.

Choose an orientation of the curves in $\partial\mathcal{D}$. Since ∂V is oriented, this choice determines a left and a right side of a component α of $\partial\mathcal{D}$ in a small annular neighborhood of α in ∂V. We then say that an endpoint $\rho(0)$ (or $\rho(1)$) of an arc ρ *lies to the right (or to the left) of* α, if a small neighborhood $\rho([0,\epsilon])$ (or $\rho([1-\epsilon,1])$) of this endpoint is contained in the right (or left) side of α in a small annulus around α. A *returning arc relative to* $\mathcal{D}$ is an arc both of whose endpoints lie on the same side of some boundary ∂D of a disk D in $\mathcal{D}$, and whose interior is disjoint from $\partial\mathcal{D}$.

Let E be a disk which is not disjoint from $\mathcal{D}$. An *outermost arc* of ∂E relative to $\mathcal{D}$ is a returning arc ρ relative to $\mathcal{D}$ such that there is a component E' of $E\setminus\mathcal{D}$ whose boundary is composed of ρ and an arc $\beta \subset D$. The interior of β is contained in the interior of D. We call such a disk E' an *outermost component* of $E \setminus \mathcal{D}$.

For every disk E which is not disjoint from $\mathcal{D}$ there are at least two distinct outermost components E', E'' of $E \setminus \mathcal{D}$. Every outermost arc of a disk is a returning arc. However, there may also be components of $\partial E \setminus \mathcal{D}$ which are returning arcs, but not outermost arcs. For example, a component of $E \setminus \mathcal{D}$ may be a rectangle bounded by two arcs contained in $\mathcal{D}$ and two subarcs of ∂E with endpoints on $\partial\mathcal{D}$ which are homotopic to a returning arc relative to $\partial\mathcal{D}$.

Let now $\mathcal{D}$ be a simple disk system and let ρ be a returning arc whose endpoints are contained in the boundary of some disk $D \in \mathcal{D}$. Then $\partial D \setminus \{\rho(0), \rho(1)\}$ is the union of two (open) intervals γ_1 and γ_2. Put $\alpha_i = \gamma_i \cup \rho$. Up to isotopy, α_1 and α_2 are simple closed curves which are disjoint from $\mathcal{D}$ (compare [St99] and [M86] for this construction). Therefore both α_1 and α_2 bound disks in the handlebody which we denote by Q_1 and Q_2. We say that Q_1 and Q_2 are obtained from D by *simple surgery along the returning arc* ρ.

The following observation is well-known (compare [M86, Lemma 3.2], or [St99]).

Lemma 2.5.2. *If Σ is a reduced disk system and ρ is a returning arc with endpoints on $D \in \Sigma$, then for exactly one choice of the disks Q_1, Q_2 defined as above, say the disk Q_1, the disk system obtained from Σ by replacing D by Q_1 is reduced.*

Proof. A reduced disk system equipped with an orientation defines a basis over $\mathbb{Z}$ for the relative homology group $H_2(V, \partial V; \mathbb{Z}) = \mathbb{Z}^n$. The homology class of the oriented disk D is the sum of the homology classes of the suitably oriented disks Q_1 and Q_2. Since D is a generator of $H_2(V, \partial V; \mathbb{Z})$, there is exactly one of the disks Q_1, Q_2, say the disk Q_1, so that the disk system $\mathcal{D}'$ obtained from $\mathcal{D}$ by replacing D by Q_1 defines a basis for $H_2(V, \partial V; \mathbb{Z})$. Then this disk system is reduced. ☐

Note that the disk Q_1 is characterized by the requirement that the two spots in the boundary of $V \setminus \Sigma$ corresponding to the two copies of D are contained in distinct connected components of $V \setminus (\Sigma \cup Q_1)$. It only depends on Σ and the returning arc ρ.

Definition 2.5.3. Let Σ be a reduced disk system. A *disk exchange move* is the replacement of a disk $D \in \Sigma$ by a disk D' which is disjoint from Σ and such that $(\Sigma \setminus D) \cup D'$ is a reduced disk system. If D' is determined as in Lemma 2.5.2 by a returning arc of a disk in a disk system $\mathcal{D}$ then the modification is called a *disk exchange move of Σ in direction of $\mathcal{D}$* or simply a *directed disk exchange move*.

A sequence (Σ_i) of reduced disk systems is called a *disk exchange sequence in direction of $\mathcal{D}$* (or *directed disk exchange sequence*) if each Σ_{i+1} is obtained from Σ_i by a disk exchange move in direction of $\mathcal{D}$.

Lemma 2.5.4. *Let Σ_1 be a reduced disk system and let $\mathcal{D}$ be any other disk system. Then there is a disk exchange sequence $\Sigma_1, \ldots, \Sigma_n$ in direction of $\mathcal{D}$ such that Σ_n is disjoint from $\mathcal{D}$.*

Proof. We define the sequence Σ_i inductively. Suppose Σ_i is already defined and not yet disjoint from $\mathcal{D}$. Then there is a outermost arc ρ of $\mathcal{D}$ with respect to Σ_i. By Lemma 2.5.2, there is a disk system Σ_{i+1} obtained by a disk exchange move along this returning arc. As a result of this surgery, the geometric intersection number between Σ_{i+1} and $\mathcal{D}$ is strictly smaller than the geometric intersection number between Σ_i and $\mathcal{D}$. Now the lemma follows by induction on the geometric intersection number between $\partial\Sigma_1$ and $\partial\mathcal{D}$. □

2.6 Racks

In this section we define and investigate combinatorial objects which serve as analogs of train tracks for handlebodies. Let again V be a handlebody of genus $g \geq 2$, perhaps with marked points on the boundary.

Definition 2.6.1. A *rack* R in V is given by a reduced disk system $\Sigma(R)$, called the *support system* of the rack R, and a collection of pairwise disjoint essential embedded arcs in $\partial V \setminus \partial\Sigma(R)$ with endpoints on $\partial\Sigma(R)$, called *ropes*, which are pairwise non-homotopic relative to $\partial\Sigma(R)$. At each side of a support disk $D \in \Sigma(R)$, there is at least one rope which ends at the disk and approaches the disk from this side.

A rack R is called *large*, if the union of $\partial\Sigma(R)$ and the set of ropes decompose ∂V into disks.

Note that the number of ropes of a rack is uniformly bounded. In the sequel we often consider isotopy classes of racks.

Explicitly, we say that two racks R, R' are *(weakly) isotopic* if their support systems $\Sigma(R), \Sigma(R')$ are isotopic and if after an identification of $\Sigma(R)$ with $\Sigma(R')$, each rope of R is freely homotopic relative to $\partial\Sigma(R)$ to a rope of R'. In Section 2.7 we will introduce a more restrictive notion of equivalence of racks.

The handlebody group $\mathrm{Map}(V)$ acts transitively on the set of reduced disk systems, and it acts on the set of weak isotopy classes of racks. For every reduced disk system Σ the stabilizer of $\partial\Sigma$ in $\mathrm{Mod}(\partial V)$ is contained in $\mathrm{Map}(V)$ (compare Proposition 2.2.1). This implies that there are only finitely many orbits for the action of $\mathrm{Map}(V)$ on the set of weak isotopy classes of racks. The stabilizer in $\mathrm{Map}(V)$ of a weak isotopy class of a rack R with support system $\Sigma(R)$ contains the group $\mathbb{Z}^n$ of Dehn twists about the components of $\partial\Sigma(R)$. In particular, this stabilizer is infinite.

Definition 2.6.2.

1. A disk system $\mathcal{D}$ (or an arbitrary geodesic lamination λ on ∂V) is *carried* by a rack R if it is in minimal position with respect to the support system $\Sigma(R)$ of R and if each component of

$\partial\mathcal{D} \setminus \partial\Sigma(R)$ (or of $\lambda \setminus \partial\Sigma(R)$) is homotopic relative to $\partial\Sigma(R)$ to a rope of R.

2. An embedded essential arc ρ in ∂V with endpoints in $\partial\Sigma(R)$ is *carried* by R if each component of $\rho \setminus \partial\Sigma(R)$ is homotopic relative to $\partial\Sigma(R)$ to a rope of R.

3. A *returning rope* of a rack R is a rope which begins and ends at the same side of some fixed support disk D (i.e. defines a returning arc relative to $\partial\Sigma(R)$).

Remark 2.6.3. i) A disk system $\mathcal{D}$ is carried by a rack R if and only if each individual disk $D \in \mathcal{D}$ is carried by R.

ii) Every disk which does not intersect the support system $\Sigma(R)$ of a rack R is not carried by R. In particular, the support system itself is not carried by R.

Let R be a rack with support system $\Sigma(R)$ and let α be a returning rope of R with endpoints on a support disk $D \in \Sigma(R)$. By Lemma 2.5.2, for one of the components γ_1, γ_2 of $\partial D \setminus \alpha$, say the component γ_1, the simple closed curve $\alpha \cup \gamma_1$ is the boundary of an embedded disk $D' \subset H$ with the property that the disk system $(\Sigma \setminus D) \cup D'$ is reduced.

A *split* of the rack R at the returning rope α is any rack R' with support system $\Sigma' = (\Sigma(R) \setminus D) \cup D'$ whose ropes are given as follows.

1. Up to isotopy, each rope ρ' of R' has its endpoints in $(\partial\Sigma(R) \setminus \partial D) \cup \gamma_1 \subset \partial\Sigma(R)$ and is an arc carried by R.

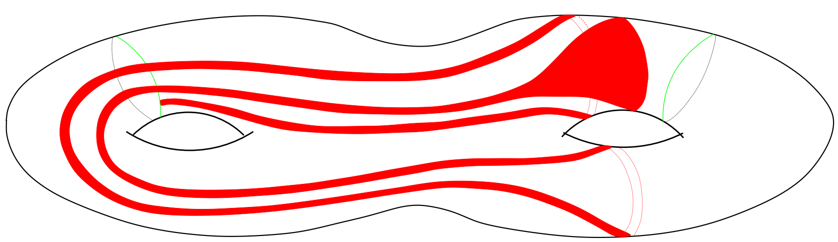
2. For every rope ρ of R there is a rope ρ' of R' such that ρ is a component of $\rho' \setminus \partial\Sigma(R)$.

The above definition implies in particular that a rope of R which does not have an endpoint on ∂D is also a rope of R'. Moreover, there is a map $\Phi : R' \to R$ which maps a rope of R' to an arc carried by R, and which maps the boundary of a support disk of R' to a simple closed curve γ of the form $\gamma_1 \circ \gamma_2$ where γ_1 either is a rope of R or trivial, and where γ_2 is a subarc of the boundary of a support disk of R (which may be the entire boundary circle). The image of Φ contains every rope of R.

Splits of racks behave differently from splits of train tracks. Although this distinction is not explicitly needed for the rest of this work, we note some important differences in the remainder of this section. For these considerations we always consider racks up to weak isotopy.

A split of a rack R at a returning rope is not unique. If R' is a split of R and if φ is a Dehn-twist about the boundary of a support disk of R then $\varphi(R')$ is a split of R as well. Moreover the following example shows that even up to the action of the group of Dehn twists about the boundaries of the support system of R, there may be infinitely many racks which can be obtained from R by a split.

Example: Let V be the handlebody of genus 2 and let Σ be a reduced disk system consisting of two disks. Let R be a rack with support system Σ which contains two distinct returning ropes α, β approaching the same support disk $D \in \Sigma$ from two distinct sides. Let $E \subset V$ be an essential disk carried by R with the following property. There is an outermost component E' of $E \setminus \Sigma$ which contains an arc homotopic to α in its boundary. Attached to $E' \subset E$ is a rectangle component $R_\beta \subset E$ of $E \setminus \Sigma$ with two opposite sides on D which is a thickening of the returning rope β. The rectangle R_β is attached to a rectangle R_α with two sides on D which is a thickening of α looping about the half-disk E'. R_α in turn is attached to a second copy of R_β etc (see the figure). A rack R' whose support system is obtained from Σ by a single disk



exchange in direction of E and which carries ∂E contains a returning rope ρ which is carried by R and so that $\rho \setminus \Sigma$ has an arbitrarily large number of components.

Another important difference between racks and train tracks concerns the relation between carrying and splitting. One the one hand, there are splits R' of R which carry disks which are not carried by R. Namely, let R be a rack and R' be a split of R. Denote the support disk of R' which is not a support disk of R by D. In particular, if D is a disk carried by both R and R', then images of D under arbitrary powers of the Dehn twist about ∂D are still carried by R', but not necessarily by R.

On the other hand, let D be a disk carried by a rack R. Then there may be no split R' of R which still carries D. Namely, R may have a single returning rope ρ and thus every split of R has the same support system Σ'. If Σ' is disjoint from D, no rack with support system Σ' carries D.

2.7 The graph of rigid racks

In this section we construct a geometric model for the handlebody group. By a geometric model we mean a connected locally finite graph on which the handlebody group acts properly and cocompactly as a group of automorphisms. The construction is similar in spirit to the construction of the train track graph in [Ha09a], which is a geometric model for the mapping class group. The model we construct admits a family of distinguished paths which are used for a coarse geometric control of the handlebody group. These paths are constructed below in Lemmas 2.7.6 and 2.7.8.

As a first step one can define a *graph of racks* $\mathcal{R}(V)$ in direct analogy to the definition of the train track graph in [Ha09a]. The vertex set of $\mathcal{R}(V)$ is the set of weak isotopy classes of large racks (satisfying a suitable completeness condition which is not important for the current work). Two such vertices are connected by an edge of length one if the corresponding racks are related by a single split. By construction, the handlebody group acts on $\mathcal{R}(V)$ as a group of automorphisms. Imitating the proof of connectivity for the train track graph from [Ha09a, Corollary 3.7] one can then show that $\mathcal{R}(V)$ is connected. Since this result is not needed in the sequel we do not include a proof here.

The graph of racks defined in this way is not a geometric model for the handlebody group, as the stabilizer of a weak isotopy class of a rack contains the group generated by Dehn twists about the support system, and thus is in particular infinite. For the same reason, the graph of racks is locally infinite. Also recall that even up to the action of the group of Dehn twists about the support system of R, there may be infinitely many different racks which can be obtained from R by a single split (as demonstrated by the example in Section 2.6).

To define a geometric model for the handlebody group using racks, we therefore have to overcome two difficulties. On the one hand, we need to record twist parameters at the support curves so that the stabilizer of a rack with a set of such twist parameters becomes finite. On the other hand, the edges have to be more restrictive than splits so that the graph becomes locally finite.

For the purposes of this chapter, these problems will be addressed by considering a more restrictive notion of equivalence of racks.

Definition 2.7.1. i) Let R be a large rack. The union of the support system and the system of ropes of R defines a cell decomposition of the surface ∂V which we call the *cell decomposition induced by R*.

ii) Let R and R' be racks. We say that R and R' are *rigidly isotopic* if the cell decompositions induced by R and R' are isotopic as cell decompositions of the surface ∂V.

In particular, if φ is a simple Dehn twist about the boundary of a support curve of a rack R, then R and $\varphi^n(R)$ are not rigidly isotopic for $n \geq 2$. This observation and the fact that the stabilizer of a reduced disk system in the mapping class group is contained in the handlebody group imply the following.

Corollary 2.7.2. *The handlebody group acts on the set of rigid isotopy classes of racks with finite quotient and finite stabilizers.*

This corollary shows that the set of rigid isotopy classes of racks can be used as the set of vertices of a Map(V)-graph which is a geometric model for Map(V).

To define a suitable set of edges for such a graph we note the following lemma.

Lemma 2.7.3. *i) There is a number $K_1 > 0$ with the following property. Let R, R' be two racks sharing the same support system. Then there is a sequence*

$$R = R_1, \dots, R_N = R'$$

of racks, such that the number of intersections between the cell decompositions induced by R_i and R_{i+1} is less than K_1 for all $i = 1, \dots, N-1$.

ii) There is a number $K_2 > 0$ with the following property. Let R be a rack and let α be a returning rope of R. Then there is a rack R' which is *obtained from R by a split along α such that the number of intersections between the cell decompositions induced by R and R' is less than K_2.*

Proof. Part *i)* of the lemma follows immediately from the fact that for every reduced disk system Σ of V, the stabilizer of $\partial\Sigma$ in the mapping class group of ∂V is contained in the handlebody group and acts with finite quotient on the set of all rigid isotopy classes of racks with a common support system.

To prove part *ii)*, let Σ' be the reduced disk system obtained from the support system of R by the disk exchange along the returning rope α. Every component of $\partial\Sigma'$ is homotopic to a union of uniformly few edges of the cell decomposition induced by R. Therefore, the number of intersections between Σ' and the cell decomposition induced by R can be uniformly bounded. Now the claim follows as in part *i)* since the stabilizer of $\partial\Sigma'$ in the mapping class group of ∂V is contained in the handlebody group. □

Definition 2.7.4. The *graph of rigid racks* $\mathcal{RR}(V)$ is the graph whose vertex set is the set of rigid isotopy classes of large racks. Two such vertices are joined by an edge if the intersection number between the cell decompositions induced by the large racks corresponding to the edges is at most K. Here K is the maximum of the constants K_1 and K_2 of Lemma 2.7.3.

Remark 2.7.5. Part *ii)* of Lemma 2.7.3 can be interpreted as the fact that twisting data about the support system of a rack R determines a finite number of splits which are adapted to these twist parameters. Furthermore, each of these possible splits carries a coarsely unique set of twist parameters induced by the original rack R.

Lemma 2.7.3 implies that $\mathcal{RR}(V)$ is connected. Since the handlebody group acts on $\mathcal{RR}(V)$ properly discontinuously and cocompactly, it is a geometric model of the handlebody group by the Svarć-Milnor-Lemma.

As a next step we define a distinguished class of paths in $\mathcal{RR}(V)$. These paths are sufficiently well-behaved to obtain a coarse geometric control for the handlebody group. The length estimates for these paths use markings and Corollary 2.A.4 which relates word norms of mapping class group elements to intersection numbers of cell decompositions. The necessary definitions and statements are given in the Appendix.

In order to simplify the notation for the rest of the chapter, we usually do not specify constants or additive and multiplicative errors in formulas, but rather state that a quantity x is "coarsely bounded" by some other quantity y (or "uniformly bounded"). By this we mean that there are constants C_1, C_2

which only depend on the genus (and the number of marked points) of V, such that x is bounded by $C_1 \cdot y + C_2$ (or C_1).

Lemma 2.7.6. *There is a number $k > 0$ satisfying the following. Let P be a pants decomposition of ∂V all of whose components bound disks in V. Let R be a large rack with support system $\Sigma(R)$. Then there is a large rack R' with the following properties.*

i) The support system $\Sigma(R')$ of R' agrees with the one of R.

ii) Each component of P which intersects the support system of R essentially is carried by R'.

iii) Each component of $P \setminus \partial\Sigma(R')$ intersects the cell decomposition induced by R' in at most k points.

iv) The distance between R and R' in $\mathcal{RR}(V)$ is coarsely bounded by the geometric intersection number between P and the 1–skeleton of the cell decomposition induced by R.

Proof. Denote the cell decomposition induced by R by C. Let S' be the surface obtained from ∂V by cutting at $\partial\Sigma(R)$. The intersection of P with S' is a union of simple closed curves and arcs connecting the boundary components of S'. We call these arcs the *arcs induced by P*. Let $\hat{R}$ be the rack whose support support system agrees with the one of R and whose ropes are given by the arcs induced by P. If $\hat{R}$ is not a large rack, then we can add ropes to $\hat{R}$ which intersect P in uniformly few points, and which intersect ropes of R in at most $i(P, C)$ points. Call the result R'.

From the construction of the rack R', properties *i)* to *iii)* are immediate. Property *iv)* follows by applying Corollary 2.A.4 to the cell decomposition C and the cell decomposition induced by R' on the subsurface S'. ☐

Definition 2.7.7. If P and R' satisfy conclusions *ii)* and *iii)* of Lemma 2.7.6 above, we say that P *is effectively carried by R'*.

The following lemma is the main step towards the upper distortion bound for the handlebody group and contains the construction of the distinguished paths in the handlebody group.

Lemma 2.7.8. *Let P be a pants decomposition all of whose components bound disks in V. Suppose P is effectively carried by a rack R with support system $\Sigma(R)$. If at least one component of P intersects $\partial\Sigma(R)$ essentially, there is a rack R' with the following properties.*

i) The support system $\Sigma(R')$ *is obtained from* $\Sigma(R)$ *by a disk exchange move in the direction of a component of* P*.*

ii) P *is effectively carried by* R'*.*

iii) The distance of R *and* R' *in* $\mathcal{RR}(V)$ *is coarsely bounded by* $i(P, \partial\Sigma(R))$*.*

Proof. Since the intersection of P with $\partial\Sigma(R)$ is nonempty, the rack R has a returning rope α corresponding to an arc induced by P.

Let Σ' be the reduced disk system obtained from $\Sigma(R)$ by a disk exchange along the returning leaf α. Each component of $\partial\Sigma'$ intersects the cell decomposition induced by R in uniformly few points. Define a rack $\hat{R}$ with support system Σ' by choosing the arcs induced by P relative to Σ' as ropes. By construction, each rope of $\hat{R}$ is obtained as a concatenation of ropes of R (as in the definition of the split of a rack). Furthermore, each rope of $\hat{R}$ intersects $\Sigma(R)$ in at most as many points as P does. Therefore, the intersection number between a rope of $\hat{R}$ and the cell decomposition induced by R can be coarsely bounded by $i(P, \partial\Sigma(R))$. We can extend $\hat{R}$ in any way to a large rack R' such that every rope of R' has the same property. Both R' and R intersect $\partial\Sigma'$ in uniformly few points. The mapping class group of $\partial V \setminus \partial\Sigma'$ is contained in the handlebody group and undistorted in the mapping class group. Hence Corollary 2.A.4 applied in the subsurface $\partial V \setminus \partial\Sigma'$ implies that the distance between R and R' in $\mathcal{RR}(V)$ is coarsely bounded by $i(P, \partial\Sigma(R))$. Now we can apply Lemma 2.7.6 to R' to obtain a rack with the desired properties. ☐

The following theorem is an easy consequence of Lemma 2.7.8.

Theorem 2.7.9. *Let* $g \geq 2$ *be arbitrary. Then the handlebody group* $\mathrm{Map}(V_g)$ *is at most exponentially distorted in the mapping class group.*

Together with the results from Sections 2.3 and 2.4 this theorem implies the main theorem from the introduction.

Proof of Theorem 2.7.9. There is a number $K > 0$ such that for every large rack R there is a pants decomposition P_R whose geometric intersection number with the cell decomposition $C(R)$ induced by R is bounded by K. This is due to the fact that the handlebody group acts cocompactly on the graph of rigid racks.

Let R_0 be a rack, and P_0 such a pants decomposition. Let f be an arbitrary element of the handlebody group. Put $P = f(P_0)$. By Proposition 2.A.3 the geometric intersection number between P and P_0 is coarsely bounded exponentially in the word norm of f in the mapping class group. Denote this bound by N.

As a first step, apply Lemma 2.7.6 to R_0 and P to construct a rack R_1 which effectively carries P and whose distance to R_0 is coarsely bounded by N. Next, use Lemma 2.7.8 to construct a rack R_2 whose distance to R_1 is again coarsely bounded by N, and such that the number of intersections between P and $\Sigma(R_2)$ is strictly less than the number of intersections between P and $\Sigma(R_1)$. Inductively repeating this procedure we find a sequence $R_1, \dots, R_K$ of racks of length K coarsely bounded by N^2, and such that P is disjoint from $\Sigma(R_K)$. In particular, there is a handlebody group element g which maps P_0 to P and whose word norm in the handlebody group is also coarsely bounded by N^2. The difference $f^{-1} \circ g$ fixes the pants decomposition P_0 and hence is a Dehn multitwist about P_0. As the group of Dehn multitwists about P_0 is contained in the handlebody group, and undistorted in the mapping class group, the word norm of $f^{-1} \circ g$ in the handlebody group is also coarsely bounded by N^2. This shows the theorem. □

2.A Markings and intersection numbers

In this Appendix we recall some facts about markings and intersection numbers which are used several times in this work.

Our terminology deviates slightly from the one used in [MM00], so we also recall the necessary definitions.

Definition 2.A.1. A *marking* μ of a surface S is a pants decomposition P of S together with a clean transversal for each curve in P. Here, a *clean transversal* to a pants curve $\gamma \in P$ is a curve c which is disjoint from all curves $\gamma' \in P \setminus \gamma$ and which intersects γ in the minimal number of points.

Two clean transversals to a curve α in a pants decomposition P differ by a Dehn twist about α (after possibly applying a half-twist about α). In this way, the set of clean transversals can be thought of as a twist normalization about the pants decomposition curves.

Note that the object we denote by "marking" is called "complete clean marking" in the terminology of [MM00]. The more general notion of marking used in [MM00] does not play any role in the present work.

Let S be a oriented surface of finite type and negative Euler characteristic (possibly with punctures and boundary components). *Subsurface projections to annuli* in S are defined in the following way (compare [MM00]). Recall that the arc complex of a closed annulus A is the graph whose vertex set is the set of arcs connecting the two boundary components of A up to isotopy fixing ∂A pointwise. Two such vertices are connected by an edge of length one, if the corresponding arcs can be realized with disjoint interior.

Let α be an essential simple closed curve on S. By S_α we denote the annular cover corresponding to α. Explicitly, S_α is a covering surface of S corresponding to the (conjugacy class of the) cyclic subgroup of $\pi_1(S)$ generated by α. Since S has negative Euler characteristic, it carries a hyperbolic metric which lifts to a hyperbolic metric on the annulus S_α. In particular, S_α has a natural boundary compactifying it to a closed annulus.

Let β be a simple closed curve or essential arc on S intersecting α. Consider the set of lifts $\widetilde{\beta}$ of β to S_α which connect the two boundary components of S_α. Every element of this set defines a vertex in the arc complex of the annulus S_α. We call the set of all these vertices the *subsurface projection of β to α*. The subsurface projection of β to α has diameter at most one as all lifts of β to S_α are disjoint.

Definition 2.A.2. The *marking graph* of S is the graph whose vertex set is the set of isotopy classes of markings. Two such markings μ and μ' are joined by an edge of length one if they differ by an *elementary move*. An elementary move from μ to μ' is one of the following two operations.

i) μ' has the same underlying pants decomposition as μ. The transversals of μ' are obtained from the ones of μ by applying one primitive Dehn twist about one of the pants curves.

ii) Replace a pants curve α by its corresponding clean transversal β in μ. Then modify α to a clean transversal of β ("cleaning the marking" in the terminology of [MM00]).

The cleaning operation is described in detail in [MM00, Lemma 2.4] (also compare the discussion on page 21 of [MM00]).

Since the details are not relevant for the current work, we do not review them here. The marking graph is a connected, locally finite graph on which the mapping class group of S acts with finite point stabilizers and finite quotient (compare [MM00]). Therefore, it is quasi-isometric to the mapping class group.

The following proposition is well-known to experts and relates distances in the marking graph to intersection numbers. Since we did not find a proof in the literature, we include one here for completeness.

Proposition 2.A.3. *Let μ_1, μ_2 be markings of a surface S. If μ_1 and μ_2 are of distance k in the marking graph, then the total number of intersections between μ_1 and μ_2 is bounded exponentially in k. Conversely, the total intersection number between μ_1 and μ_2 is a coarse upper bound for the distance between μ_1 and μ_2 in the marking graph of S.*

Proof. We begin with the lower bound for the distance in the marking graph. Let μ_1 and μ_2 be two markings. For a number $\epsilon > 0$, we say a marked Riemann surface X belongs to the ϵ-thick part of Teichmüller space if the length of each simple closed geodesic on X is at least ϵ. We will simply speak of the thick part, if the corresponding ϵ is understood from the context. There are points X_i the ϵ-thick part of Teichmüller space for S such that each curve in μ_i is shorter than some universal constant C on X_i. Here, ϵ is a universal constant depending only on the genus of the surface S. Explicitly, let P_i be the underlying pants decomposition of the marking μ_i. The pants decomposition P_i defines Fenchel-Nielsen coordinates for the Teichmüller space of S. This implies that there is a marked Riemann surface X_i' such that each curve in P_i has hyperbolic length 1 on X_i'. On a hyperbolic pair of pants all of whose boundary components have lengths equal one the distance between any two boundary components is uniformly bounded. This implies that on X_i' there are clean transversals to P_i whose hyperbolic length is also uniformly bounded. By changing the marking on X_i' by Dehn twists about P_i we obtain the desired surfaces X_i.

If the distance between μ_1 and μ_2 in the marking graph is bounded by k, then the Teichmüller distance between X_1 and X_2 is also coarsely bounded by k since the mapping class group acts properly and cocompactly on the thick part of Teichmüller space. Thus the total hyperbolic length of μ_2 on X_1 is bounded by $e^{2k} \cdot C$ by Wolpert's lemma ([W79, Lemma 3.1]). But each curve in μ_1 has a collar of definite width on X_1 since its length is bounded by C, and therefore the total number of intersections of μ_1 and μ_2 is also coarsely bounded by e^{2k}.

Next we show the upper bound for the distance in the marking graph. In the proof we will use singular Euclidean structures as in [B06] and the relation between the mapping class group of a surface and the corresponding Teichmüller space.

Let P_1 and P_2 be the underlying pants decompositions of the markings μ_1, μ_2. We may assume that $P_1 \cup P_2$ fills the surface, i.e. that all components of $S \setminus (P_1 \cup P_2)$ are simply connected. Namely, if $P_1 \cup P_2$ does not fill, then P_1 and P_2 share a common curve α. We can then change the transversal to α in μ_1 such that the diameter of the subsurface projection to α of the transversals to α in μ_1 and μ_2 is at most one. The number of steps necessary for this modification is bounded by the intersection number between the two transversals. We can then pass to the subsurface obtained by cutting S along the common curve α and discarding the corresponding transversal. Repeat this procedure until $P_1 \cup P_2$ fills.

Furthermore, we can assume that the twist about a pants curve $\delta \in P_1$ defined by μ_1 coarsely agrees with the one defined by P_2. By this we mean the following. Since P_1 and P_2 fill the surface, there is at least one curve of P_2 which intersects δ. Denote by c_δ the transversal to δ in μ_1. The diameter of the subsurface projection of P_2 and c_δ to δ is bounded from above by the intersection number between μ_1 and μ_2. Hence, after modifying the transversal to δ in μ_1 by at most $i(\mu_1, \mu_2)$ Dehn twists about δ, the diameter of the projection is at most 3. Similarly, we modify μ_2 such that the twist about the pants curves in P_2 given by μ_2 agrees with the one defined by P_1.

For a pair of measured laminations λ_1, λ_2 which jointly fill the surface and satisfy $i(\lambda_1, \lambda_2) = 1$ we denote by $q(\lambda_1, \lambda_2)$ the quadratic differential whose horizontal measured lamination is λ_1 and whose vertical measured lamination is λ_2. Now let ρ be the Teichmüller geodesic defined by P_1 and P_2; that is $\rho_t = q(e^{-t}P_1, e^t/i(P_1, P_2)P_2)$ (compare the construction in [B06] for pairs of curves). Recall that on every hyperbolic surface of genus g there is a pants decomposition such that the hyperbolic length of each pants curve is bounded by a universal constant B (the Bers constant) which depends only on the genus. By the collar lemma, a curve whose hyperbolic length is bounded by B has extremal length coarsely bounded by B. Thus the length of such a curve in any singular Euclidean metric in the same conformal class is bounded by a universal constant B'.

We set $T = \log(2B')$. Then for the singular Euclidean metric defined by ρ_{-T}, a curve whose length is smaller than B' cannot intersect P_1. Hence, P_1 is the only Bers short pants decomposition for ρ_{-T}. Similarly, P_2 is the only Bers short pants decomposition on $\rho_{\log(i(P_1,P_2))+T}$. In particular, there are two points X_1, X_2 in Teichmüller space, whose Teichmüller distance is bounded by $2T + \log(i(P_1, P_2))$ and such that P_i is Bers short on X_i.

Now for any k which is sufficiently large, by work of Rafi we have the following estimate for the Teichmüller distance $d_{\mathcal{T}}(X_1, X_2)$ (compare [R07, Equation (19)]).

$$d_{\mathcal{T}}(X_1, X_2) \succ \sum_{Y} [d_Y(\mu'_1, \mu'_2)]_k + \sum_{\alpha \notin \Gamma} \log [d_\alpha(\mu'_1, \mu'_2]_k .$$

Here, μ'_1 and μ'_2 are shortest markings on X_1 and X_2, respectively, and $[x]_k$ is a cutoff function which is 0 if $x \leq k$ and x otherwise. The expression $a \succ b$ means that a is coarsely bounded by b. The first sum is taken over all subsurfaces $Y \subset S$, while the indexing set Γ of the second sum is the set of (isotopy classes of) simple closed curves which are short on either X_1 or X_2. Note that in our case Γ agrees with the union of the pants curves in P_1 and P_2. In both cases d_Y (or d_α) denotes the diameter of the set of subsurface projections of μ'_1 and μ'_2 to Y (or α).

In our case, since P_1 and P_2 fill, we can replace the subsurface projections of μ'_i by those of P_i, except maybe in the cases where the subsurface is bounded by curves contained in Γ. Hence we get

$$d_{\mathcal{T}}(X_1, X_2) \succ \sum_{\partial Y \not\subset \Gamma} [d_Y(P_1, P_2)]_k + \sum_{\alpha \notin \Gamma} \log [d_\alpha(P_1, P_2)]_k .$$

Now, since $d_{\mathcal{T}}(X_1, X_2) \prec \log(i(P_1, P_2))$ we have

$$i(P_1, P_2) \succ \sum_{\partial Y \not\subset \Gamma} [d_Y(P_1, P_2)]_k + \sum_{\alpha \notin \Gamma} [d_\alpha(P_1, P_2)]_k .$$

Since the number of subsurfaces whose boundary is completely contained in Γ is uniformly bounded, and the total intersection of μ_1 and μ_2 bounds each of these projections, we get

$$i(\mu_1, \mu_2) \succ \sum_{Y} [d_Y(\mu_1, \mu_2)]_k + \sum_{\alpha} [d_\alpha(\mu_1, \mu_2)]_k .$$

where now the sums are taken over all subsurfaces and all curves respectively. By [MM00, Theorem 6.12], the right hand side of this inequality is coarsely equal to the distance of μ_1 and μ_2 in the marking graph. This shows the first claim. □

In the proof of the upper bound on distortion of the handlebody group the following corollary is used in an essential way.

Corollary 2.A.4. *Let $N > 0$ be given. Let C be a cell decomposition of the surface S with at most N cells. Let $f \in \mathrm{Map}(S)$ be arbitrary. The intersection number between C and $f(C)$ is coarsely bounded by an exponential of the word norm of f. Here, the constants depend on the genus of S and the number N.*

Similarly, let C and C' are cell decomposition with at most N cells and which intersect in K points. Then there is a mapping class g whose word norm is bounded coarsely in K, and such that $g(C)$ and C' intersect in uniformly few points.

Proof. Note that up to the action of the mapping class group there are only finitely many cell decompositions C of S with at most N cells. Hence, there is a constant $K > 0$ such that for any such cell decomposition C there is a marking μ_C whose intersection number with C is bounded by K.

By the preceding Proposition 2.A.3 the number of intersections between μ_C and $f(\mu_C)$ is coarsely bounded exponentially in the word norm of f. Since the intersection number between $f(\mu_C)$ and $f(C)$ is uniformly bounded, the corollary follows.

Similarly, if C and C' intersect in K points, then the intersection number between μ_C and $\mu_{C'}$ can be coarsely bounded by K. Hence, Proposition 2.A.3 implies the second claim of the corollary. □

Chapter 3

Distorted stabilizers in the handlebody group

3.1 Introduction

Let V_g be a handlebody of genus $g \geq 3$. In this chapter we consider the stabilizer of certain type of curve in the handlebody group. To describe the class of curves we are interested in, let D_1 be a properly embedded, nonseparating disk in V. Let α_1 be a simple closed curve which intersects ∂D_1 in a single point. We call such curves *V_g–primitive*, since they define conjugacy classes of primitive elements in the fundamental group of the handlebody V_g (see Lemma 3.5.4).

The main result of this chapter is the following

Theorem 3.1.1. *Let V_g be a handlebody of genus $g \geq 3$. Then the stabilizer of a V_g–primitive curve in* $\mathrm{Map}(V_g)$ *is exponentially distorted.*

The relevance of this theorem stems from the following observation. Instead of considering the stabilizer of α in the handlebody group, we could consider the stabilizer of α in the full mapping class group of the boundary surface ∂V_g. Here, the stabilizer of α is undistorted (see [HM10] for a complete proof of this fact, which is an easy consequence of the work in [MM00]). Hence, by Theorem 3.1.1 the extrinsic geometry of curve stabilizers is different in the handlebody group and in surface mapping class groups.

On the other hand, the analog of Theorem 3.1.1 is true for outer automorphism groups of free groups. By work of Handel and Mosher [HM10], the stabilizer of the conjugacy class of a free factor of rank 1 in a free group of rank ≥ 3 is exponentially distorted. The stabilizer of α in the handlebody group projects (via the action on the fundamental group) to the stabilizer of such a free factor.

In fact, this geometric property of $\mathrm{Out}(F_g)$ is one of the two main ingredients to the proof of Theorem 3.1.1. To transfer it to the handlebody group, we explicitly construct elements of the handlebody group which project in $\mathrm{Out}(F_g)$ to the distorted sequence that Handel and Mosher consider.

The upper distortion bound in Theorem 3.1.1 follows using a geometric model for the handlebody group, namely the graph of rigid racks (see [HH11a]). We identify the stabilizer of a V–primitive curve α as a suitable subgraph of the graph of racks. Then one can employ a surgery method, together with an estimate of intersection numbers, to obtain the upper distortion bound.

This chapter is organized as follows. In Section 3.2 we recall disk and point pushing maps for surface mapping class groups and handlebody groups. These form the basic building blocks for the construction of Nielsen twists in the handlebody group in Section 3.3. In Section 3.4 we use these elements to show the lower bound on distortion. Section 3.5 finishes the proof of Theorem 3.1.1 by proving the upper distortion bound.

3.2 Point and disk pushing homeomorphisms

Let S^2_{g-1} be a closed surface of genus $g-1 \geq 1$ with two disjoint marked disks D^+, D^-. We use the convention that both isotopies and homeomorphisms of surfaces and handlebodies fix the set of marked points setwise and each marked disk pointwise. Let $S_{g-1,2}$ be the surface of genus $g-1$ with two marked points x^+, x^- obtained by collapsing each of the marked disks of S^2_{g-1} to a point. It is well-known (compare e.g. [FM11, Proposition 3.19]) that there is a short exact sequence relating the mapping class groups of S^2_{g-1} to the pure mapping class group of $S_{g-1,2}$ as follows. By the *pure mapping class group* PMap we mean the subgroup of the mapping class group that fixes each marked point.

$$1 \to \mathbb{Z}^2 \to \mathrm{Map}(S^2_{g-1}) \to \mathrm{PMap}(S_{g-1,2}) \to 1 \tag{3.1}$$

Here, the kernel $\mathbb{Z}^2$ is generated by Dehn twists about the marked disks D^+ and D^-. Next, we describe how the mapping class group of a surface with marked points changes when a marked point is removed.

This is described by the Birman exact sequence. For its formulation in our context, let $x = x^{\pm}$ be one of the marked points of $S_{g-1,2}$ and let $S_{g-1,1}$ be the surface obtained from $S_{g-2,2}$ by forgetting the marked point x.

Theorem 3.2.1 (compare [FM11, Theorem 4.6] and [Bi74])**.** *There is a short exact sequence sequence*

$$1 \to \pi_1(S_{g-1,1}, x) \to \mathrm{PMap}(S_{g-1,2}) \to \mathrm{Map}(S_{g-1,1}) \to 1.$$

The map $\mathcal{P} : \pi_1(S_{g-1,1}, x) \to \mathrm{Map}(S_{g-1,2})$ is the so-called *point-pushing map* and can be constructed as follows. Let $\gamma : [0,1] \to S_{g-1,1}$ be a loop based at x. Choose an isotopy F_t of $S_{g-1,1}$ with the property that $F_t(x) = \gamma(t)$ and $F_0 = \mathrm{id}$. The endpoint F_1 of this isotopy is a homeomorphism of $S_{g-1,2}$, which we call the point pushing homeomorphism $P(\gamma)$ along γ. We let $\mathcal{P}(\gamma)$ be the image of $P(\gamma)$ in the mapping class group of $S_{g-1,2}$. In fact, $\mathcal{P}(\gamma)$ only depends on the homotopy class of γ, and thus defines a homomorphism $\mathcal{P} : \pi_1(S_{g-1,1}, x) \to \mathrm{PMap}(S_{g-1,2})$.

We call a preimage of $\mathcal{P}(\gamma)$ in $\mathrm{Map}(S^2_{g-1})$ under the map in the sequence (3.1) a *disk pushing mapping class*.

It is well-known that point pushing homeomorphisms act on the fundamental group by conjugations (compare e.g. [FM11, page 247]). The following lemma establishes a similar property for the action on arcs connecting the two marked points.

Lemma 3.2.2. *Let $S_{g-1,2}$ be a surface with two marked points x^+, x^-, and let $\alpha : [0,1] \to S_{g-1,2}$ be an arc connecting x^- to x^+. Let γ be a loop based at x^+ and let $P(\gamma)$ be the corresponding point-pushing homeomorphism.*

Then the arc $P(\gamma)(\alpha)$ is homotopic on S_{g-1} relative to its endpoints to the arc obtained by concatenating α and γ.

Proof. Suppose first that γ is an embedded loop. Then there is a regular tubular neighborhood U of the loop γ which is orientation preserving homeomorphic to $S^1 \times [-1,1]$ such that $S^1 \times \{0\}$ is the loop γ. Up to isotopy, the intersection of α with this neighborhood is a disjoint union of arcs, one of which has the form $\{s_0\} \times [-1,0]$ for a point $s_0 \in S^1$ and possibly others of the form $\{s_i\} \times [-1,1]$ for $s_1, \ldots, s_k \in S^1$.

The point pushing homeomorphism along γ can be chosen to be supported in the tubular neighborhood U, and there takes the form of a left Dehn twist on $S^1 \times [-1, 0]$ followed by a right Dehn twist on $S^1 \times [0, 1]$ (compare Fact 4.7 of [FM11]). From this the claim is immediate. (see also the discussion in the proof of Lemma 4.2 of [HH11a]).

The general case follows since every loop γ can be written as a concatenation of embedded loops. □

At this point we want to emphasize that the image arc $P(\gamma)(\alpha)$ is in general not homotopic to the concatenation of α and γ as an arc on $S_{g-1,2}$, since the homotopy constructed in the proof above may pass through the marked points.

The whole discussion so far is in fact compatible with handlebody groups. Namely, suppose that S^2_{g-1} is the boundary of a handlebody V^2_{g-1} of genus $g-1$ with two marked disks on its boundary. Then the Dehn twist about each of the marked disks is contained in the handlebody group of V^2_{g-1}. Thus, the short exact sequence (3.1) implies that there is a short exact sequence

$$1 \to \mathbb{Z}^2 \to \mathrm{Map}(V^2_{g-1}) \to \mathrm{Map}(V_{g-1,2}) \to 1$$

where $V_{g-1,2}$ is the handlebody of genus $g-1$ with two marked points on its boundary surface obtained by collapsing the two marked disks on ∂V^2_{g-1} to points.

Point pushing homeomorphisms extend to the interior of the handlebody (to see this, simply choose the defining isotopy to be an isotopy of the handlebody and not just the boundary surface). Hence, Theorem 3.2.1 implies that there is a Birman exact sequence for handlebody groups:

$$1 \to \pi_1(S_{g-1,1}, x) \to \mathrm{PMap}(V_{g-1,2}) \to \mathrm{Map}(V_{g-1,1}) \to 1$$

Because of these two sequences, disk pushing mapping classes are contained in the handlebody groups.

3.3 Nielsen twists in the handlebody group

The main technical step in lower distortion bound of Theorem 3.1.1 is an explicit construction of elements in the handlebody group which act on the fundamental group in an easy fashion.

To be describe the outer automorphisms we will use, let F_g be the free group on g generators, and let $e_1, \ldots, e_g$ be a generating set. Pick one of these generators, say e_g, and consider a word G in $e_1, \dots, e_{g-1}$. Then there is an automorphism of F_g acting on the basis (e_i) as follows:

$$\begin{array}{rcl} e_1 & \mapsto & Ge_1 \\ e_2 & \mapsto & e_2 \\ & \dots & \\ e_g & \mapsto & e_g \end{array}$$

For the purpose of this chapter, we call such an automorphism (and its image in the outer automorphism group) a *Nielsen twist*.

Now let V_g be a handlebody if genus g, and let D be an essential nonseparating disk in V_g. The complement U of D in V_g is a handlebody of genus $g-1$ with two disks removed from its boundary. Let V' be a handlebody of genus $g-1$ with two distinguished disks D^+ and D^- on its boundary. The inclusion of U in V induces a continuous map $F : V' \to V$ which maps both D^+ and D^- to D, and which induces a homeomorphism of the complement of D^+ and D^- to U. We will consider V' as if it were a submanifold of V and often call it the complement of D in V if no confusion can occur. Note that the homeomorphism F induces a homomorphism F_* of the handlebody group of V' into the handlebody group of V.

We next describe a convenient loop system on ∂V_g which generates the fundamental group of V_g. Let $\Sigma = \{D_1, \dots, D_g\}$ be a *cut system* of V_g, i.e. a collection of pairwise disjoint, nonisotopic disks which cut V_g into a spotted ball. Since D is nonseparating, we may assume that $D_1 = D$.

We choose a base point $y \in \partial D$. Let $e_1, \dots, e_g$ be a collection of disjointly embedded loops on ∂V_g with the following properties. Suppose that e_1 approaches D from two different sides at its two endpoints, and that all other e_i, $i \neq 1$ approach D from the same side at both of their endpoints. Furthermore, the loop e_i is disjoint from D_k for all $k \neq i$ (except possibly at the basepoint if $k = 1$) and the loop e_i intersects the disk D_i in a single point. Since Σ decomposes V_g into a ball, such a collection of loops generate $\pi_1(V_g, y)$.

The preimages of the $e_2, \dots, e_g$ under F thus are loops on V' based at one of the distinguished disks, say D^+. The preimage of e_1 under F is an arc connecting the marked disk D^- to the marked disk D^+. If no confusion can occur, we denote these preimages by the same symbol. Let $\widetilde{V}$ be the handlebody of genus $g-1$ with two marked points x^+, x^- obtained by collapsing

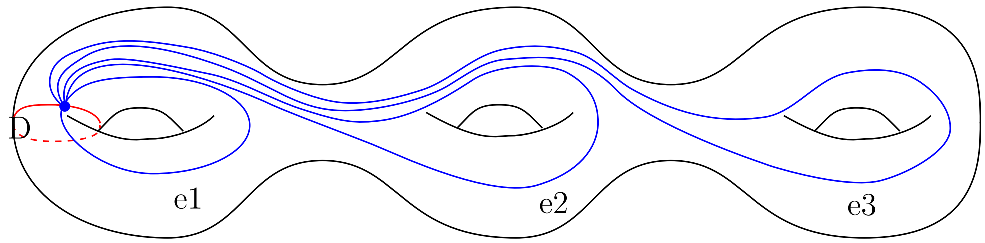


Figure 3.1: The setup for Lemma 3.3.1

the marked disk $D^{\pm}$ of V' to the point $x^{\pm}$. The loops $e_2, \ldots, e_g$ project to loops based at x^+ on $\widetilde{V}$, while e_1 projects to an arc connecting x^- to x^+ (up to possibly reversing the orientation of e_1). Again, we will denote these loops and arcs by the same symbol, if the context is clear.

The next lemma is the main step in understanding the action of F_*-images of disk pushing mapping classes on the fundamental group of V_g.

Lemma 3.3.1. *Let γ be a loop on $\partial V'$ based at a point y on D^+. Let $[\gamma] \in \pi_1(S_{g,1}, x^+)$ be the homotopy class defined by the projection of γ to $\widetilde{V}$.*

Let φ be a disk-pushing map defined by γ on $S_g^2 = \partial V'$. Then the map $F_*\varphi$ has the following properties.

i) $F_*\varphi(D) = D$.

ii) $F_*\varphi(e_i)$ is homotopic on ∂V to the loop obtained by conjugating e_i by $[\gamma]$, for $i = 2, \ldots, g$.

iii) $F_*\varphi(e_1)$ is homotopic in V' to the loop obtained by concatenating e_1 and γ relative to its endpoints.

Proof. The first property is immediate from the definition of F_*. Property *ii)* follows since point pushing maps act by conjugation on the fundamental group (compare e.g. [FM11, page 247]). Property *iii)* is a consequence of Lemma 3.2.2. There is a slight subtlety hidden here, as Lemma 3.2.2 only states that the arc $\varphi(e_1)$ is homotopic to the concatenation of α and $[\gamma]$ on S_{g-1}. In particular, such a homotopy may pass over the marked points. Therefore, the arc defined by $\varphi(e_1)$ on $\widetilde{V}$ is homotopic to the concatenation of e_1 and γ only if we allow the arc to pass through the marked disks D^+ and D^-. However, since we are only interested in the homotopy class of the of the loop $F_*\varphi(e_1)$ in the handlebody and not on its boundary surface, this is not an issue. ∎

The following corollary is immediate from Lemma 3.3.1.

Corollary 3.3.2. *The map $F_*\varphi$ from Lemma 3.3.1 maps in* $\mathrm{Out}(F_n)$ *to the image of a Nielsen twist*

$$\begin{array}{ccc} e_1 & \mapsto & Ge_1 \\ e_2 & \mapsto & e_2 \\ & \dots & \\ e_n & \mapsto & e_n \end{array}$$

where G is the image of γ in the fundamental group $\pi_1(V_g, y)$ of the handlebody.

3.4 The lower distortion bound

In this section we prove the lower distortion bound for the curves described in the introduction. As a first step, we note the following lemma, which motivates the notation V_g–primitive curve.

Lemma 3.4.1. *Let V_g be a handlebody of genus g and let α be a V_g–primitive curve. Then the projection of α in $\pi_1(V)$ defines the conjugacy class of a primitive element.*

Proof. Let D_1 be the nonseparating disk which intersects α in a single point. Put δ_1 to be the boundary of a regular neighborhood of $\partial D_1 \cup \alpha_1$. The curve δ_1 bounds a separating essential disk D' in V_g. The disk D' cuts V_g into a spotted solid torus T and a handlebody V' of genus $g-1$. The curve α_1 defines the conjugacy class of a generator of $\pi_1(T)$. Complete D_1 to a cut system $\Sigma = \{D_1, \dots, D_g\}$ of V_g by adding $g-1$ disks which are contained in V'. Let $p \in \partial V_g$ be a basepoint on the boundary of the handlebody. Then there is a basis of $\pi_1(V_g, p)$ defined by disjointly embedded loops $\gamma_1, \dots, \gamma_g$ on ∂V_g based at p with the following properties. The loop γ_1 is freely homotopic to α_1. Furthermore, a $\gamma_i, i > 1$ is disjoint from $D_j, j \neq i$ and intersects D_i in one point. Hence α defines the conjugacy class of a primitive element. □

We call a basis of loops for $\pi_1(V_g, p)$ as in the previous proof a *dual basis to* Σ (or a *dual basis to an extension of D*).

The goal of this section is to prove the following:

Proposition 3.4.2. *Let $g \geq 3$ and V_g be a handlebody of genus g. Let α be a V_g–primitive curve. Then the stabilizer of the homotopy class of α is at least exponentially distorted in the handlebody group.*

The proof of this theorem uses the work of Handel and Mosher on distorted stabilizers of (conjugacy classes of) primitive elements in $\mathrm{Out}(F_g)$ to show the lower bound on distortion.

We begin by reviewing the construction from Section 4.3 (Case 1) of [HM10] that we will use.

Let $\gamma_1, \ldots, \gamma_g$ be a dual basis of loops to an extension Σ of D as defined above. We denote by $e_1, \ldots, e_g$ the images of the γ_i in $F_g = \pi_1(V_g, p)$. Handel and Mosher consider the following automorphism $\Theta : F_g \to F_g$:

$$\begin{aligned}
\Theta(e_1) &= e_1 e_2 \\
\Theta(e_2) &= e_1 \\
\Theta(e_i) &= e_i \quad \text{for } 3 \leq i \leq g
\end{aligned}$$

and a sequence of automorphisms $\Phi_k : F_g \to F_g$ defined as

$$\begin{aligned}
\Phi_k(e_i) &= e_i \quad \text{for } i < n \\
\Phi_k(e_n) &= e_n \Theta^k(e_1)
\end{aligned}$$

By slight abuse of notation we will denote the images of Θ and Φ_k in the outer automorphism group of F_g by the same symbols.

The following lemma is proven in the discussion of Case 1 of Section 4.3 of [HM10].

Lemma 3.4.3. *Let G be the stabilizer of the conjugacy class of e_1 in the outer automorphism group of F_g. Then the word norm $\|\Phi_k\|_G$ of Φ_k as an element of G grows exponentially with k.*

We want to construct suitable mapping classes φ_k of the handlebody V_g inducing Φ_k on the fundamental group level.

Let δ_i be the boundary of a regular neighborhood of $\gamma_i \cup \partial D_i$. For each $1 \leq i \leq g$, the curve δ_i is essential, separating and diskbounding in V_g.

We first choose a homeomorphism $f : V_g \to V_g$ which fixes δ_g, α_g and ∂D_g pointwise and which induces Θ on the fundamental group level. To see that this is possible, observe that the disk bounded by δ_g decomposes V_g into two complementary components. Let V_{g-1} be the component not containing γ_g

and $\hat{V}$ be the other component. The inclusion $V_{g-1} \to V_g$ induces an isomorphism of the fundamental group of V_{g-1} onto the subgroup of F_g generated by $\alpha_1, \ldots, \alpha_{g-1}$. Since the homomorphisms from the handlebody group to the outer automorphism group of the fundamental group is onto, we can choose a homeomorphism $f' : V_{g-1} \to V_{g-1}$ which induces the restriction of Θ to $\langle e_1, \ldots, e_{g-1} \rangle$.

Now extend this homeomorphism to $\hat{V}$ by the identity to obtain the desired map f. Let ϑ be the mapping class defined by f.

To define φ_k we use the construction of maps in the handlebody group described in Section 3.3. To stay consistent with the notation used in that section, put $D = D_g$ and choose a basepoint $p \in \partial D_g$.

Since the homeomorphism f fixes the disk D_g, it induces a homeomorphism of V' to itself, which we denote by the same letter.

Let V' be the complement (as defined in Section 3.3) of D_g and let D^+ and D^- be the two distinguished spots on $\partial V'$. Let $p \in D^+$ be a base point and let γ be a loop based at p such that the image of γ in $\pi_1(V_g)$ is conjugate to e_1.

Let $F : V' \to V'$ be a disk pushing homeomorphism defined by γ. Then $F^{(k)} := f^k \circ F \circ f^{-k}$ is a disk pushing homeomorphism along $f^k(\gamma_1)$ on V'.

Define φ_k to be the mapping class V_g obtained from $F^{(k)}$ by regluing the two spots D^+ and D^- as described in Section 3.3.

In the handlebody group, $\varphi_k = \vartheta^k \circ \varphi_0 \circ \vartheta^{-k}$ by construction. Hence, the word length of φ_k grows linearly in k by the triangle inequality.

By Corollary 3.3.2, the element φ_k induces a Nielsen twist on the level of fundamental groups, appending the image of $\Theta^k(e_1)$ to e_g. In particular, φ_k induces Φ_k on $\pi_1(V_g, y)$ and fixes the simple closed curve α_1 (this follows from the fact that φ_k acts as conjugation on the loop γ_1 on the boundary surface).

Hence the word norm of φ_k in the stabilizer of α_1 in the handlebody group is coarsely bounded from below by the word norm of Φ_k in the stabilizer of e_1 in the outer automorphism group of the free group. By Lemma 3.4.3 the latter grows exponentially. On the other hand, the word norm of φ_k in the handlebody group grows linearly. As a consequence, Proposition 3.4.2 follows

3.5 The upper distortion bound

In this section we show that stabilizer of an V_g–primitive curve is at most exponentially distorted in the handlebody group, completing the proof of Theorem 3.1.1.

To do so, we employ the results of [HH11a]. In particular, we use the graph of rigid racks as a geometric model of the handlebody group. For convenience, we recall the basic definitions and facts here.

Definition 3.5.1. A *rack* R is a cut system $\Sigma(R)$ (called the *support system*) together with a collection of pairwise disjoint arcs on $\partial V \setminus \partial\Sigma(R)$ with endpoints on $\partial\Sigma(R)$ (called the *ropes*) such that on each side of each support disk there is at least one rope with an endpoint approaching the support disk on that side. A rack R is called *large* if $\partial\Sigma(R)$ together with the collection of ropes decomposes ∂V into disks. The cell decomposition of ∂V defined by these curves and arcs is called the cell decomposition *induced by the rack R*. Two large racks are called *rigidly isotopic* if the cell decompositions of ∂V induced by them are isotopic.

The importance of racks stems from the following construction which is inspired by the train track graph for closed surfaces.

Definition 3.5.2. The graph of rigid racks is the graph whose vertex set is the set of rigid isotopy classes of large racks. Two such vertices are connected by an edge, if the cell decompositions induced by the corresponding racks intersect in at most K points (Here, K is a uniform constant chosen in such a way that the resulting graph is connected; compare [HH11a, Lemma 7.3] for details).

By Corollary 7.2 of [HH11a], the handlebody group acts on the graph of rigid racks properly discontinuously and cocompactly as a group of automorphisms. Hence, by the Svarc-Milnor lemma it is quasi-isometric to the graph of rigid racks.

As a next step, we define a geometric model of the stabilizer of α in the graph of rigid racks.

Definition 3.5.3. Let α be a V_g–primitive curve. We say that a rack R is *adapted to α* if the following holds. The support system of R contains a disk D which intersects α in a single point. Furthermore, R contains a rope which is freely homotopic to α.

Lemma 3.5.4. *The stabilizer of α in the handlebody group acts on the set of isotopy classes of racks which are adapted to α with finite quotient.*

Proof. As a first step, note that the stabilizer of α acts transitively on the set of isotopy classes of reduced disk systems which intersect α in a single point. Namely, let Σ and Σ' be two such reduced disk systems, and let $D \in \Sigma$ (respectively $D' \in \Sigma'$) be the disks which intersect α.

The complements of $\partial\Sigma \cup \alpha$ and $\partial\Sigma' \cup \alpha$ on ∂V_g are homeomorphic as surfaces. Therefore, there is is a mapping class of the boundary surface ∂V_g which fixes α and maps $\partial\Sigma$ to $\partial\Sigma'$. Since every homeomorphism of the boundary of a spotted ball extends to the interior of the ball, such a mapping class is contained in the handlebody group.

Hence, it suffices to show the statement of the lemma for racks adapted to α which have a common support system. The stabilizer of $\partial\Sigma$ in the mapping class group is contained in the handlebody group. Now the claim is an immediate consequence of the fact that there are only finitely many cell decompositions of a surface with uniformly few cells up to the action of the mapping class group. ☐

Lemma 3.5.4 allows to estimate the word norm of an element in the stabilizer of α in the handlebody group. Namely, let φ be such a handlebody group element. Suppose that there is an edge-path in the graph of rigid racks connecting R_0 to $\varphi(R_0)$, where R_0 is adapted to α. Furthermore, assume that we can arrange the edge-path in such a way that every vertex corresponds to a rack which is adapted to α. Then the length of such a path is a coarse upper bound for the word norm of φ in the stabilizer of α.

This strategy is executed in the following proposition, which finishes the proof of Theorem 3.1.1.

Proposition 3.5.5. *The stabilizer of the free homotopy class of α in the handlebody group is at most exponentially distorted.*

Proof. Let D be a nonseparating disk intersecting the simple closed curve $\alpha \subset \partial V_g$ in one point. Extend $D = D_1$ to a reduced disk system $\Sigma_0 = \{D_1, \dots, D_g\}$ which is disjoint from α (see the discussion at the beginning of Section 3.4). We further choose simple closed curves $\alpha_1, \dots, \alpha_g$ which are pairwise disjoint, and such that $\alpha_1 = \alpha$ and the geometric intersection number between α_i and ∂D_j is 1 if $i = j$ and 0 otherwise.

Let δ_k be the boundary of a regular neighborhood of $\alpha_k \cup \partial D_k$. The curve δ_k bounds a disk in V. We let $\mathcal{D}_0$ be the disk system bounded by $\{\delta_1, \dots, \delta_g\}$.

By construction of the curves δ_k, the complement of $\mathcal{D}_0$ consists of g solid once-spotted tori and a spotted ball. As a consequence, Σ_0 is the unique reduced disk system which is disjoint from $\mathcal{D}_0$.

Let R_0 be a large rigid rack whose support system is Σ_0, which contains the curves α_i as a rope (i.e. R_0 is adapted to α) and which intersects $\mathcal{D}_0$ in uniformly few points.

Let φ be a handlebody group element fixing α. We will prove that there is a path in $\mathcal{RR}(V)$ connecting R_0 to $f(R_0)$ such that all racks occurring in this path are adapted to α. Furthermore, the length of this path can be bounded by an exponential function of the distance between R_0 and $f(R_0)$. By the discussion above, this suffices to show the proposition.

The existence of the desired path is a consequence of the proof of the upper distortion bound in [HH11a]. We repeat the construction here. Put $\mathcal{D} = f(\mathcal{D}_0)$ and $\Sigma = f(\Sigma_0)$. By Lemma 7.6 of [HH11a] there is a large rack R_1 which carries $\mathcal{D}'$ and such that the distance between R_0 and R_1 in the graph of rigid racks is bounded by the geometric intersection number between $\partial\mathcal{D}'$ and the one-skeleton of the cell decomposition induced by R_0.

Proposition A.3 from [HH11a] implies that this intersection number N_1 is bounded exponentially in the word norm of f in the handlebody group. Since α is disjoint from $\mathcal{D}'$ we can assume without loss of generality that R still contains α as a rope. In fact, the proof of Lemma 7.6 also yields that there is a path in the graph of rigid racks connecting R_0 to R_1 such that each rack in this path is adapted to α and such that the length of this path is also bounded by N_1.

As in the proof of Theorem 7.9 of [HH11a] we now apply Lemma 7.8 of [HH11a] repeatedly to obtain an edge-path of racks R_i, $i = 1, \dots, N$, such that $\mathcal{D}$ is disjoint from the support system of R_N. Again, since $\mathcal{D}$ is disjoint from α, all racks R_i can be chosen to be adapted to α.

By choice of $\mathcal{D}$, the reduced disk system Σ is the unique reduced disk system disjoint from $\mathcal{D}$. Hence, the rigid rack R_N has Σ as its support system and contains α as a rope. Thus, we can connect R_N to $f(R_0)$ with a path of rigid racks, all of which have Σ as their support system and contain α as a rope. The distance between R_N and $f(R_0)$ is at most $N_1 + N + d$ by the triangle inequality, where d is the distance between R_0 and $f(R_0)$.

The stabilizer of $\partial\Sigma$ is undistorted in the handlebody group and furthermore is equal to the stabilizer of $\partial\Sigma$ in the mapping class group of ∂V. Hence, the stabilizer of $\partial\Sigma \cup \alpha$ is also undistorted in the handlebody group.

This implies that the path connecting R_N to $f(R_0)$ as above can be chosen to have length coarsely bounded by $N_1 + N + d$. $\square$

Chapter 4

Isoperimetric inequalities for handlebody groups[1]

4.1 Introduction

A handlebody of genus $g \geq 2$ is a compact orientable 3-manifold V whose boundary ∂V is a closed surface of genus g. The *handlebody group* $\mathrm{Map}(V)$ is the group of isotopy classes of orientation preserving homeomorphisms of V. Via the natural restriction homomorphism, the group $\mathrm{Map}(V)$ can be viewed as a subgroup of the mapping class group $\mathrm{Map}(\partial V)$ of ∂V. This subgroup is of infinite index, and it surjects onto the outer automorphism group of the fundamental group of V which is the free group with g generators.

The handlebody group is finitely presented. Thus it can be equipped with a word norm that is unique up to quasi-isometry. Hence, the handlebody group carries a well-defined large-scale geometry. However, this large scale geometry is not compatible with the large-scale geometry of the ambient group $\mathrm{Map}(\partial V)$. Namely, we showed in [HH11a] (compare Chapter 2) that the handlebody group is an exponentially distorted subgroup of the mapping class group of the boundary surface for every genus $g \geq 2$. Here, a finitely generated subgroup $H < G$ of a finitely generated group G is called *exponentially distorted* if the following holds. First, the word norm in H of

[1]This chapter is identical with the preprint [HH11c]
Ursula Hamenstädt and Sebastian Hensel,
Isoperimetric inequalities for the handlebody groups,
arXiv:1109.5255

every element $h \in H$ can be bounded from above by an exponential function in the word norm of h in G. On the other hand, there is no such bound with sub-exponential growth rate.

As a consequence, it is not possible to directly transfer geometric properties from the mapping class group to the handlebody group. In this chapter we initiate an investigation of the intrinsic large-scale geometry of the handlebody group.

A particularly useful geometric invariant of a finitely presented group G is its *Dehn function*, which can be defined as the isoperimetric function of a presentation complex for G (see Section 4.4 for a complete definition). Although the Dehn function itself depends on the choice of a finite presentation of G, the growth type of the Dehn function does not. In fact, the growth type of the Dehn function is a quasi-isometry invariant of G.

The mapping class group $\mathrm{Map}(\partial V)$ of ∂V is automatic [Mo95] and hence has quadratic Dehn function. Since $\mathrm{Map}(V)$ is exponentially distorted in $\mathrm{Map}(\partial V)$, this fact does not provide any information on the Dehn function of $\mathrm{Map}(V)$. On the other hand, for $g \geq 3$ the Dehn function of the outer automorphism group $\mathrm{Out}(F_g)$ of a free group F_g on g generators is exponential [HV96, BV95, BV10, HM10]. However, since the kernel of the projection from the handlebody group to $\mathrm{Out}(F_g)$ is infinitely generated [McC85], this fact also does not restrict the Dehn function of $\mathrm{Map}(V)$.

The goal of this chapter is to give an upper bound for the Dehn function of $\mathrm{Map}(V)$. We show

Theorem 4.1.1. *The handlebody group* $\mathrm{Map}(V)$ *satisfies an exponential isoperimetric inequality, i.e. the growth of its Dehn function is at most exponential.*

The strategy of proof for Theorem 4.1.1 is similar to the strategy employed in [HV96] to show an exponential upper bound for the Dehn function of outer automorphism groups of free groups. We construct a graph which is a geometric model for the handlebody group (a similar construction is used in [HH11a] in order to show exponential distortion of handlebody groups). Vertices of this graph correspond to isotopy classes of special cell decompositions of ∂V containing the boundary of a *simple disk system* in their one-skeleton. A simple disk system is a collection of pairwise disjoint, pairwise non-homotopic embedded disks in V which decompose V into simply connected regions. We then use a a surgery procedure for disk systems to define a distinguished class of paths in this geometric model. Although these paths are in general not quasi-geodesics for the handlebody group (see the example at the end of this note), they are sufficiently well-behaved so that they can be used to fill a cycle with area bounded by an exponential function in the length of the cycle.

The organization of this chapter is as follows. In Section 2 we introduce disk systems and special paths in the disk system graph. Section 3 discusses a geometric model for the handlebody group. This model is used in Section 4 for the proof of Theorem 4.1.1.

4.2 Disk exchange paths

In this section we collect some facts about properly embedded disks in a handlebody V of genus $g \geq 2$. In particular, we describe a surgery procedure that is central to the construction of paths in the handlebody group.

A disk D in V is called *essential* if it is properly embedded and if ∂D is an essential simple closed curve on ∂V. A *disk system* for V is a set of pairwise disjoint essential disks in V no two of which are homotopic. A disk system is called *simple* if all of its complementary components are simply connected. It is called *reduced* if in addition it has a single complementary component.

We usually consider disks and disk systems only up to proper isotopy. Furthermore, we will always assume that disks and disk systems are in minimal position if they intersect. Here we say that two disk systems $\mathcal{D}_1, \mathcal{D}_2$ are in *minimal position* if their boundary multicurves intersect in the minimal number of points in their respective isotopy classes and if every component of $\mathcal{D}_1 \cap \mathcal{D}_2$ is an embedded arc in $\mathcal{D}_1 \cap \mathcal{D}_2$ with endpoints in $\partial\mathcal{D}_1 \cap \partial\mathcal{D}_2$. Note that minimal position of disks is not unique; in particular the intersection pattern $\mathcal{D}_1 \cap \mathcal{D}_2$ is not determined by the isotopy classes of $\mathcal{D}_1$ and $\mathcal{D}_2$.

The following easy fact will be used frequently throughout the chapter.

Lemma 4.2.1. *The handlebody group acts transitively on the set of isotopy classes of reduced disk systems. Every mapping class of ∂V that fixes the isotopy class of a simple disk system is contained in the handlebody group.*

Proof. The first claim follows from the fact that the complement of a reduced disk system in V is a ball with $2g$ spots and any two such manifolds are homeomorphic. The second claim is immediate since every homeomorphism of the boundary of a spotted ball extends to the interior. □

Let $\mathcal{D}$ be a disk system. An *arc relative to* $\mathcal{D}$ is a continuous embedding $\rho : [0,1] \to \partial V$ whose endpoints $\rho(0)$ and $\rho(1)$ are contained in $\partial\mathcal{D}$. An arc ρ is called *essential* if it cannot be homotoped into $\partial\mathcal{D}$ with fixed endpoints. In the sequel we always assume that arcs are essential and that the number of intersections of ρ with $\partial\mathcal{D}$ is minimal in its isotopy class.

Choose an orientation of the curves in $\partial\mathcal{D}$. Since ∂V is oriented, this choice determines a left and a right side of a component α of $\partial\mathcal{D}$ in a small annular neighborhood of α in ∂V. We then say that an endpoint $\rho(0)$ (or $\rho(1)$) of an arc ρ *lies to the right (or to the left) of* α, if a small neighborhood $\rho([0,\epsilon])$ (or $\rho([1-\epsilon,1])$) of this endpoint is contained in the right (or left) side of α in a small annulus around α. A *returning arc relative to* $\mathcal{D}$ is an arc both of whose endpoints lie on the same side of the boundary ∂D of a disk D in $\mathcal{D}$, and whose interior is disjoint from $\partial\mathcal{D}$.

Let E be a disk which is not disjoint from $\mathcal{D}$. An *outermost arc* of ∂E relative to $\mathcal{D}$ is a returning arc ρ relative to $\mathcal{D}$, with endpoints on the boundary of a disc $D \in \mathcal{D}$, such that there is a component E' of $E \setminus \mathcal{D}$ whose boundary is composed of ρ and an arc $\beta \subset D$. The interior of β is contained in the interior of D. We call such a disk E' an *outermost component* of $E \setminus \mathcal{D}$.

For every disk E which is not disjoint from $\mathcal{D}$ there are at least two distinct outermost components E', E'' of $E \setminus \mathcal{D}$. There may also be components of $\partial E \setminus \mathcal{D}$ which are returning arcs, but not outermost arcs. For example, a component of $E \setminus \mathcal{D}$ may be a rectangle bounded by two arcs contained in $\mathcal{D}$ and two subarcs of ∂E with endpoints on $\partial\mathcal{D}$ which are homotopic to a returning arc relative to $\partial\mathcal{D}$.

Let now $\mathcal{D}$ be a simple disk system and let ρ be a returning arc whose endpoints are contained in the boundary of some disk $D \in \mathcal{D}$. Then $\partial D \setminus \{\rho(0), \rho(1)\}$ is the union of two (open) intervals γ_1 and γ_2. Put $\alpha_i = \gamma_i \cup \rho$. Up to isotopy, α_1 and α_2 are simple closed curves in ∂V which are disjoint from $\mathcal{D}$ (compare [St99] for this construction). Therefore both α_1 and α_2 bound disks in the handlebody which we denote by Q_1 and Q_2. We say that Q_1 and Q_2 are obtained from D by *simple surgery along the returning arc* ρ.

The following observation is well known (compare [M86, Lemma 3.2], [St99] or [HH11a]).

Lemma 4.2.2. *If* Σ *is a reduced disk system and* ρ *is a returning arc with endpoints on* $D \in \Sigma$*, then for exactly one choice of the disks* Q_1, Q_2 *defined as above, say the disk* Q_1*, the disk system obtained from* Σ *by replacing* D *by* Q_1 *is reduced.*

The disk Q_1 is characterized by the requirement that the two spots in the boundary of $V \setminus \Sigma$ corresponding to the two copies of D are contained in distinct connected components of $H \setminus (\Sigma \cup Q_1)$. It only depends on Σ and the returning arc ρ. We call the interval γ_1 used in the construction of the disk Q_1 the *preferred interval defined by the returning arc*.

Definition 4.2.3. Let Σ be a reduced disk system. A *disk exchange move* is the replacement of a disk $D \in \Sigma$ by a disk D' which is disjoint from Σ and such that $(\Sigma \setminus D) \cup D'$ is a reduced disk system. If D' is determined as in Lemma 4.2.2 by a returning arc of a disk in a disk system $\mathcal{D}$ then the modification is called a *disk exchange move of* Σ *in direction of* $\mathcal{D}$ or simply a *directed disk exchange move*.

A sequence (Σ_i) of reduced disk systems is called a *disk exchange sequence in direction of* $\mathcal{D}$ (or *directed disk exchange sequence*) if each Σ_{i+1} is obtained from Σ_i by a disk exchange move in direction of $\mathcal{D}$.

The following lemma is an easy consequence of the fact that simple surgery reduces the geometric intersection number (see [HH11a] for a proof).

Lemma 4.2.4. *Let Σ_1 be a reduced disk system and let $\mathcal{D}$ be any other disk system. Then there is a disk exchange sequence $\Sigma_1, \dots, \Sigma_n$ in direction of $\mathcal{D}$ such that Σ_n is disjoint from $\mathcal{D}$.*

To estimate the growth rate of the Dehn function of the handlebody group we will need to compare disk exchange sequences starting in disjoint reduced disk systems. This is made possible by considering another type of surgery sequence for disk systems, which we describe in the remainder of this section.

To this end, let $\mathcal{D}$ be any simple disk system and let ρ be a returning arc. A *full disk replacement defined by* ρ modifies a simple disk system $\mathcal{D}$ to a simple disk system $\mathcal{D}'$ as follows. Let $D \in \mathcal{D}$ be the disk containing the endpoints of the returning arc ρ. Replace D by both disks Q_1, Q_2 obtained from D by the simple surgery defined by ρ. The disks Q_1, Q_2 are disjoint from each other and from $\mathcal{D}$. If one (or both) of these disks is isotopic to a disk Q from $\mathcal{D} \setminus D$ then this disk will be discarded (i.e. we retain a single copy of Q; compare [Ha95] for a similar construction). We say that a sequence $(\mathcal{D}_i)$ is a *full disk replacement sequence in direction of* $\mathcal{D}$ (or *directed full disk replacement sequence*) if each $\mathcal{D}_{i+1}$ is obtained from $\mathcal{D}_i$ by a full disk replacement along a returning arc contained in $\partial\mathcal{D}$.

The following two lemmas relate full disk replacement sequences to disk exchange sequences. Informally, these lemmas state that every directed disk exchange sequence may be extended to a full disk replacement sequence, and conversely every full disk replacement sequence contains a disk exchange sequence. To make this idea precise, we use the following

Definition 4.2.5. Let $\mathcal{D}$ be a disk system. Suppose that $\mathcal{D}_0, \dots, \mathcal{D}_n$ is a full disk replacement sequence in direction of $\mathcal{D}$ and that $\Sigma_1, \dots, \Sigma_k$ is a disk exchange sequence in direction of $\mathcal{D}$.

We say that the sequences $(\mathcal{D}_i)$ and (Σ_i) are *compatible*, if there is a non-decreasing surjective map $r : \{0, \dots, n\} \to \{1, \dots, k\}$ such that $\Sigma_{r(i)} \subset \mathcal{D}_i$ for all i.

Lemma 4.2.6. *Let Σ be a reduced disk system, let $\mathcal{D}$ be a simple disk system containing Σ and let $\mathcal{D} = \mathcal{D}_0, \mathcal{D}_1, \dots, \mathcal{D}_m$ be a full disk replacement sequence in direction of a disk system $\mathcal{D}'$. Then there is a disk exchange sequence $\Sigma = \Sigma_0, \Sigma_1, \dots, \Sigma_u$ in direction of $\mathcal{D}'$ which is compatible with (D_i).*

Proof. We proceed by induction on the length of the full disk replacement sequence $(\mathcal{D}_i)$. If this length equals zero there is nothing to show. Assume that the claim holds true whenever this length does not exceed $m - 1$ for some $m > 0$.

Let $\mathcal{D}_0, \dots, \mathcal{D}_m$ be a full disk replacement sequence of length m and let $\Sigma \subset \mathcal{D}_0$ be a reduced disk system. Let $D \in \mathcal{D}_0$ be the disk replaced in the full disk replacement move connecting $\mathcal{D}_0$ to $\mathcal{D}_1$.

If $D \in \Sigma$ then for one of the two disks obtained from D by simple surgery, say the disk D', the disk system $\Sigma_1 = (\Sigma \setminus D) \cup D'$ is reduced. However, $\Sigma_1 \subset \mathcal{D}_1$ and the claim now follows from the induction hypothesis.

If $D \notin \Sigma$ then $\Sigma \subset \mathcal{D}_1$ by definition and once again, the claim follows from the induction hypothesis. □

Lemma 4.2.7. *Let $\Sigma_0, \dots, \Sigma_m$ be a disk exchange sequence of reduced disk systems in direction of a disk system $\mathcal{D}'$. Then for every simple disk system $\mathcal{D}_0 \supset \Sigma_0$ there is a full disk replacement sequence $\mathcal{D}_0, \dots, \mathcal{D}_k$ in direction of $\mathcal{D}'$ which is compatible with (Σ_i).*

Proof. We proceed by induction on the length m of the directed disk exchange sequence.

The case $m = 0$ is trivial, so assume that the lemma holds true for directed disk exchange sequences of length at most $m - 1$ for some $m \geq 1$. Let $\Sigma_0, \dots, \Sigma_m$ be a directed disk exchange sequence of length m. Suppose Σ_1 is obtained from Σ_0 by replacing a disk $D \in \Sigma_0$. Let ρ be the returning arc with endpoints on D defining the disk replacement, and let D_1 be the disk in Σ_1 which is the result of the simple surgery.

We distinguish two cases. In the first case, $\rho \cap \mathcal{D}_0 = \rho \cap D$. Then ρ is a returning arc relative to $\mathcal{D}_0$. Let $\mathcal{D}_1$ be the disk system obtained from $\mathcal{D}_0$ by the full disk replacement defined by ρ. One of the two disks obtained by simple surgery along ρ is the disk D_1 and hence $D_1 \in \mathcal{D}_1$. The claim now follows from the induction hypothesis, applied to the disk exchange sequence $\Sigma_1, \dots, \Sigma_m$ of length $m-1$ and the simple disk system $\mathcal{D}_1$ containing Σ_1.

In the second case, the returning arc ρ intersects $\mathcal{D}_0 \setminus D$. Then $\rho \setminus (\mathcal{D}_0 \setminus D)$ contains a component ρ' which is a returning arc with endpoints on a disk $Q \in \mathcal{D}_0 \setminus \{D\}$. A replacement of the disk Q by both disks obtained from Q by simple surgery using the returning arc ρ' reduces the number of intersection components of $\rho \cap (\mathcal{D}_0 \setminus D)$. Moreover, the resulting disk system contains D. In finitely many surgery steps, say $s \geq 1$ steps, we obtain a simple disk system $\mathcal{D}_s$ with the following properties.

1. $\mathcal{D}_s$ contains D and is obtained from $\mathcal{D}_0$ by a full disk replacement sequence.

2. $\rho \cap (\mathcal{D}_s - D) = \emptyset$.

Define $r(i) = 0$ for $i = 0, \dots, s$, where r is the function required in the definition of compatibility. We now can use the procedure from the first case above, applied to $\Sigma_0, \mathcal{D}_s$ and ρ to carry out the induction step.

This completes the proof of the lemma. □

4.3 The graph of rigid racks

The goal of this section is to describe a construction of paths in the handlebody group whose geometry is easy to control. A version of these paths was already used in [HH11a] to establish an upper bound for the distortion of the handlebody group in the mapping class group.

The main objects are given by the following

Definition 4.3.1. A *rack* R in V is given by a reduced disk system $\Sigma(R)$, called the *support system* of the rack R, and a collection of pairwise disjoint essential embedded arcs in $\partial V \setminus \partial\Sigma(R)$ with endpoints on $\partial\Sigma(R)$, called *ropes*, which are pairwise non-homotopic relative to $\partial\Sigma(R)$. At each side of a support disk $D \in \Sigma(R)$, there is at least one rope which ends at the disk and approaches the disk from this side. A rack is called *large* if the set of ropes decomposes $\partial V \setminus \partial\Sigma(R)$ into simply connected regions.

We will consider racks up to an equivalence relation called "rigid isotopy" which is defined as follows.

Definition 4.3.2. i) Let R be a large rack. The union of the support system and the system of ropes of R defines the 1–skeleton of a cell decomposition of the surface ∂V which we call the *cell decomposition induced by* R.

ii) Let R and R' be racks. We say that R and R' are *rigidly isotopic* if there is an isotopy of ∂V which maps the support system of R to the support system of R' and defines an isotopy of the cell decompositions induced by R and R'.

In particular, if T is a simple Dehn twist about the boundary of a support disk of a rack R, then R and $T^n(R)$ are not rigidly isotopic for $n \geq 2$. This observation and the fact that the stabilizer in the mapping class group of a reduced disk system is contained in the handlebody group imply the following

Lemma 4.3.3. *The handlebody group acts on the set of rigid isotopy classes of racks with finite quotient and finite point stabilizers.*

For simplicity of notation, we call a rigid isotopy class of a large rack simply a *rigid rack*. Lemma 4.3.3 allows us to use rigid racks as the vertex set of a $\mathrm{Map}(V)$–graph. More precisely, we make the following

Definition 4.3.4. The *graph of rigid racks* $\mathcal{RR}_K(V)$ is the graph whose vertex set is the set of rigid racks. Two such vertices are joined by an edge if up to isotopy, the 1–skeleta of the cell decompositions induced by the racks intersect in at most K points.

It follows easily from Lemma 4.3.3 that the number K may be chosen in such a way that the graph $\mathcal{RR}_K(V)$ is connected. In Lemma 7.3 of [HH11a] such a number $K > 0$ is constructed explicitly. In the sequel, we will always use this choice of K and suppress the mention of K from our notation. It then follows from Lemma 4.3.3 and the Svarc-Milnor lemma that the graph $\mathcal{RR}(V)$ is quasi-isometric to $\mathrm{Map}(V)$.

Next we construct a family of distinguished paths in the graph of rigid racks. The paths are inspired by splitting sequences of train tracks on surfaces. To this end, we first define a notion of "carrying" for racks.

Definition 4.3.5. 1. A disk system $\mathcal{D}$ is *carried* by a rigid rack R if it is in minimal position with respect to the support system $\Sigma(R)$ of R and if each component of $\partial\mathcal{D} \setminus \partial\Sigma(R)$ is homotopic relative to $\partial\Sigma(R)$ to a rope of R.

2. An embedded essential arc ρ in ∂V with endpoints on $\partial\Sigma(R)$ is *carried* by R if each component of $\rho \setminus \partial\Sigma(R)$ is homotopic relative to $\partial\Sigma(R)$ to a rope of R.

3. A *returning rope* of a rigid rack R is a rope which begins and ends at the same side of some fixed support disk D (i.e. defines a returning arc relative to $\partial\Sigma(R)$).

Let R be a rigid rack with support system $\Sigma(R)$ and let α be a returning rope of R with endpoints on a support disk $D \in \Sigma(R)$. By Lemma 4.2.2, for one of the components γ_1, γ_2 of $\partial D \setminus \alpha$, say the component γ_1, the simple closed curve $\alpha \cup \gamma_1$ is the boundary of an embedded disk $D' \subset H$ with the property that the disk system $(\Sigma \setminus D) \cup D'$ is reduced.

A *split* of the rigid rack R at the returning rope α is any rack R' with support system $\Sigma' = (\Sigma(R) \setminus D) \cup D'$ whose ropes are given as follows.

1. Up to isotopy, each rope ρ' of R' has its endpoints in $(\partial\Sigma(R) \setminus \partial D) \cup \gamma_1 \subset \partial\Sigma(R)$ and is an arc carried by R.

2. For every rope ρ of R there is a rope ρ' of R' such that up to isotopy, ρ is a component of $\rho' \setminus \partial\Sigma(R)$.

The above definition implies in particular that a rope of R which does not have an endpoint on ∂D is also a rope of R'. Moreover, there is a map $\Phi : R' \to R$ which maps a rope of R' to an arc carried by R, and which maps the boundary of a support disk of R' to a simple closed curve γ of the form $\gamma_1 \circ \gamma_2$ where γ_1 either is a rope of R or trivial, and where γ_2 is a subarc of the boundary of a support disk of R (which may be the entire boundary circle). The image of Φ contains every rope of R.

We are now ready to recall the construction of a distinguished class of edge-paths in the graph of rigid racks from [HH11a]. These paths are sufficiently well-behaved to yield some geometric control of the handlebody group.

For a reduced disk system Σ let $\mathcal{RR}(V, \Sigma)$ be the complete subgraph of $\mathcal{RR}(V)$ whose vertices are marked rigid racks with support system Σ.

Definition 4.3.6. Let $\mathcal{D}$ be a simple disk system. A *$\mathcal{D}$-splitting sequence* of racks is an edge-path R_i in the graph of rigid racks with the following properties.

i) There is a disk exchange sequence Σ_i in direction of $\mathcal{D}$ and a sequence of numbers $1 = r_1 < \cdots < r_k$ such that the support system of R_j is Σ_i for all $r_i \leq j \leq r_{i+1} - 1$. The sequence Σ_i is called the *associated disk exchange sequence*.

ii) For $r_i \leq j \leq r_{i+1} - 1$, the sequence R_j is a uniform quasi-geodesic in the graph $\mathcal{RR}(V, \Sigma_i)$.

Here and in the sequel, we say that a path is a *uniform* quasigeodesic if the quasigeodesic constants of the path depend only on the genus of the handlebody. Similarly, we say that a number is *uniformly bounded*, if there is a bound depending only on the genus of V.

We showed in [HH11a] that any two points in the graph of rigid racks can be connected by a splitting sequence. More precisely, the proof of Theorem 7.9 of [HH11a] yields

Theorem 4.3.7. *Let R, R' be two rigid racks. Then there is a disk system $\mathcal{D}$ depending only on the support system of R' with the following property. Let $\Sigma(R) = \Sigma_1, \Sigma_2, \ldots, \Sigma_n$ be a disk exchange sequence in direction of $\mathcal{D}$ such that Σ_n is disjoint from $\mathcal{D}$.*

Then there is a splitting sequence connecting R to R' whose associated disk exchange sequence is (Σ_i). The length of such a sequence is bounded uniformly exponentially in the distance between R and R' in the graph of rigid racks.

In Section 4.2 we saw that $\mathcal{D}$–disk exchange sequences starting in disjoint reduced disk systems can be compared using full disk replacement sequences. In the rest of this section we develop a slight generalization of racks, which will allow to similarly compare $\mathcal{D}$–splitting sequences starting in adjacent vertices of $\mathcal{RR}(V)$.

Namely, define an *extended rack* R in the same way as a rack except that now the support system $\mathcal{D}(R)$ of R may be any simple disk system instead of a reduced disk system. The cell decomposition induced by an extended rack is defined in the obvious way, and similarly we can talk about rigid isotopies between extended racks. The rigid isotopy class of an extended rack is called a *rigid extended rack*.

Rigid extended racks can be used in the same way as racks to define a geometric model for the handlebody group.

Definition 4.3.8. The graph of rigid extended racks $\mathcal{RER}_K(V)$ is the graph whose vertex set is the set of large rigid extended racks. Two such vertices are connected by an edge of length one if up to isotopy the 1–skeleta of the cell decompositions induced by the corresponding vertices intersect in at most K points.

Again, the constant K is chosen in such a way that the graph of rigid extended racks is connected. We denote the resulting graph by $\mathcal{RER}(V)$. For future use, we choose the constant K big enough such that in addition the following holds. For a simple disk system $\mathcal{D}$ let $\mathcal{RER}(V, \mathcal{D})$ be the complete subgraph of $\mathcal{RER}(V)$ whose vertices are rigid extended racks with support system $\mathcal{D}$. We may choose K large enough such that for any simple disk system $\mathcal{D}$ the subgraph $\mathcal{RER}(V, \mathcal{D})$ is connected.

An analog of Lemma 4.3.3 holds for rigid extended racks as well, and implies that the handlebody group acts on the graph of rigid extended racks with finite quotient. Thus, the graph of rigid extended racks is quasi-isometric to the handlebody group. Note also that every large rack is a large extended rack. Thus the graph of rigid extended racks embeds as a subgraph in the graph of rigid extended racks. This inclusion is a quasi-isometry.

A *full split* of a rigid extended rack is defined as follows. Let R be a rigid extended rack and let α be a returning rope of R. A rigid extended rack R' is called a full split of R at α if the support system of R' is obtained from $\Sigma(R)$ by a full disk replacement along α. Moreover, we require that the ropes of R' satisfy the analogous conditions as the ropes of a split of a rigid rack.

The following is a natural generalization of splitting paths to extended racks.

Definition 4.3.9. Let $\mathcal{D}$ be a simple disk system. A *full $\mathcal{D}$-splitting sequence* of racks is an edge-path (R_i) in the graph of rigid extended racks with the following properties.

i) There is a full disk exchange sequence $(\mathcal{D}_i)$ in direction of $\mathcal{D}$ and a sequence of numbers $1 = r_1 < \cdots < r_k$ such that the support system of R_j is $\mathcal{D}_i$ for all $r_i \leq j \leq r_{i+1} - 1$. The sequence $(\mathcal{D}_i)$ is called the *associated full disk exchange sequence.*

ii) For $r_i \leq j \leq r_{i+1} - 1$, the sequence (R_j) is a uniform quasi-geodesic in the graph $\mathcal{RER}(V, \Sigma_i)$.

The proof of Theorem 7.9 of [HH11a] implies the following theorem which allows to connect two rigid racks with a full splitting sequence.

Theorem 4.3.10. *There is a number k_1 with the following property. Let R, R' be two rigid racks. Then there is a simple disk system $\hat{\mathcal{D}}$ depending only on the support system of R' with the following property. Let $\mathcal{D}(R) = \mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_n$ be a full disk exchange sequence in direction of $\hat{\mathcal{D}}$ such that $\mathcal{D}_n$ is disjoint from $\hat{\mathcal{D}}$.*

Then there is an extended rigid rack $\hat{R}$ which is at distance at most k_1 to R' in $\mathcal{RER}(V)$, and there is a full splitting sequence connecting R to $\hat{R}$ whose associated full disk replacement sequence is $(\mathcal{D}_i)$. The length of any such sequence is bounded by $e^{k_1 d}$, where d is the distance between R and R' in the graph of rigid extended racks.

Combining Lemmas 4.2.6 and 4.2.7 with Theorem 4.3.10 above, we obtain the following.

Corollary 4.3.11. *There is a number $k_2 > 0$ with the following property.*

i) Let $(R_i), i = 1, \ldots N$ be a $\mathcal{D}$–splitting sequence of racks with associated disk exchange sequence (Σ_j). Let $(\mathcal{D}_j)$ be a full disk replacement sequence compatible with (Σ_j). Then there is a full $\mathcal{D}$–splitting sequence $\widetilde{R}_k, k = 1, \ldots K$ such that the following holds. The associated full disk replacement sequence to $(\widetilde{R}_k)$ is $\mathcal{D}_j$. Furthermore, $\widetilde{R}_1 = R_1$ and the distance between $\widetilde{R}_K$ and R_N is at most k_1. The length K of any such sequence is at most $e^{k_2 d}$, where d is the distance between R_0 and R_K in the graph of rigid racks.

ii) Conversely, suppose that $\widetilde{R}_k, k = 1, \ldots, K$ is a full $\mathcal{D}$–splitting sequence with associated full disk replacement sequence $\mathcal{D}_j$. Suppose further that (Σ_j) is a disk exchange sequence compatible with $(\mathcal{D}_i)$. If $\widetilde{R}_1$ is a large rack, then there is a $\mathcal{D}$–splitting sequence $R_1, R_2, \ldots, R_N$ whose associated disk exchange sequence is (Σ_j) such that R_N is of distance at most k_1 to $\widetilde{R}_K$. The length N of any such sequence is at most $e^{k_2 d}$, where d is the distance between R_0 and R_N in the graph of rigid racks.

4.4 The Dehn function of $\mathrm{Map}(V_g)$

In this section we prove the main result of this chapter.

Theorem 4.4.1. *The Dehn function of the handlebody group has at most exponential growth rate.*

To begin, we recall the definitions of the Dehn function and growth rate. Let G be a finitely presented group. Choose a finite generating $\mathcal{S}$ and let $\mathcal{R}$ be a finite defining set of relations for G. This means the following. The set $\mathcal{R}$ generates a subgroup R_0 of the free group $F(\mathcal{S})$ with generating set $\mathcal{S}$. Denote by R the normal closure of R_0 in $F(\mathcal{S})$. The set $\mathcal{R}$ is called a defining set of relations for G if the quotient $F(\mathcal{S})/R$ is isomorphic to G.

Every $r \in R < F(\mathcal{S})$ can be written as a product of conjugates of elements in $\mathcal{R}$:

$$r = \Pi_{i=1}^{n} r_i^{\gamma_i}, \qquad r_i \in \mathcal{R}, \gamma_i \in G.$$

We call the minimal length n of such a product *the area* $\mathrm{Area}(r)$ *needed to fill the relation* r. On the other hand, r can be written as a word in the elements of $\mathcal{S}$. We call the minimal length of such a word the *the length* $l(r)$ *of the loop* r.

The *Dehn function of* G is then be defined by

$$\delta(n) = \max\{\mathrm{Area}(r) | r \in R \text{ with } l(r) \leq n\}.$$

The function δ depends on the choice of the generating set $\mathcal{S}$ and the set of relations $\mathcal{R}$. However, the Dehn function obtained from different generating sets and defining relations are equivalent in the following sense. Say that two functions $f, g : \mathbb{N} \to \mathbb{N}$ are *of the same growth type*, if there are numbers $K, L > 0$ such that

$$L^{-1} \cdot g(K^{-1} \cdot x - K) - L \leq f(x) \leq L \cdot g(K \cdot x + K) + L$$

for all $x \in \mathbb{N}$.

In this section we use the graph $\mathcal{RER}(V)$ of rigid extended racks as a geometric model for the handlebody group.

To estimate the Dehn function, we consider a loop γ in $\mathcal{RER}(V)$ of length $R > 0$. We have to show that there is a number $k > 0$ and that there are at most e^{kR} loops $\zeta_1, \ldots, \zeta_m$ of length at most k so that γ can be contracted to a point in m steps consisting each of replacing a subsegment of ζ_i by another subsegment of ζ_i. This suffices, since each loop ζ_i as above corresponds to a cycle in the handlebody group which can be filled with uniformly small area.

Recall from Section 4.3 the definition of the graph $\mathcal{RER}(V, \mathcal{D})$. The following lemma allows to control the isoperimetric function of these subgraphs.

Lemma 4.4.2. *Let $\mathcal{D}$ be a simple disk system for V.*

i) $\mathcal{RER}(V, \mathcal{D})$ is a connected subgraph of $\mathcal{RER}(V)$ which is equivariantly quasi-isometric to the stabilizer of $\partial\mathcal{D}$ in the mapping class group of ∂V.

ii) $\mathcal{RER}(V, \Sigma)$ is quasi-isometrically embedded in $\mathcal{RER}(V)$.

iii) Any loop in $\mathcal{RER}(V, \Sigma)$ can be filled with area coarsely bounded quadratically in its length.

Proof. $\mathcal{RER}(V, \Sigma)$ is connected by definition of the graph of rigid extended racks (see Section 4.3).

Let G be the stabilizer of $\partial\mathcal{D}$ in the mapping class group of ∂V. The group G is contained in the handlebody group since every homeomorphism of the boundary of a spotted ball extends to the interior. The group G acts on $\mathcal{RER}(V, \mathcal{D})$ with finite quotient and finite point stabilizers. To show this, note that up to the action of the mapping class group, there are only finitely many isotopy classes of cell decompositions of a bordered sphere with uniformly few cells. Thus by the Svarc-Milnor lemma, $\mathcal{RER}(V, \mathcal{D})$ is equivariantly quasi-isometric to G, showing *i)*.

The stabilizer G of $\partial\mathcal{D}$ is quasi-isometrically embedded in the full mapping class group of ∂V (see [MM00] or [Ha09b, Theorem 2]). Hence G is also quasi-isometrically embedded in the handlebody group. Together with *i)* this shows *ii)*.

The group G is a Lipschitz retract of the mapping class group of ∂V (see [HM10] for a detailed discussion of this fact which is a direct consequence of the work of Masur and Minsky [MM00]). Mapping class groups are automatic [Mo95] and hence have quadratic Dehn function. Then the same holds true for G (compare again [HM10]). This implies claim *iii)*. $\square$

As the next step, we use Corollary 4.3.11 to control splitting paths starting at adjacent points in the graph of marked racks. We show that these paths can be constructed in such a way that the resulting loop can be filled with controlled area. Together with the length estimate for marked splitting paths from Theorem 4.3.10 this will imply the exponential bound for the Dehn function.

The main technical tool in this approach is given by the following lemma.

Lemma 4.4.3. *For each $k > 0$ there is a number $k_3 > 0$ with the following property.*

Let $\mathcal{D}$ be a simple disk system. Let $R_i, i = 1, \dots, N$ be a $\mathcal{D}$–splitting sequence of rigid racks and let $\widetilde{R}_j, j = 1, \dots, M$ be a full $\mathcal{D}$–splitting sequence of extended racks such that the following holds.

i) The rigid extended racks R_1 and $\widetilde{R}_1$ (respectively R_N and $\widetilde{R}_M$) have distance at most k in the graph of rigid extended racks.

ii) The associated disk exchange sequences of R_i and $\widetilde{R}_j$ are compatible.

Then the loop γ in $\mathcal{RER}(V)$ formed by the sequences (R_i), $(\widetilde{R}_j)$ and geodesics between R_1 and $\widetilde{R}_1$ and R_N and $\widetilde{R}_M$ can be filled with area $k_3(N+M)^3$.

Proof. The idea of the proof is to inductively decompose the loop γ into smaller loops, each of which can be filled with area at most $k_3(N+M)^2$ for a suitable k_3.

Denote the disk exchange sequence associated to R_i by $\Sigma_i, i = 1, \dots, n)$ and the full disk replacement sequence associated to $\widetilde{R}_j$ by $\mathcal{D}_j, j = 1, \dots m$. Let $r : \{1, \dots, m\} \to \{1, \dots, n\}$ be the monotone non-decreasing surjective function given by compatibility, i.e. $\Sigma_{r(j)} \subset \mathcal{D}_j$ for all $j = 1, \dots, m$.

We define

$$I(i) = \{k \mid \Sigma(R_k) = \Sigma_i\}$$

and

$$J(i) = \{k \mid \mathcal{D}(\widetilde{R}_k) = \mathcal{D}_l \text{ and } r(l) = i\}.$$

Put $i_k = \max I(k)$ and $j_k = \max J(k)$. We will inductively choose paths d_k connecting R_{i_k} to $\widetilde{R}_{j_k}$ and paths c_k connecting R_{i_k+1} to $\widetilde{R}_{j_k+1}$ with the following properties.

i) The path c_k is a uniform quasigeodesic in $\mathcal{RER}(V, \Sigma_{k+1})$.

ii) The path d_k is a uniform quasigeodesic in $\mathcal{RER}(V, \Sigma_k)$.

iii) The paths c_{k+1}, d_k are uniform fellow-travelers, i.e. the Hausdorff distance between c_{k+1} and d_k is uniformly bounded.

A family of paths with these properties implies the statement of the lemma in the following way.

The restriction of the sequence R_i to $I(k)$ and the restriction of $\widetilde{R}_j^{-1}$ to $J(k)$ form together with c_{k-1}^{-1} and d_k a loop γ_k in $\mathcal{RER}(V, \Sigma_k)$. The length of c_{k-1} and d_k is coarsely bounded by $N+M$ by the triangle inequality. Hence, the length of γ_k can be coarsely bounded by $4(N+M)$. Since $\mathcal{RER}(V, \Sigma_k)$ admits a quadratic isoperimetric function, this loop can thus be filled with area bounded by $k_3(N+M)^2$ for some uniform constant k_3.

Similarly, the paths d_k^{-1} and c_{k+1} , together with the edges connecting R_{i_k} to R_{i_k+1} and $\widetilde{R}_{j_k}$ to $\widetilde{R}_{j_k+1}$ form a loop δ_k. The length of δ_k can again be coarsely bounded by $2(N+M)$ using the triangle inequality. Since the paths d_k and c_k are fellow-travelers, δ_k can be filled with area depending linearly on its length.

There are at most $2\max(N, M)$ loops γ_k, δ_k. Hence, the concatenation of all the loops γ_i and δ_j can be filled with area at most $k_3(N+M)^3$ (after possibly enlarging the constant k_2). The paths c_i and d_i occur in the concatenation of γ_i and δ_j twice, with opposite orientations, except for c_0 and the last occurring arc d_L. As a consequence, the concatenation of the loops γ_i and δ_j is, after erasing these opposite paths, uniformly close to γ in the Hausdorff metric. Thus, γ may also be filled with area bounded by $k_3(N+M)^2$ (again possibly increasing k_3).

We now describe the inductive construction of the paths c_k and d_k. We set $c_0 = d_0$ to be the constant path R_1. Suppose that the paths c_i, d_i are already constructed for $i = 0, \ldots, k-1$.

The support systems of R_{i_k} and $\widetilde{R}_{j_k}$ both contain Σ_k. We first construct the path d_k connecting R_{i_k} and $\widetilde{R}_{j_k}$.

Namely, the reduced disk systems Σ_k and Σ_{k+1} are disjoint. The simple disk system $\Sigma_k \cup \Sigma_{k+1}$ is disjoint from the support systems of R_{i_k}, R_{i_k+1} and $\widetilde{R}_{j_k}, \widetilde{R}_{j_k+1}$ by definition of a split. Furthermore, the 1-skeleta of the cell decompositions of all four of these extended racks intersect $\partial\Sigma_k \cup \partial\Sigma_{k+1}$ in uniformly few points. Hence, there are rigid extended racks U_1, U_2 which have $\Sigma_k \cup \Sigma_{k+1}$ as their support system and such that U_1 is uniformly close to R_{i_k}, and U_2 is uniformly close to $\widetilde{R}_{j_k}$ in $\mathcal{RER}(V)$. Let e be a geodesic path in $\mathcal{RER}(V, \Sigma_k \cup \Sigma_{k+1})$ connecting U_1 and U_2. Since $\mathcal{RER}(V, \Sigma_k \cup \Sigma_{k+1})$ is undistorted in $\mathcal{RER}(V)$ by Lemma 4.4.2, the length of e is coarsely bounded by $N+M+1$. By adding uniformly short geodesic segments in $\mathcal{RER}(V, \Sigma_k)$ at the beginning and the end of e, we obtain the path d_k with property *ii)*.

By definition of i_k and j_k, we have $i_k+1 \in I(k+1)$ and $j_k+1 \in J(k+1)$. Hence, both R_{i_k+1} and $\widetilde{R}_{j_k+1}$ contain Σ_{k+1} in their support systems. We can thus define c_k with properties *i)* and *iii)* by adding uniformly short geodesic segments in $\mathcal{RER}(V, \Sigma_{k+1})$ to the beginning and the end of e. □

We have now collected all the tools for the proof of the main theorem.

Proof of Theorem 4.4.1. Recall that it suffices to show that every loop in the graph of rigid racks can be filled with area coarsely bounded by an exponential function of its length.

Let R_i be a loop of length L in the graph of rigid racks based at $R_0 = \hat{R}$. Let $\hat{\Sigma}$ be the disk system given by Theorem 4.3.10 applied to $R' = R_0$. Since the graph of rigid racks is quasi-isometric to the graph of extended rigid racks, we can consider R_i as a loop in $\mathcal{RER}(V)$ and it suffices to show that this loop can be filled in $\mathcal{RER}(V)$ with area bounded exponentially in its length.

The strategy of this proof is similar to the proof of Lemma 4.4.3: we will write the loop (R_i) as a concatenation of smaller loops whose area we can control.

We will define paths c_i in $\mathcal{RER}(V)$ with the following properties.

1. The path c_i connects R_i to a rack which is uniformly close to $\hat{R}$ in $\mathcal{RER}(V)$.

2. The path c_i is a $\hat{\Sigma}$–splitting sequence of racks.

3. The loop formed by c_i, c_{i+1}, the edge between R_i and R_{i+1}, and a geodesic connecting other pair of endpoints of c_i, c_{i+1} can be filled with area bounded by $e^{k_3 L}$.

As a consequence, the loop (R_i) itself can be filled with area at most $Le^{k_3 L}$, proving the theorem.

The construction of the paths c_i is again by induction. We set c_0 to be the constant path R_0. Suppose now that the path c_k is already constructed. Since R_k and R_{k+1} are connected by an edge in the graph of rigid racks, their support systems Σ_k and Σ_{k+1} are disjoint. Let $\Sigma_k^{(i)}, i = 1, \dots, n$ be the disk exchange sequence associated to the splitting sequence c_k. Put $\mathcal{D}_1 = \Sigma_k \cup \Sigma_{k+1}$. Using Lemma 4.2.7 we obtain a full disk replacement sequence $(\mathcal{D}_i)$ compatible with $(\Sigma_k^{(i)})$. Corollary 4.3.11 part *i)* then yields a full splitting sequence $\widetilde{R}_k, k = 1, \dots, M$ with associated full disk exchange sequence $(\mathcal{D}_i)$. By Lemma 4.4.2, the loop formed by c_k and $\widetilde{R}_k, k = 1, \dots, M$ can be filled with area bounded by $k_2(N+M)^3$, where N is the length of the path c_k.

Using Lemma 4.2.6 on the sequence $(\mathcal{D}_i)$ and the initial reduced disk system (Σ_{k+1}) we obtain a $\hat{\Sigma}$–splitting sequence $(\Sigma^{(i)}_{k+1})$ compatible with $(\mathcal{D}_i)$, which starts in Σ_{k+1}. Corollary 4.3.11 part *ii)* now yields a $\hat{\Sigma}$–splitting sequence c_{k+1} starting in R_{k+1} and ending uniformly close to $\hat{R}$. Applying Lemma 4.4.3 again, we see that the loop formed by c_{k+1} and $\widetilde{R}_k, k = 1, \dots, M$ can be filled with area bounded by $k_3(N' + M)^3$, where N' is the length of the path c_{k+1}.

Since both c_k and c_{k+1} are splitting sequences connecting points which are of distance at most L, their lengths can be bounded by $k_4 e^{k_4 L}$ for a suitable k_4 by Theorem 4.3.7. As a consequence, the paths c_k and c_{k+1} satisfy condition *iii)*. This concludes the inductive construction of c_i and the proof of the theorem. □

The proof of the theorem would give a polynomial bound for the Dehn function provided that the length of the splitting paths used to fill in loops had a length which is polynomial in the distance between their endpoints. Unfortunately, however, the following example show that such a bound does not exist. This is similar to the behavior of paths of sphere systems used in [HV96] to show an exponential upper bound for the Dehn function of $\mathrm{Out}(F_n)$,

For simplicity of exposition, we do not construct these paths in the graph of rigid racks (or the handlebody group), but instead in a slightly simpler graph. The example given below can be extended to the full graph of rigid racks in a straightforward fashion.

We define $\mathcal{RD}(V)$ to be the graph of reduced disk systems in V. The vertex set of $\mathcal{RD}(V)$ is the set of isotopy classes of reduced disk systems, and two such vertices are connected by an edge of length one if the corresponding disk systems are disjoint. Every directed disk exchange sequence defines an edge-path in $\mathcal{RD}(V)$. The following example shows that the length of these edge-paths may be exponential in the distance between their endpoints.

Example 4.4.4. Consider a handlebody V of genus 4. For each $n \in \mathbb{N}$ we will construct a disk exchange sequence $\Sigma^{(n)}_1, \dots, \Sigma^{(n)}_{N(n)}$ such that on the one hand, the length $N(n)$ of the sequence growth exponentially in n. On the other hand, the distance between endpoints $\Sigma_1^{(n)}$ and $\Sigma_{N(n)}^{(n)}$ in $\mathcal{RD}(V)$ grows linearly in n. To simplify the notation, in this example we will only construct the endpoint $\Sigma_{N(n)}^{(n)}$ and denote it by Σ_n.

We choose three disjoint simple closed curves $\alpha_1, \alpha_2, \alpha_3$ which decompose the surface ∂V into a pair of pants, two once-punctured tori and a once-punctured genus 2 surface (see Figure 4.1). We may choose the α_i such that they bound disks in V. We denote the two solid tori in the complement of these disks by T_1, T_2 and the genus 2 subhandlebody by V'.

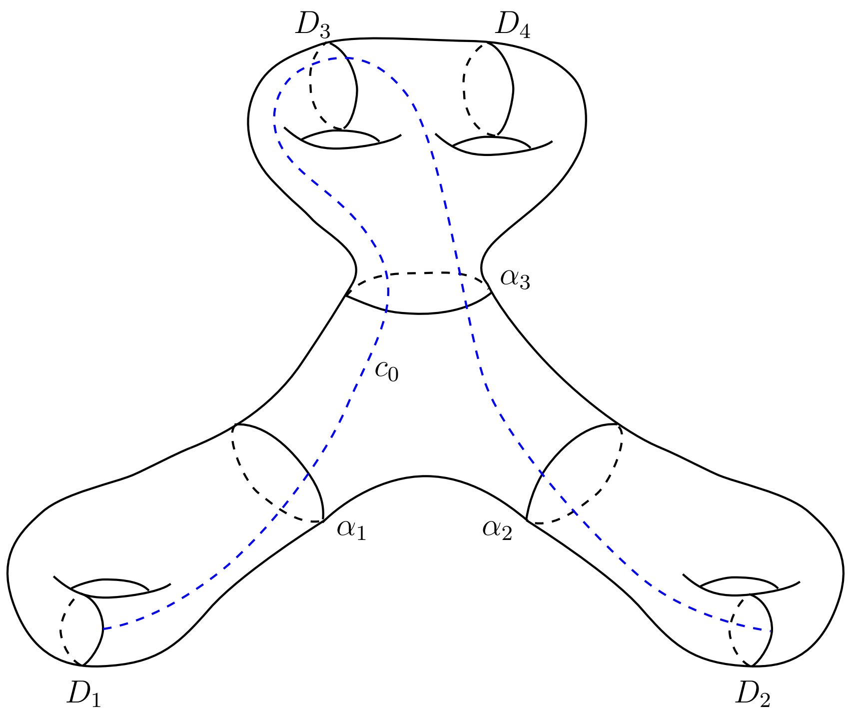


Figure 4.1: The setup for the example of a non-optimal disk exchange path. An admissible arc is drawn dashed.

Let $\Sigma_0 = \{D_1, D_2, D_3, D_4\}$ be a reduced disk system such that $D_1 \subset T_1, D_2 \subset T_2$ and $D_3, D_4 \subset V'$. Choose a base point p on α_3. Let γ_1, γ_2 be two disjointly embedded loops on $\partial V \cap V'$ based at p with the following property. The loop γ_1 intersects the disk D_3 in a single point and is disjoint from D_4, while γ_2 intersects D_4 in a single point and is disjoint from D_3. Since the complement of $D_3 \cup D_4$ in V' is simply connected, such a pair of loops generates the fundamental group of V'. Denote the projections of γ_1

and γ_2 to $\pi_1(V', p)$ by A_1 and A_2, respectively.

Let c be an embedded arc on ∂V. We say that c is *admissible* if the following holds. The arc c connects the disk D_1 to the disk D_2. The interior of c intersects α_1 and α_2 in a single point each. Furthermore it intersects α_3 in two points, and its interior is disjoint from both D_1 and D_2.

Let c be an admissible arc. The intersection of c with V' is an embedded arc c' connecting α_3 to itself. The arc c' may be turned into an embedded arc in V' based at p by connecting the two endpoints of c' to p along α_3. Since the curve α_3 bounds a disk in V', the image of this loop in $\pi_1(V', p)$ is determined by the homotopy class of the arc c relative to $\partial D_1, \partial D_2$. We call this image the *element induced by the arc c*.

Choose an admissible arc c_0 in such a way that it intersects the disk D_3 in a single point, and is disjoint from D_4 (see Figure 4.1 for an example). Up to changing the orientation of γ_1 we may assume that the element induced by c_0 is A_1.

We now describe a procedure that produces essential disks from admissible arcs. To this end, let c be an admissible arc. Consider a regular neighborhood U of $D_1 \cup c \cup D_2$. Its boundary consists of three simple closed curves. Two of them are homotopic to either ∂D_1 or ∂D_2. The third one we denote by $\beta(c)$. Note that $\beta(c)$ bounds a nonseparating disk in V.

Choose a fixed element φ of the handlebody group of V with the following properties. The mapping class φ fixes the isotopy classes of the curves α_1, α_2 and α_3. The restriction of φ to the complement of V' is isotopic to the identity. The restriction of φ to V' induces an automorphism of exponential growth type on $\pi_1(V')$. To be somewhat more precise, we may choose φ such that it acts on the basis A_i as the following automorphism Φ:

$$\begin{aligned} A_1 &\mapsto A_1 A_2 \\ A_2 &\mapsto A_1^2 A_2 \end{aligned}$$

Put $c_n = \varphi^n(c_0)$ and $\beta_n = \beta(c_n)$. We claim that a disk exchange sequence in direction of β_n that makes β_n disjoint from Σ_0 has length at least 2^n.

To this end, note that the arc c_n intersects the disks D_3 and D_4 in at least 2^n points. Namely, the element of $\pi_1(V', p)$ induced by $\varphi^n(c_0)$ is equal to $\Phi^n(A_1)$. The cyclically reduced word describing $\Phi^n(A_1)$ in the basis A_1, A_2 has length at least 2^n by construction of Φ.

Hence, the curve β_n can be described as follows. Choose a parametrization $\beta_n : [0, 1] \to \partial V$. Then there are numbers $0 < t_1 < \cdots < t_N < t_{N+1} <$

$\cdots < t_{2N} < 1$ such that the following holds. Each subarc $\beta_n([t_i, t_{i+1}])$ intersects Σ_0 only at its endpoints. The subarcs $\beta_n([t_N, t_{N+1}])$ and $\beta_n([0, t_1] \cup [t_{2N}, 1])$ are returning arcs to Σ_0. Furthermore, the arcs $\beta_n([t_i, t_{i+1}])$ and $\beta_n([t_{2N-i}, t_{2N+1-i}])$ are homotopic relative to Σ_0 for all $i = 1, \ldots, N-1$. More generally, if there are numbers t_i with these properties for a reduced disk system Σ we say that β_n *is a long string of rectangles with respect to* Σ. The number N is then called *the length of the string of rectangles.* By construction, the length N of the string of rectangles β_n defines with respect to Σ_0 is at least 2^n.

The curve β_n has two returning arcs with respect to Σ_0. Choose one of them, say $\beta_n([t_N, t_{N+1}])$, and denote it by a. Let $\sigma \in \Sigma_0$ denote the disk containing the endpoints of a. One of the disks obtained by simple surgery along a is isotopic to either D_1 or D_2 (depending on which returning arc we chose). The preferred interval defined by a contains every intersection point of β_n with σ except the endpoints of a.

Denote by Σ_1 the reduced disk system obtained by simple surgery along a. By construction, the subarc $\beta_n(t_{N-1}, t_{N+2})$ now defines a returning arc with respect to Σ_1. One of the disks obtained by simple surgery along this returning arc is still properly isotopic to D_1. Furthermore, the subarcs $\beta_n([t_i, t_{i+1}])$ and $\beta_n([t_{2N-i}, t_{2N+1-i}])$ are still arcs with endpoints on Σ_1 which are homotopic relative to Σ_1 for all $i = 1, \ldots, N-2$. Each of these arcs cannot be homotoped into $\partial\Sigma_1$.

Hence the curve β_n has a description as a string of rectangles of length $N-1$ with respect to Σ_1 and the argument can be iterated. By induction, it follows that any disk exchange sequence starting in Σ_0 which ends in a disk system disjoint from β_n has length at least 2^n.

On the other hand, the growth of the distance between Σ_0 and $\varphi^n(\Sigma_0)$ in the graph of reduced disk systems is linear in n by the triangle inequality. The curve β_n intersects $\varphi^n(\Sigma_0)$ in uniformly few points, and thus the disk system $\varphi^n(\Sigma_0)$ is uniformly close to a reduced disk system that is disjoint from β_n. Thus the disk systems Σ_n have the properties described in the beginning of the example.

Chapter 5

$\mathrm{Out}(F_n)$ as a mapping class group[1]

5.1 Introduction

The mapping class group $\mathrm{Map}(S_g)$ of a closed surface S_g of genus g is defined in topological terms: it is the quotient of the group of homeomorphisms of S_g by the connected component of the identity. The classical Dehn-Nielsen-Baer theorem identifies $\mathrm{Map}(S_g)$ with a purely algebraic object, namely the outer automorphism group $\mathrm{Out}(\pi_1(S_g, p))$ of the fundamental group of the surface S_g.

The mapping class group is finitely presented and hence it admits a family of left invariant metrics which are unique up to quasi-isometry. Such a metric can be investigated using simple topological objects as the main tool. In [MM00] the authors construct explicit families of quasi-geodesics in $\mathrm{Map}(S_g)$ using the combinatorics of isotopy classes of simple closed curves on S_g. This approach leads to a geometric understanding of the mapping class group and of many of its natural subgroups.

The outer automorphism group $\mathrm{Out}(F_n)$ of the free group with $n \geq 2$ generators is a finitely presented group which also has a topological description. To this end, let M_n be the connected sum of n copies of $S^1 \times S^2$. By a

[1] This chapter is identical with the preprint [HH11b]
Ursula Hamenstädt and Sebastian Hensel,
Sphere systems, intersections and the geometry of $\mathrm{Out}(F_n)$,
arXiv:1109.2687

theorem of Laudenbach [L74], Out(F_n) is a cofinite quotient of the group of all isotopy classes of orientation preserving homeomorphism of M_n.

As in the case of surface mapping class groups, the geometry of Out(F_n) can be investigated using as a tool the simplest essential submanifolds of M_n, namely embedded spheres. This idea was used by Hatcher in [Ha95] to show homological stability for Out(F_n). A geometric application of this approach includes an upper bound for the growth rate of the Dehn function of Out(F_n) (see [HV96]).

The main goal of this chapter is to initiate an investigation of the large-scale geometry of Out(F_n) from this topological point of view. Explicitly, we analyze the extrinsic geometry of two families of subgroups of Out(F_n) which can be described as follows.

The fundamental group of a surface $S_{g,1}$ of genus $g \geq 1$ with one puncture is the free group F_{2g}. A version of the Dehn-Nielsen-Baer theorem for the mapping class group Mod($S_{g,1}$) of $S_{g,1}$ states that there is a group isomorphism ι of Mod($S_{g,1}$) onto the subgroup of Out(F_n) of all outer automorphisms which preserve the conjugacy class defined by a puncture parallel simple closed curve in $S_{g,1}$. We show

Theorem 5.3.2. *The homomorphism ι is a quasi-isometric embedding.*

In fact, for any number $m \geq 1$, the mapping class group of a surface S of genus $g \geq 0$ with $m \geq 0$ punctures and fundamental group F_n embeds onto a subgroup of Out(F_n). However, we do not investigate such subgroups in the case $m \geq 2$ here.

There is an analog of Theorem 5.3.2 for graphs which admit cofinite actions of Mod($S_{g,1}$) and Out(F_{2g}), respectively. Namely, let $\mathcal{AG}(S_{g,1})$ be the arc graph of $S_{g,1}$. The vertex set of $\mathcal{AG}(S_{g,1})$ is the set of isotopy classes of essential embedded arcs connecting the puncture of $S_{g,1}$ to itself. Two such vertices are connected by an edge if the corresponding arcs are disjoint up to homotopy. The mapping class group Map($S_{g,1}$) of a once-punctured surface acts on $\mathcal{AG}(S_{g,1})$. We define the sphere graph $\mathcal{SG}(M_{2g})$ of M_{2g} as the graph whose vertex set is the set of isotopy classes of embedded essential spheres in M_{2g}. Two such vertices are connected by an edge if the corresponding spheres are disjoint up to homotopy. The tools developed for the proof of Theorem 5.3.2 also yield

Proposition 5.3.9. *There is a* Map($S_{g,1}$)*–equivariant quasi-isometric embedding of the arc graph $\mathcal{AG}(S_{g,1})$ into the sphere graph $\mathcal{SG}(M_{2g})$.*

The main idea for the proof of Theorem 5.3.2 is as follows. The mapping class group of the 3-manifold M_{2g} acts properly and cocompactly on the graph $\mathcal{S}_0(M_{2g})$ whose vertices are *reduced sphere systems*, i.e. systems of $2g$ pairwise non-isotopic essential spheres which cut M_{2g} into a single simply connected region. Consider an embedding $\varphi : S_g^1 \to M$ of a surface S_g^1 of genus g with one boundary component into M_{2g} which induces an isomorphism on the level of fundamental groups. The intersection of a simple sphere system with the image of φ (where both surfaces are supposed to be in general position) defines an embedded system of arcs on S_g^1 which decomposes S_g^1 into simply connected regions. We can use the set of isotopy classes of such arc systems as a vertex set for a $\mathrm{Map}(S_{g,1})$–complex on which $\mathrm{Map}(S_{g,1})$ acts properly and cocompactly. The main task is now to show that edge paths in $\mathcal{S}_0(M_{2g})$ can be arranged to trace out edge paths of the same length in this complex. We also have to establish a topological characterization of those edge-paths in $\mathcal{S}_0(M_{2g})$ which connect two points in a fixed orbit of $\iota(\mathrm{Mod}(S_{2g}))$.

Investigating $\mathrm{Out}(F_n)$ via sphere systems and intersections can also be used to give a short proof of a recent result of Handel and Mosher [HM10]. We define a procedure which makes a simple sphere system disjoint from a given essential 2–sphere σ in M_n. This procedure allows us to show that stabilizers of homotopy classes of essential spheres in the mapping class group of M_n are undistorted. Recall that a finitely generated subgroup H of a finitely generated group G is said to be undistorted if the inclusion map of H into G is a quasi-isometric embedding.

Namely, the stabilizer of the homotopy class of sphere σ in the mapping class group of M_n is equivariantly quasi-isometric to the complete subgraph $\mathcal{S}(M_n, \sigma)$ of $\mathcal{S}(M_n)$ whose vertices correspond to simple sphere systems containing σ. Let Σ, Σ' be two simple sphere systems containing σ and let $\Sigma_1, \ldots, \Sigma_N$ is a shortest path in $\mathcal{S}(M_n)$ connecting Σ to Σ'. Applying the intersection procedure to each Σ_i we obtain a path of length N in $\mathcal{S}(M_n, \sigma)$ connecting Σ to Σ'. Thus, the subgraph $\mathcal{S}(M_n, \sigma)$ is undistorted in $\mathcal{S}(M_n)$ and therefore the stabilizer of σ is undistorted in the mapping class group of M_n. By rephrasing this result in group theoretic terms, we obtain the following result of [HM10].

Theorem 5.2.1. *i) The stabilizer of the conjugacy class of a free splitting $F_n = G * H$ is undistorted in* $\mathrm{Out}(F_n)$.

ii) Let $G < F_n$ be a free factor of corank 1*. Then the stabilizer of the conjugacy class of G is undistorted in* $\mathrm{Out}(F_n)$.

This chapter is organized as follows. In Section 5.2, we first give some background on the manifold M_n and sphere systems. This section also contains the proof of Theorem 5.2.1. Section 5.3 is devoted to the proof of Theorem 5.3.2 and Proposition 5.3.9. Appendix 5.A contains a topological lemma about stabilizers of spheres in M_n which is used in Section 5.2.

5.2 Stabilizers of spheres

Let F_n be the free group on n generators. By $\mathrm{Out}(F_n)$ we denote the outer automorphism group of F_n. Explicitly, $\mathrm{Out}(F_n)$ is the quotient of the group $\mathrm{Aut}(F_n)$ of all automorphisms of F_n by the subgroup of inner automorphisms.

The purpose of this section is to give a short topological proof of a theorem of Handel and Mosher [HM10]. For its formulation, we use the following definitions. A *free splitting* of the free group F_n consists of two subgroups $G, H < F_n$ such that $F_n = G * H$. By this we mean the following: the inclusions of G and H into F_n induce a natural homomorphism $G * H \to F_n$, where $*$ denotes the free product of groups. By stating that $F_n = G * H$ we require that this homomorphism is an isomorphism.

We say that an automorphism φ of F_n preserves the free splitting $F_n = G * H$, if φ preserves the groups G and H. It is possible to define free splittings in a more general way using actions of F_n on trees (see [HM10, Section 1.4]) but in this chapter we use the definition given above.

A *corank* 1 *free factor* is a subgroup G of F_n of rank $n-1$ such that there exists a cyclic subgroup H of F_n with $F_n = G * H$. We say that an automorphism φ of F_n preserves this corank 1 free factor, if φ preserves the group G. We emphasize that φ is not required to preserve the cyclic group H, and that the group H is not uniquely determined by G.

An element $[\varphi] \in \mathrm{Out}(F_n)$ is said to preserve the conjugacy class of the free splitting $G * H$ (or corank 1 free factor G), if there is a representative φ of $[\varphi]$ which preserves the free splitting $G * H$ (or the corank 1 free factor G).

A finite, symmetric generating set of a group G defines a word norm on G. We call the metric induced by such a norm a *word metric* on G. Two different finite generating sets of G give rise to quasi-isometric metrics. Recall that a map $f : X \to Y$ between metric spaces is called a *quasi-isometric embedding*, if there is a number $K > 0$ such that

$$\frac{1}{K}d_Y(f(x), f(x')) - K \leq d_X(x, x') \leq K d_Y(f(x), f(x')) + K$$

for all $x, x' \in X$. A finitely generated subgroup $H < G$ of a finitely generated group G is called *undistorted* if the inclusion homomorphism $H \to G$ is a quasi-isometric embedding.

We can now state the main theorem of this section.

Theorem 5.2.1. *i) The stabilizer of the conjugacy class of a free splitting $F_n = G * H$ is undistorted in* $\mathrm{Out}(F_n)$.

ii) Let $G < F_n$ be a free factor of corank 1. *Then the stabilizer of the conjugacy class of G is undistorted in* $\mathrm{Out}(F_n)$.

As indicated in the introduction, we will prove Theorem 5.2.1 using the topology of the connected sum M_n of n copies of $S^2 \times S^1$ (where S^k denotes the k–sphere). Alternatively, M_n can be obtained by doubling a handlebody of genus n along its boundary. Since $\pi_1(M_n) = F_n$, there is a natural homomorphism from the group $\mathrm{Diff}^+(M_n)$ of orientation preserving diffeomorphisms of M_n to $\mathrm{Out}(F_n)$. This homomorphism factors through the mapping class group $\mathrm{Map}(M_n) = \mathrm{Diff}^+(M_n)/\mathrm{Diff}_0(M_n)$ of M_n, where $\mathrm{Diff}_0(M_n)$ is the connected component of the identity in $\mathrm{Diff}^+(M_n)$. In fact, Laudenbach [L74, Théorème 4.3, Remarque 1)] showed that the following stronger statement is true.

Theorem 5.2.2. *There is a short exact sequence*

$$1 \to K \to \mathrm{Diffeo}^+(M_n)/\mathrm{Diffeo}_0(M_n) \to \mathrm{Out}(F_n) \to 1$$

where K is a finite group, and the map $\mathrm{Diffeo}^+(M_n)/\mathrm{Diffeo}_0(M_n) \to \mathrm{Out}(F_n)$ *is induced by the action on the fundamental group.*

By [L74, Théorème 4.3, part 2)], we can replace diffeomorphisms by homeomorphisms in the definition of the mapping class group of M_n.

An embedded 2-sphere in M_n is called *essential*, if it does not bound a 3-ball in M_n. Throughout the chapter we assume that 2-spheres are smoothly embedded, essential and are two-sided in M_n.

The following observation identifies the stabilizers in $\mathrm{Map}(M_n)$ which occur in Theorem 5.2.1. The statement is an immediate consequence of

Corollary 21 of [HM10] and a standard topological argument which is for example presented in [AS09]. For completeness of exposition we provide a purely topological proof in Appendix 5.A.

Lemma 5.A.1. *i) Let σ be an essential separating sphere in M_n. Then the stabilizer of σ in* Map(M_n) *projects onto the stabilizer of the conjugacy class of a free splitting in* Out(F_n). *Furthermore, every stabilizer of a conjugacy class of a free splitting arises in this way.*

ii) Let σ be a nonseparating sphere in M_n. Then the stabilizer of σ in Map(M_n) *projects onto the stabilizer of the conjugacy class of a corank* 1 *free factor in* Out(F_n). *Furthermore, every stabilizer of a conjugacy class of a corank* 1 *free factor arises in this way.*

To study stabilizers of essential spheres in M_n we use the following geometric model for the mapping class group of M_n (compare [Ha95] and [HV96]).

A *sphere system* is a set $\{\sigma_1, \ldots, \sigma_m\}$ of essential spheres in M_n no two of which are homotopic. A sphere system is called *simple* if its complementary components in M_n are simply connected. The *sphere system graph* $\mathcal{S}(M_n)$ is the graph whose vertex set is the set of homotopy classes of simple sphere systems. Two such vertices are joined by an edge of length 1 if the corresponding sphere systems are disjoint up to homotopy.

The mapping class group of M_n acts on $\mathcal{S}(M_n)$ properly discontinuously and cocompactly (see e.g. the proof of Corollary 4.4 of [HV96] for details on this). Furthermore, the surgery procedure described in Section 3 of [HV96] shows that $\mathcal{S}(M_n)$ is connected. The finite subgroup K occurring in the statement of Theorem 5.2.2 of Map(M_n) acts trivially on isotopy classes of spheres and hence this action factors through an action of Out(F_n).

For an essential sphere σ, let $\mathcal{S}(M_n, \sigma)$ be the complete subgraph of $\mathcal{S}(M_n)$ whose vertex set is the set of homotopy classes of simple sphere systems containing σ. The surgery procedure described in [HV96] shows that the graph $\mathcal{S}(M_n, \sigma)$ is connected. The stabilizer of σ in Out(F_n) acts cocompactly on $\mathcal{S}(M_n, \sigma)$. Thus the Svarc-Milnor lemma immediately implies the following.

Lemma 5.2.3. *i) The sphere system graph $\mathcal{S}(M_n)$ is quasi-isometric to* Out(F_n).

ii) The graph $\mathcal{S}(M_n, \sigma)$ is equivariantly quasi-isometric to the stabilizer of σ in Out(F_n).

Combining Lemma 5.A.1 and Lemma 5.2.3, Theorem 5.2.1 thus reduces to the following.

Theorem 5.2.4. *The inclusion of* $\mathcal{S}(M_n, \sigma)$ *into* $\mathcal{S}(M_n)$ *is a quasi-isometric embedding.*

The main tool used in the proof of this statement is a surgery procedure that makes a given simple sphere system disjoint from the sphere σ. On the one hand, this surgery procedure is inspired by the construction used in [HV96] to show that the sphere system complex is contractible. On the other hand, it is motivated by the subsurface projection methods of [MM00].

To describe this surgery procedure we fix an essential sphere σ in M_n for the rest of this section. We treat separating and nonseparating spheres in a unified manner using the following notation. If σ is separating, let M^1 and M^2 be its complementary components in M_n and put $N_i = M^i \cup \sigma$. We then let N be the disjoint union of N_1 and N_2. If σ is nonseparating, let M be its complement. There is a canonical way to add two copies of σ to M to obtain a compact three-manifold N whose boundary consists of two spheres.

In both cases, N is a compact three-manifold whose boundary consists of two copies of σ. If no confusion can occur we will often treat N as if it were a submanifold of M_n and call it *the complement of* σ. In particular, we simply speak of the intersection of a sphere system with N.

Consider now a simple sphere system Σ of M_n. By applying a homotopy, we may assume that all intersections between Σ and σ are transverse. The intersection of the spheres in Σ with N is a disjoint union of properly embedded surfaces $C_1, \dots, C_m$, possibly with boundary. Each C_i is a subsurface of a sphere in Σ, and thus it is a bordered sphere. If Σ contains spheres disjoint from σ then some of the C_i may be spheres without boundary components. We call the C_i the *sphere pieces* defined by Σ.

We say that Σ and σ *intersect minimally* if the number of connected components of $\Sigma \cap \sigma$ is minimal among all sphere systems homotopic to Σ which intersect σ transversely.

Every simple sphere system Σ can be changed by a homotopy to intersect σ minimally. Unless stated otherwise, we will assume from now on that spheres and sphere systems intersect minimally. Let $\Sigma' \supset \Sigma$ be a simple sphere system and suppose that Σ intersects σ minimally. Then Σ' can be homotoped relative to Σ to intersect σ minimally.

Details on the construction of such a homotopy can be found in [Ha95]. Hatcher also shows the existence of a unique normal form of spheres with respect to simple sphere systems which gives more information than minimal intersection. Since we do not use this normal form here, we refer the reader to [Ha95] for details.

Let C be one of the sphere pieces of Σ, and let $\alpha_1, \dots, \alpha_k$ be its boundary components on ∂N. A *gluing datum for* C is a set of disks $D_1, \dots, D_k$ contained in ∂N such that $\partial D_i = \alpha_i$ for all $1 \leq i \leq k$. The disks D_i are called *closing disks*. Let C' be the surface obtained from C by gluing D_i along ∂D_i to α_i. Since C is a bordered sphere, the surface C' is an immersed sphere in N (which may be inessential). We say that C' is obtained from C by *capping off the boundaries according to the gluing datum*. By a *gluing datum* $\mathcal{D}$ *for* $\{C_i, 1 \leq i \leq m\}$ (or *gluing datum for* Σ) we mean a set of disks on σ consisting of a gluing datum for each sphere piece C_i of Σ.

We say that a gluing datum $\mathcal{D}$ for Σ is *admissible* if it satisfies the following compatibility property: if $D, D' \in \mathcal{D}$ are two disks which intersect nontrivially then $D \subset D'$ or $D' \subset D$. We say that any set of disks with this property is *properly nested*.

Note that if Σ is disjoint from σ then the empty set is the only admissible gluing datum for Σ.

Lemma 5.2.5. *If $\mathcal{D}$ is an admissible gluing datum for Σ then every sphere obtained by capping off the boundary components of a sphere piece according to $\mathcal{D}$ is embedded up to homotopy. Furthermore, the spheres obtained by capping off the boundary components of all sphere pieces according to $\mathcal{D}$ can be embedded disjointly.*

Proof. If $\mathcal{D}$ is empty, there is nothing to show.

Otherwise, say that a disk $D \in \mathcal{D}$ is *innermost* if $D \subset D'$ for every $D' \in \mathcal{D}$ with $D \cap D' \neq \emptyset$. Since $\mathcal{D}$ is admissible, there is at least one innermost disk D_1 bounded by a curve α_1.

In N, the curve α_1 occurs twice as the boundary of a sphere piece, once on each boundary component of N. Let C^1 and C^2 be the two sphere pieces having a copy of α_1 contained in their boundary.

The disk D_1 also occurs on both boundary components of N, and both of these disks have the property that they only intersect a single sphere piece in N, namely one of the C^j. Let D^j be the copy of D_1 intersecting C^j.

We glue D^j to the corresponding sphere piece C^j and then slightly push D^j inside N with a homotopy to obtain a properly embedded bordered sphere C'^j in N. Since D is innermost, this sphere is disjoint from all sphere pieces

$C_k \neq C^j$, and has one less boundary component than C^j. We replace the sphere piece C^j by the bordered sphere C'^j for $j = 1, 2$.

The collection of bordered spheres obtained in this way is a collection of disjointly embedded sphere pieces, and $\mathcal{D} \setminus \{D_1\}$ is an admissible gluing datum for this collection. The lemma now follows by induction on the number of elements in $\mathcal{D}$. $\square$

For an admissible gluing datum $\mathcal{D}$ for Σ, let $\mathcal{S}(\mathcal{D})$ be the collection of disjointly embedded spheres obtained by capping off the boundaries of each sphere piece according to $\mathcal{D}$. The set $\mathcal{S}(\mathcal{D})$ may contain inessential spheres and parallel spheres in the same homotopy class. We denote by $\pi_\sigma(\Sigma, \mathcal{D})$ the union of σ with one representative for each essential homotopy class of spheres occurring in $\mathcal{S}(\mathcal{D})$.

To show that the sphere system obtained in this way from a simple sphere system Σ is again simple, we require the following topological lemma.

Lemma 5.2.6. *Let C be a sphere piece in N intersecting the boundary of N in at least one curve α. Let $D \subset \partial N$ be a disk with $\partial D = \alpha$. Let C' be the sphere piece obtained by gluing D to C and slightly pushing D into N (which might be a sphere without boundary components).*

Then every closed curve in N which can be homotoped to be disjoint from C' can also be homotoped to be disjoint from C.

Proof. Pushing the disk D slightly inside of N with a homotopy traces out a three-dimensional cylinder Q in N. The boundary of Q consists of two disks (the disk D, and the image of D under the homotopy) and an annulus A which can be chosen to lie in C (see Figure 5.1 for an example).

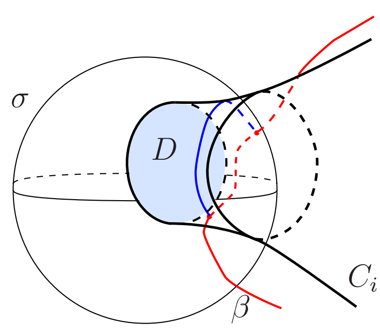


Figure 5.1: Reducing the number of boundary components of a sphere piece.

Suppose that β is a closed curve in N which is disjoint from C' but not from C. Then any intersection point between β and C is contained in the annulus A. Up to homotopy, the intersection between β and Q is a disjoint union of arcs connecting A to itself. Since Q is simply connected, each of these arcs can be moved by a homotopy relative to its endpoints to be contained entirely in A. Slightly pushing each of these arcs off A then yields the desired homotopy that makes β disjoint from C. □

Lemma 5.2.7. *Let Σ be a simple sphere system, and let $\mathcal{D}$ be an admissible gluing datum for Σ. Then $\pi_\sigma(\Sigma, \mathcal{D})$ is a simple sphere system.*

Proof. Let Σ be a simple sphere system, and let $\mathcal{D}$ be an admissible gluing datum. As $\pi_\sigma(\Sigma, \mathcal{D})$ contains σ by construction, it suffices to show that the spheres $S \in \pi_\sigma(\Sigma, \mathcal{D})$ which are distinct from σ decompose N into simply connected regions.

Since the fundamental group of N injects into the fundamental group of M and Σ is a simple sphere system, no essential simple closed curve in N is disjoint from $\Sigma \cap N$. In other words, no essential simple closed curve in N is disjoint from all sphere pieces defined by Σ.

By Lemma 5.2.6, this property is preserved under capping off one boundary component on a sphere piece. By induction, no essential simple closed curve in N is disjoint from all spheres $S \in \mathcal{S}(\mathcal{D})$. Removing inessential spheres and parallel copies of the same sphere from $\mathcal{S}(\mathcal{D})$ does not affect this property.

This implies that $\pi_\sigma(\Sigma, \mathcal{D})$ is a simple sphere system as claimed. □

It is not hard to show that for each Σ there is an admissible gluing datum (e.g by considering the dual graph to the intersection of Σ with σ as in the upcoming proof of Lemma 5.2.8). Since we do not need this statement in the sequel, we do not give a proof here.

We do however need the following relative version of this statement, which is the central ingredient for the proof of Theorem 5.2.4.

Lemma 5.2.8. *Let Σ be a simple sphere system, and let $\mathcal{D}$ be an admissible gluing datum for Σ. Suppose that Σ' is a simple sphere system which is disjoint from Σ up to homotopy. Then there is an admissible gluing datum $\mathcal{D}'$ for Σ' such that $\pi_\sigma(\Sigma, \mathcal{D})$ and $\pi_\sigma(\Sigma', \mathcal{D}')$ are disjoint up to homotopy.*

Proof. As a first step, note that if Σ' is obtained from Σ by removing some spheres then the claim is immediate – one can simply take $\mathcal{D}'$ as a subset of $\mathcal{D}$.

Since Σ' is disjoint from Σ, the union $\Sigma \cup \Sigma'$ is a simple sphere system (where we discard multiple copies of the same homotopy class of a sphere). The sphere system Σ' is then obtained from $\Sigma \cup \Sigma'$ by removing some number of spheres. Therefore, to show the lemma it suffices to consider the case that $\Sigma' \supset \Sigma$.

Let $\Sigma = \{\sigma_1, \dots, \sigma_r\}$ and let $\Sigma' = \{\sigma_1, \dots, \sigma_r, \sigma_1', \dots, \sigma_s'\}$. We call the sphere pieces defined by one of the σ_i *old sphere pieces* and those defined by one of the σ_j' *new sphere pieces*.

The gluing datum $\mathcal{D}$ contains a disk for each boundary component of every old sphere piece. We will construct $\mathcal{D}'$ inductively as an extension of $\mathcal{D}$. The boundary components of new sphere pieces fall into two different classes: those that are contained in some disk from the collection $\mathcal{D}$ and those that are disjoint from any disk in $\mathcal{D}$.

Let $D_1, \dots, D_k$ be the maximal disks in $\mathcal{D}$ with respect to the partial order defined by inclusion (this makes sense as $\mathcal{D}$ is admissible). Let α be a boundary component of a new sphere piece C' such that $\alpha \subset D_l$ for some l. We then choose the closing disk $D(\alpha)$ to be the unique embedded disk in σ bounded by α which is contained in D_l. Let $\mathcal{D}_1$ be the union of $\mathcal{D}$ with the set of all disks obtained in this way. By construction, $\mathcal{D}_1$ is properly nested in the sense defined before Lemma 5.2.5.

If $\mathcal{D}_1$ is a gluing datum for Σ' then we are done. Otherwise, consider the set of those boundary components $\alpha_1, \dots, \alpha_k$ of new sphere pieces which are disjoint from every disk in $\mathcal{D}$ and hence also from every disk in $\mathcal{D}_1$. If $k = 0$ there is nothing to show, so we may assume $k \geq 1$.

Let $I = \bigcup_{i=1}^k \alpha_i \subset \sigma$ and let T be the dual graph to I. Explicitly, T is the graph whose vertex set is the set of connected components of $\sigma \setminus I$. Two such vertices corresponding to components U_1, U_2 are joined by an edge if there is a component of I contained in the closure of both U_i. As every circle on a sphere is separating, the graph T is in fact a tree (see Figure 5.2).

Let v be a leaf of T, corresponding to a complementary component whose closure is a disk $D(v)$. This disk $D(v)$ intersects I in a single component $\alpha(v)$.

If $k \geq 2$ then $D(v)$ is the unique disk on σ bounded by $\alpha(v)$ which is disjoint from all $\alpha_i \neq \alpha(v)$. If $k = 1$ then $D(v)$ is one of the two embedded disks in σ bounded by α_1.

In both cases, if $D(v)$ intersects a disk $D \in \mathcal{D}_1$, then $D \subset D(v)$ since otherwise $\alpha(v) \subset D$.

Hence, the set of disks $\mathcal{D}_2 = \mathcal{D}_1 \cup \{D(v)\}$ is properly nested. Furthermore, the set of boundary components of new sphere pieces that are not contained

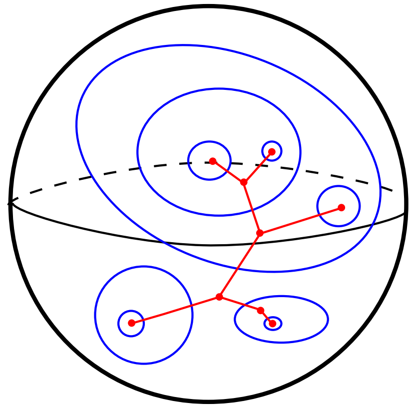


Figure 5.2: An example of a collection of boundaries of sphere pieces and the corresponding dual tree.

in any disk of $\mathcal{D}_2$ is $\{\alpha_1, \dots, \alpha_k\} \setminus \{\alpha(v)\}$.

The lemma now follows by repeating this procedure, assigning a closing disk to each α_i. ☐

Proof of Theorem 5.2.4. Let Σ, Σ' be two simple sphere systems containing σ. Choose an edge-path $\Sigma = \Sigma_1, \dots, \Sigma_L = \Sigma'$ of shortest length connecting Σ to Σ' in the sphere system graph $\mathcal{S}(M_n)$.

Since Σ_1 is disjoint from σ by assumption, $\mathcal{D}_1 = \emptyset$ is an admissible gluing datum for Σ_1.

By Lemma 5.2.8 there is an admissible gluing datum $\mathcal{D}_2$ for Σ_2 such that $\Sigma_1 = \pi_\sigma(\Sigma_1, \mathcal{D}_1)$ is disjoint from $\pi_\sigma(\Sigma_2, \mathcal{D}_2)$.

Inductively applying Lemma 5.2.8, one obtains admissible gluing data $\mathcal{D}_i$ for Σ_i such that $\pi_\sigma(\Sigma_i, \mathcal{D}_i)$ is disjoint from $\pi_\sigma(\Sigma_{i+1}, \mathcal{D}_{i+1})$ for all $i = 2, \dots, L-1$.

As Σ_L is disjoint from σ, the only admissible gluing datum is the empty set, and hence $\pi_\sigma(\Sigma_L, \mathcal{D}_L) = \Sigma_L$.

By construction, the sequence $\pi_\sigma(\Sigma_i, \mathcal{D}_i)$ for $1 \leq 1 \leq L$ defines an edge-path in $\mathcal{S}(M_n, \sigma)$ connecting Σ to Σ'. Thus the distance between Σ and Σ' in $\mathcal{S}(M_n)$ equals the distance between Σ and Σ' in $\mathcal{S}(M_n, \sigma)$ and the theorem is shown. ☐

5.3 Mapping class groups in $\mathrm{Out}(F_n)$

In this section we study an embedding of a surface mapping class group into $\mathrm{Out}(F_n)$. Let S_g^1 be a surface of genus g with one boundary component, and let $S_{g,1}$ be the surface obtained by collapsing the boundary component of S_g^1 to a marked point. We often view the marked point as a puncture of the surface, so that the fundamental group of $S_{g,1}$ is the free group F_{2g} on $2g$ generators.

A simple closed curve on $S_{g,1}$ which bounds a disk containing the marked point defines a distinguished conjugacy class in $\pi_1(S_{g,1})$ called the *cusp class*. The mapping class group of $S_{g,1}$ preserves the cusp class.

The following analog of the Dehn-Nielsen-Baer theorem for punctured surfaces is well-known (see e.g. Theorem 8.8 of [FM11]).

Theorem 5.3.1. *The homomorphism*

$$\iota : \mathrm{Map}(S_{g,1}) \to \mathrm{Out}(F_{2g})$$

induced by the action on the fundamental group of $S_{g,1}$ is injective. Its image consists of those outer automorphisms which preserve the cusp class.

The goal of this section is to prove

Theorem 5.3.2. *The homomorphism ι is a quasi-isometric embedding.*

We employ the following geometric model for the mapping class group of $S_{g,1}$. A *binding loop system for $S_{g,1}$* is defined to be a collection of embedded loops $\{a_1, \dots, a_n\}$ based at the marked point of $S_{g,1}$ which intersect only at the marked point and which decompose $S_{g,1}$ into a disjoint union of disks.

Let $\mathcal{BL}(S_{g,1})$ be the graph whose vertex set is the set of isotopy classes of binding loop systems. Here isotopies are required to fix the marked point. Two such systems are connected by an edge if they intersect in at most K points. As the mapping class group of $S_{g,1}$ acts with finite quotient on the set of isotopy classes of binding loop systems, we can choose the number $K > 0$ such that the following lemma is true.

Lemma 5.3.3. *The graph $\mathcal{BL}(S_{g,1})$ is connected. The mapping class group of $S_{g,1}$ acts on $\mathcal{BL}(S_{g,1})$ with finite quotient and finite point stabilizers.*

Instead of working with binding loop systems of $S_{g,1}$ directly we will frequently use binding arc systems of S_g^1. By this we mean a collection A of disjointly embedded arcs $\{a_1, \dots, a_n\}$ connecting the boundary component of S_g^1 to itself which decompose S_g^1 into simply connected regions. We will consider such binding arc systems up to isotopy of properly embedded arcs. A binding arc system for S_g^1 defines a binding loop system for $S_{g,1}$ by collapsing the boundary component of S_g^1 to the marked point. Note that if A_1, A_2 are two disjoint binding arc systems for S_g^1 then the corresponding binding loop systems for $S_{g,1}$ are uniformly close in $\mathcal{BL}(S_{g,1})$. By this we mean that there is a number $K > 0$ depending only on g such that the distance between the two binding loop systems in $\mathcal{BL}(S_{g,1})$ is at most K. The Dehn twist about the boundary component of S_g^1 acts trivially on the isotopy class of any arc system. Thus the action of the mapping class group of S_g^1 on binding arc systems factors through an action of $\mathrm{Map}(S_{g,1})$.

We can now describe the strategy of the proof of Theorem 5.3.2; details for each step will be given below. As in Section 5.2, we use simple sphere systems in M_{2g} to build a graph $\mathcal{S}_0(M_{2g})$ which is quasi-isometric to $\mathrm{Out}(F_{2g})$ (for technical reasons we choose a subgraph of $\mathcal{S}(M_{2g})$ in this section, see below for the definition). Let $\phi \in \mathrm{Map}(S_{g,1})$ be given and let F be a diffeomorphism of M_{2g} which represents the outer automorphism $\iota(\phi)$ of the fundamental group F_{2g}. We may choose F in such a way that it preserves an embedded surface $S_g^1 \subset M_{2g}$ and such that F restricts to a representative of ϕ on S_g^1. Now consider a shortest path $\Sigma_0, \Sigma_1, \dots, \Sigma_N = F(\Sigma_0)$ connecting a base sphere system to its image under F in $\mathcal{S}_0(M_{2g})$. The intersections of Σ_i with the surface $S_g^1 \subset M_{2g}$ then yield a sequence of binding arc systems $A_0, \dots, A_N$ on the surface S_g^1 and therefore a path of length coarsely bounded by N in the graph of binding loop systems of $S_{g,1}$. Each of the Σ_i is only determined up to homotopy and therefore the arc systems A_i are not defined canonically. The main technical difficulty now consists in obtaining enough control on the representatives of the homotopy classes to ensure that $A_N = \Sigma_N \cap S_g^1$ defines the homotopy class $\phi(A_0)$. To this end we also have to successively modify the surface S_g^1 by homotopies. Once this is done, the word norm of ϕ is coarsely bounded by N, and hence by the word norm of $\iota(\phi)$ in $\mathrm{Out}(F_{2g})$, showing Theorem 5.3.2.

We now define the geometric model of $\mathrm{Out}(F_{2g})$ used in this section. Let $M = M_{2g}$ be the connected sum of $2g$ copies of $S^2 \times S^1$. Say that a simple sphere system Σ for M is *reduced* if $M \setminus \Sigma$ is connected. Let $\mathcal{S}_0(M)$ be the complete subgraph of $\mathcal{S}(M)$ whose vertices correspond to reduced sphere systems. We call $\mathcal{S}_0(M)$ the *reduced sphere system graph*.

Lemma 5.3.4. *The graph* $\mathcal{S}_0(M)$ *is connected.*

This lemma can be shown using a surgery argument which is well-known in the analogous case of reduced disk systems for handlebodies (see e.g. [HH11a, Lemma 5.2], [St99] or [M86, Lemma 3.2]) For convenience of the reader we sketch a proof in the sphere system case here.

Proof. Let Σ, Σ' be two reduced sphere systems. We may assume without loss of generality that Σ and Σ' are in general position and hence the intersection is a disjoint union of finitely many circles. We prove the lemma by induction on the number of such intersection circles.

Let $\sigma' \in \Sigma'$ be a sphere which intersects Σ. The intersection $\sigma' \cap \Sigma$ is a disjoint union of finitely many circles $\alpha_1, \dots, \alpha_k$. There is at least one such circle $\alpha = \alpha_i$ which bounds a disk $D' \subset \sigma'$ whose interior contains no other intersection circle $\alpha_j, j \neq i$. Suppose that α is contained in the sphere $\sigma \in \Sigma$. Denote by D_1, D_2 the two embedded disks in σ bounded by α and put $\sigma_j = D_j \cup D'$ for $j = 1, 2$. Up to homotopy, the surface σ_j is a sphere which is disjoint from Σ.

We claim that the sphere system $\Sigma_j = \Sigma \cup \{\sigma_j\} \setminus \{\sigma\}$ is reduced for a suitable j. To prove the claim, note that a sphere system with $2g$ components is reduced if it defines a basis of $H_2(M, \mathbb{Z})$. Since Σ is reduced, for exactly one choice of $j = 1, 2$ the system Σ_j defines a basis of $H_2(M, \mathbb{Z})$ (the corresponding sphere σ_j has to separate the two sides of σ in the complement of Σ). This shows the lemma. ☐

The graph $\mathcal{S}_0(M)$ is $\mathrm{Out}(F_{2g})$-invariant. Moreover, the action of $\mathrm{Out}(F_{2g})$ on $\mathcal{S}_0(M)$ is properly discontinuous and thus $\mathcal{S}_0(M)$ is equivariantly quasi-isometric to $\mathrm{Out}(F_{2g})$.

The advantage of using reduced sphere systems is that they make it easy to encode free homotopy classes of curves in M. Namely, let $\Sigma = \{\sigma_1, \dots, \sigma_n\}$ be a reduced sphere system. We choose a transverse orientation for each sphere σ_i so we may speak of a positive and a negative side of σ_i.

Let $p \in M$ be a base point in the complement of Σ. A *basis dual* to Σ is a set of loops $\gamma_1, \dots, \gamma_n$ in M based at p such that the loop γ_i is disjoint from σ_j for all $j \neq i$ and intersects σ_i in a single point. We orient γ_i such that it approaches σ_i from the positive side. Since the complement of Σ is simply connected, the loops γ_i define a basis of $\pi_1(M, p)$.

Now let α be a closed curve in M. Choose an orientation of α. Up to applying a homotopy to α we may assume that α and Σ are in general position and thus intersect in a finite set of points. Apply a homotopy to α in the complement of Σ such that α passes through the basepoint p. The resulting based loop $\hat{\alpha}$ is a representative of the free homotopy class defined by α. Since the complement of Σ is simply connected, the sequence of (oriented) spheres in Σ which are consecutively hit by $\hat{\alpha}$ (and hence α) defines a word in the $\gamma_i^{\pm}$ representing $\hat{\alpha}$. In other words, the free homotopy class defined by α is determined by the sequence of sides of spheres in Σ that α intersects.

We next put α in tight position with respect to Σ as follows. Let M_Σ be the complement of Σ in the sense described for a single sphere in Section 5.2 – that is, M_Σ is a compact connected three-manifold whose boundary consists of $2n$ boundary spheres $\sigma_1^+, \sigma_1^-, \dots, \sigma_n^+, \sigma_n^-$. The boundary spheres σ_i^+ and σ_i^- correspond to the two sides of σ_i. If α is not disjoint from Σ then the intersection of α with M_Σ is a disjoint union of arcs connecting the boundary components of M_Σ. We call these arcs the *Σ–arcs of α*. An orientation of α induces a cyclic order on the Σ–arcs of α.

We say that α *intersects Σ minimally* if no Σ-arc of α connects a boundary component of M_Σ to itself.

Lemma 5.3.5. *Every closed curve α in M can be modified by a homotopy to intersects Σ minimally. Let α and α' be two simple closed curves which are freely homotopic and which intersect Σ minimally. Then there is a bijection f between the Σ–arcs of α and the Σ–arcs of α' such that $f(a)$ is homotopic to a for each Σ–arc a of α. If orientations of α and α' are chosen appropriately, f may be chosen to respect the cyclic orders on the Σ–arcs.*

Proof. Since M_Σ is simply connected, an arc in M_Σ connecting a boundary component to itself is homotopic into that boundary component. This shows the first claim.

To see the other claims, let p be a base point in the complement of Σ and let γ_i be a basis of $\pi_1(M,p)$ dual to Σ. The sequence of oriented spheres from Σ determined by the consecutive intersections of α defines a word in the γ_i representing the conjugacy class of α.

If α intersects Σ minimally, this word representing the conjugacy class of α is reduced and cyclically reduced. The analogous statements are also true for α'. Since α and α' are freely homotopic, they define the same conjugacy class in $\pi_1(M,p)$. Up to cyclic permutation, a conjugacy class in a free group contains a unique cyclically reduced word. Therefore, the words in γ_i defined by α and α' are equal up to cyclic permutation and possibly reversing the orientation of α'. This implies the lemma. □

Now let $V = S_g^1 \times [0,1]$ be the trivial oriented interval bundle over S_g^1. We identify $M = M_{2g}$ with the three-manifold obtained by doubling V along its boundary. To simplify notation, we put $n = 2g$.

As can be seen from the description of M as the double of V, the surface $S_g^1 \times \{\frac{1}{2}\}$ is incompressible in M. Let $\varphi_0 : S_g^1 \to M$ be the thus defined embedding of S_g^1 into M. Let β be the boundary curve of S_g^1. The image $\varphi_0(\beta)$ is an embedded closed curve in M which maps to the cusp class in $\pi_1(S_{g,1}) = \pi_1(M)$.

Next we put $\varphi_0(S_g^1)$ and $\varphi_0(\beta)$ in good position with respect to a given reduced sphere system. Since for surfaces and curves in M homotopy is in general not the same as isotopy, we need to take some care in defining these notions.

Consider the more general case of a surface $\varphi(S_g^1)$, where $\varphi : S_g^1 \to M$ is any embedding of S_g^1 into M which is homotopic to φ_0 (note that such an embedding need not be isotopic to φ_0). Up to modifying φ with a small isotopy, we may assume that Σ intersects the surface $\varphi(S_g^1)$ transversely. Then the preimage $\varphi^{-1}(\Sigma)$ is a one-dimensional submanifold of S_g^1, and hence it is a disjoint union of simple closed curves and properly embedded arcs.

Definition 5.3.6. We say that φ is in *ribbon position with respect to* Σ if each component of $\varphi^{-1}(\Sigma)$ is a properly embedded arc. It is said to be in *minimal position* if in addition $\varphi(\beta)$ is in minimal position with respect to Σ. In either case, we call the preimage $\varphi^{-1}(\Sigma)$ the *arc system induced by* Σ *and* φ.

Note that a priori the homotopy class of the arc system induced by Σ and φ need not be determined by the isotopy class of Σ even if φ is in minimal position with respect to Σ.

A main step towards the proof of Theorem 5.3.2 consists in establishing some control over the homotopy class of the arc system induced by φ and a sphere system associated to the an element in the image of $\mathrm{Map}(S_{g,1})$. To make this precise, fix an embedded binding arc system $A_0 = \{a_1^0, \dots, a_{2g}^0\}$ of S_g^1 consisting of precisely $2g$ arcs which cut $S_{g,1}$ into a single disc. We require that β intersects each arc a_i^0 in two points, and that a subarc of β defined by the intersection points with a_i^0 approaches a_i^0 from the same side at both of its endpoints. Such a binding arc system can easily be obtained from the standard description of S_g^1 as a one-holed $4g$–gon with opposite sides identified.

The interval bundle over A_0 is a disk system in the handlebody V which cuts V into a ball. Doubling this disk system across the boundary of V, we obtain a reduced simple sphere system Σ_0 in M_{2g} such that the arc system induced by φ_0 and Σ_0 is the arc system A_0. Similarly, any diffeomorphism f of $S_{g,1}$ extends in this way to a diffeomorphism $I(f)$ of M by first extending f to a product map of $S_{g,1}$ and then extending further by doubling. The action of $I(f)$ on homotopy classes of sphere systems then coincides with the action of the image of the projection of f to $\mathrm{Map}(S_g^1)$ under the inclusion $\iota : \mathrm{Map}(S_g^1) \to \mathrm{Out}(F_n)$. This observation is used in the following lemma which gives control on some classes of arc systems.

Lemma 5.3.7. *Let f be an orientation preserving diffeomorphism of S_g^1 and let Σ be a simple sphere system which is homotopic to $I(f)(\Sigma_0)$. Suppose that $\varphi : S_g^1 \to M$ is homotopic to φ_0 and in minimal position with respect to Σ. Then the arc system induced by φ and Σ is homotopic to $f(A_0)$.*

Proof. Let f be an orientation preserving diffeomorphism of S_g^1 and let $F = I(f)$. The sphere system $F(\Sigma_0)$ is then in minimal position with respect to φ_0, and $\varphi_0^{-1}(F(\Sigma_0)) = f(A_0)$ by construction. By applying an isotopy to M which maps Σ to $F(\Sigma_0)$ we may assume without loss of generality that $\Sigma = F(\Sigma_0)$. If we replace φ by its composition with this isotopy then this does not change $\varphi^{-1}(\Sigma)$.

With respect to the sphere system $\Sigma = F(\Sigma_0)$, the curve $\varphi_0(\beta)$ is in minimal position. Furthermore, by choice of the arc system A_0, the curve $\varphi_0(\beta)$ intersects each sphere σ_j in Σ in exactly two points x_j' and y_j'. Let $\beta_j'^1$ and $\beta_j'^2$ be the two subarcs of $\varphi_0(\beta)$ defined by these intersection points.

By Lemma 5.3.5, minimal position of curves is unique and only depends on the homotopy class of the curve and of the sphere system. Therefore the curve $\varphi(\beta)$ intersects each sphere $\sigma_j \in \Sigma$ also in two points, say x_j and y_j. Denote by β_j^1 and β_j^2 the two subarcs of $\varphi(\beta)$ defined by these intersection points. Again by uniqueness of minimal position of curves, the arc β_j^r is homotopic to $\beta_j'^r$ with endpoints sliding on σ_j for $r = 1, 2$ (after possibly exchanging β_j^1 and β_j^2).

Let $a_j \subset S_g^1$ be the preimage of σ_j under φ and let a_i' be the preimage of σ_j under φ_0. The boundary of a regular neighborhood of $\beta \cup a_j$ in S_g^1 is the union of two simple closed curves d_j^1, d_j^2 and the boundary curve β, and the boundary of a regular neighborhood of $\beta \cup a_j'$ consists of two simple closed curves $d_j'^1, d_j'^2$ and the boundary curve β.

Up to exchanging d_j^1 and d_j^2, the curve $\varphi(d_j^k) \subset M$ is freely homotopic to a curve $\delta_j^k = \beta_j^k * \alpha_j$ obtained by concatenating β_j^k and an embedded arc α_j on σ_j. Similarly, the curve $\varphi_0(d_j'^k)$ is freely homotopic to a curve $\delta_j'^k = \beta_j'^k * \alpha_j'$ obtained by concatenating $\beta_j'^k$ and an embedded arc α_j' on σ_i.

Since σ_j is simply connected and $\beta_j'^k$ is homotopic to β_j^k relative to σ_j, the curves $\delta_j'^k$ and δ_j^k are freely homotopic. Since φ and φ_0 induce the same isomorphism on the level of fundamental groups, this implies that also the simple closed curves d_j^k and $d_j'^k$ in S_g^1 are freely homotopic.

The curves β, d_j^1 and d_j^2 bound a pair of pants P_i on S_g^1. The arc a_j is up to isotopy the unique essential embedded arc in P_j connecting β to itself. Similarly, β, $d_j'^1$ and $d_j'^2$ bound a pair of pants P_j', which is isotopic to P_j. As a_j' is the unique essential embedded arc in P_j' connecting β to itself, it is therefore isotopic to a_j.

Hence we have shown that the arc system induced by a map φ and a sphere system Σ as in the statement is isotopic to the arc system induced by φ_0 and Σ, hence isotopic to $f(A_0)$. □

To apply Lemma 5.3.7 we have to keep φ in minimal position when changing the sphere system. For this we use an inductive method which is described in the next lemma. For its proof, we need the following observation, which also motivates the terminology "ribbon position".

Suppose that φ is in ribbon position with respect to the reduced sphere system Σ. Since φ is homotopic to φ_0, it induces an isomorphism between the fundamental groups of S_g^1 and M. The intersection of $\varphi(S_g^1)$ with M_Σ is a union of surfaces $P_1, \dots, P_k$. As Σ is a simple sphere system, the arc system $\varphi^{-1}(\Sigma)$ on S_g^1 is binding and hence each of the surfaces P_i is a disk whose boundary is not completely contained in a boundary component of M_Σ.

Pick one such disk, say P_i, and consider its boundary curve δ_i. We can write this curve in the form

$$\delta_i = a_1 * b_1 * \cdots * a_r * b_r$$

where each a_i is an arc contained in one of the boundary spheres of M_Σ, and each b_i is a properly embedded arc in M_Σ. Let $\Gamma_i \subset P_i$ be an embedded graph in P_i defined in the following way. The graph Γ_i has one distinguished vertex v_0 contained in the interior of P_i and one vertex v_r contained in each arc a_r. Each vertex v_r ($r \geq 1$) is connected by an edge to the vertex v_0. The oriented surface P_i determines a *ribbon structure* on Γ_i. Here a ribbon structure on Γ_i is simply a cyclic order of the half-edges at v_0.

To reconstruct $\varphi(S_g^1)$ from the ribbon graphs Γ_i we equip Γ_i with a *twisting datum.* Namely, fix the arcs a_r and an orientation of each of the arcs a_r, so that we can refer to the left and right endpoint of each a_r. A twisting datum on Γ_i associates to each edge of Γ_i a sign $+$ or $-$. We call the graph Γ_i equipped with a twisting datum a *decorated ribbon graph.*

The *surface associated to a decorated graph* Γ_i is defined in the following way. Put a small embedded oriented disk D at the central vertex v_0 of Γ_i containing a neighborhood of v_0 so that the cyclic order of the edges at v_0 corresponds to the counterclockwise order on D. Connect each arc a_r to the disk D with a *band*, i.e. an embedded product of two intervals $[0,1] \times [0,1]$ in M, as follows. One of the sides of B_r is the arc a_r, and the opposite side is contained in ∂D. We call these sides the *horizontal sides.* Correspondingly, the *vertical sides* are properly embedded arcs in M. The orientation of ∂D determines a left and right endpoint of each of these intervals. Up to homotopy, there are two ways to glue a band between two prescribed horizontal sides which correspond to the two ways of pairing the endpoints of these intervals. If the edge corresponding to the band B_r is decorated with a $+$, we match the left endpoint of a_r with the left endpoint of the interval on ∂D, otherwise we pair the left with the right endpoint. If the twisting data on Γ_i is chosen appropriately, the surface associated to Γ_i is homotopic to U_i relative to ∂M_Σ to U_i.

Lemma 5.3.8. *Suppose that φ is in minimal position with respect to Σ. Let σ' be an embedded sphere disjoint from Σ. Suppose that there is a sphere $\sigma \in \Sigma$ such that $\Sigma' = \Sigma \cup \{\sigma'\} \setminus \{\sigma\}$ is a reduced sphere system.*

Then there is an embedding $\varphi' : S_g^1 \to M$ with the following properties.

i) φ' is homotopic to φ.

ii) φ' is in minimal position with respect to Σ'.

iii) The arc system induced by φ' and Σ is the same as the arc system induced by φ and Σ.

Proof. By assumption, φ is in ribbon position with respect to Σ. Denote the components of $\varphi(S_g^1) \cap M_\Sigma$ by $P_1, \ldots, P_k$. By applying an isotopy to φ that does not change $\varphi^{-1}(\Sigma)$, we may assume that each P_i is the surface associated to a decorated ribbon graph Γ_i as described after Definition 5.3.6. We may choose Γ_i in such a way that no intersection point of σ' with Γ_i is a vertex of Γ_i and that σ' intersects each Γ_i transversely. Hence the intersection between σ' and Γ_i consists of a finite union of points, and the intersection between P_i and σ' consists of a disjoint union of arcs. Namely, the surface associated to a decorated ribbon graph may be chosen to lie in an arbitrarily small neighborhood of the graph.

As a consequence, the sphere σ' intersects each component of $\varphi(S_g^1)\cap M_\Sigma$ in a disjoint union of arcs. Each component of $\varphi(S_g^1) \cap M_{\Sigma\cup\{\sigma'\}}$ is a disk whose boundary contains a subarc of β and hence φ is in ribbon position with respect to $\Sigma \cup \{\sigma'\}$ and thus also with respect to Σ'.

It remains to show that φ can be changed by a homotopy as claimed in the lemma.

Let b be a Σ'–arc of β. Assume first that b also is a Σ–arc. Then b has both endpoints on a sphere distinct from σ. By assumption on Σ, the arc b does not connect the same boundary component of M_Σ to itself. This then also holds true for b viewed as a Σ'-arc.

If b is not of this form, at least one of its endpoints is contained in the sphere σ'. Suppose that both endpoints of b are contained on the same side of σ' (alternatively, on the same boundary component of $M_{\Sigma'}$). We call such subarcs of β *problematic*. A problematic subarc b does not intersect the sphere σ. Namely, we observed in the proof of Lemma 5.3.4 that in M_Σ, the sphere σ' separates the two boundary components corresponding to σ. Thus if b intersected σ, a subarc of b would return to the same side of σ. By assumption on Σ, this is not the case.

Let P_i be the component of $\varphi(S_g^1) \cap M_\Sigma$ containing b in its boundary. Choose small open tubular neighborhoods $\mathcal{U}_1 \subset \mathcal{U}_2$ of Σ so that $\overline{\mathcal{U}}_1 \subset \mathcal{U}_2$ and that $\overline{\mathcal{U}}_2$ is disjoint from σ'. We also assume that $\varphi(S_g^1)$ intersects $\overline{\mathcal{U}}_2$ in a union of disjoint embedded rectangles with two opposite sides on two different boundary components of $\overline{\mathcal{U}}_2$. Choose a homotopy H supported in the complement of $\mathcal{U}_1$ such that in the complement of $\mathcal{U}_2$ the image of P_i under this homotopy is Γ_i.

We compose this homotopy with φ such that the resulting map collapses P_i to the graph Γ_i. Explicitly, this means that we modify φ with a homotopy to obtain a map $\varphi_1 : S_g^1 \to M$ in the following way. On the set $\varphi^{-1}((M \setminus P_i) \cup \mathcal{U}_1)$, the maps φ and φ_1 coincide. On $\varphi^{-1}(P_i \setminus \mathcal{U}_2)$ we set φ_1 to be the postcomposition of φ with the endpoint of the homotopy H.

The map φ_1 is not an embedding of S_g^1 into M since it collapses the region $\varphi^{-1}(P_i \cap \mathcal{U}_2)$ to the graph $\Gamma_i \cap \mathcal{U}_2$. However, by construction, φ_1 is homotopic to φ_0 and the preimage $\varphi_1^{-1}(\Sigma)$ is the same as $\varphi^{-1}(\Sigma)$.

Since b is contained in a boundary arc of P_i and connects a side of σ' to itself, the same is true for the graph Γ_i. Since each complementary component of σ' in M_Σ is simply connected, the graph Γ_i is therefore homotopic with fixed endpoints to a graph Γ_i' in M_Σ which intersects σ' in fewer points than Γ_i and which is disjoint from all other $\Gamma_j, j \neq i$. The graph Γ_i' inherits the structure of a decorated ribbon graph from Γ_i.

We now modify φ_1 using this homotopy (in the same way that we constructed φ_1) to obtain a map $\varphi_2 : S_g^1 \to M$ which maps $\varphi^{-1}(P_i \cap \mathcal{U}_2)$ to $\Gamma_i' \cap \mathcal{U}_2$ and still agrees with φ on $S_g^1 \setminus \varphi^{-1}(P_i)$.

As a last step, we modify φ_2 by a homotopy to make it again an embedding. Namely, let P_i' be the surface defined by the decorated graph Γ_i' as described above. Then P_i' is homeomorphic to the disk P_i with a homeomorphism that induces an isomorphism of the decorated graphs Γ_i' and Γ_i and restricts to the identity on each component $P_i \cap \partial M_\Sigma$.

Hence we can apply a homotopy to the map φ_2 (supported on $\varphi^{-1}(P_i \setminus \mathcal{U}_1)$) to obtain a embedding $\varphi_3 : S_g^1 \to M$ with the following properties. On $\varphi^{-1}(M_\Sigma \setminus P_i)$, the maps φ_3 and φ agree. Furthermore, the set $\varphi^{-1}(P_i \cap \mathcal{U}_2)$ is mapped to the surface P_i' which can be chosen to be contained in a small regular neighborhood of Γ_i' in M. Finally, $\varphi^{-1}(\Sigma) = \varphi_3^{-1}(\Sigma)$. We can choose this homotopy such that φ_3 is in ribbon position with respect to Σ' by the same argument as before.

By construction, the image of β under φ_3 has fewer problematic arcs than the image of β under φ. The existence of the desired φ' follows then by inductively applying this procedure (with φ_3 in the place of φ). □

We now have collected all the necessary tools to prove the main theorem of this section.

Proof of Theorem 5.3.2. Let $f \in \mathrm{Map}(S_{g,1})$ be given. To prove the theorem, we need to show that the word norm of f as an element of the surface mapping class group is coarsely bounded by the word norm of $\iota(f)$ in $\mathrm{Out}(F_{2g})$.

The word norm of $\iota(f)$ in $\mathrm{Out}(F_{2g})$ is coarsely equal to the distance between Σ_0 and $\iota(f)(\Sigma_0)$ in the reduced sphere system graph.

Choose a shortest path connecting Σ_0 to $\iota(f)(\Sigma_0)$ in the reduced sphere system graph, and denote by $\Sigma_0, \Sigma_1, \ldots, \Sigma_N$ the corresponding sphere systems.

We now inductively define a sequence of binding arc systems. By construction, φ_0 is in minimal position with respect to Σ_0. As Σ_1 is connected to Σ_0 by an edge in the reduced sphere system graph, Σ_1 is obtained from Σ_0 by replacing a single sphere.

Thus Lemma 5.3.8 applies, and yields a reduced sphere system Σ'_1 which is homotopic to Σ_1 and disjoint from Σ_0, and furthermore an embedding φ_1. This embedding is homotopic to φ_0, in minimal position with respect to Σ_1 and such that $\varphi_1^{-1}(\Sigma_0) = \varphi_0^{-1}(\Sigma_0)$. Put $A_1 = \varphi_1^{-1}(\Sigma'_1)$. By the choice of φ_1, the arc system A_1 is binding and disjoint from A_0.

Inductively applying Lemma 5.3.8, we obtain a sequence of sphere systems Σ'_i and embeddings $\varphi_i : S^1_g \to M$ such that the following holds. Each Σ'_i is homotopic to Σ_i and each φ_i is homotopic to φ_0. Furthermore, the arc systems A_i induced by Σ'_i and φ_i define a path in the graph $\mathcal{BL}(S_{g,1})$ whose length is coarsely bounded by N.

By Lemma 5.3.7 the arc system A_N is homotopic to $f(A_0)$. Hence, as the binding loop system graph is quasi-isometric to $\mathrm{Map}(S_{g,1})$, the theorem follows. □

The method employed in the proof of Theorem 5.3.2 has another application. For its formulation, recall that the *arc graph of* S^1_g is the graph whose vertex set is the set of isotopy classes of embedded essential arcs connecting the boundary of S^1_g to itself. Again, isotopies are only required to fix the boundary component setwise. Two such vertices are joined by an edge if the corresponding arcs can be embedded disjointly. Similarly, define the *sphere graph of* M to be the graph whose vertex set is the set of isotopy classes of essential 2-spheres in M. Two such vertices are connected by an edge if the corresponding spheres can be realized disjointly.

Let a be an arc representing a vertex of the arc graph of S^1_g. The interval bundle over a is a disk $D(a)$ in the handlebody $V = S^1_g \times [0,1]$. The isotopy class of this disk only depends on the isotopy class of a, since the Dehn twist about the boundary of S^1_g is contained in the kernel of the map $\mathrm{Map}(S^1_g) \to \mathrm{Map}(V)$. We let $\sigma(a)$ be the essential sphere in M which is obtained by doubling $D(a)$ along ∂V.

Proposition 5.3.9. *The map sending a to $\sigma(a)$ induces a quasi-isometric embedding of the arc graph of S^1_g into the sphere graph of M.*

In particular, this theorem immediately implies the following.

Corollary 5.3.10. *For each $g \geq 1$ the sphere graph of M_{2g} has infinite diameter.*

Proof of Proposition 5.3.9. Let a, a' be two essential arcs in S_g^1. Since the mapping class group of S_g^1 acts transitively on the set of isotopy classes of essential arcs in S_g^1, there is a mapping class f such that $f(a) = a'$. Furthermore, we may assume that a is contained in the standard arc system A_0.

A single arc in S_g^1 does not separate the surface S_g^1. Thus the sphere $\sigma(a)$ is a nonseparating essential sphere in M.

Let $\sigma(a) = \sigma_1, \sigma_2, \ldots, \sigma_N = \sigma(a')$ be a shortest path in the sphere graph of M. We may assume without loss of generality that each σ_i is a nonseparating sphere. Namely, suppose that σ_i is separating and let M_1, M_2 be its two complementary components. If σ_{i-1} and σ_{i+1} are contained in different components, then they are connected by an edge in the sphere graph. In this case, the sphere σ_i can be removed from the edge-path. If σ_{i-1} and σ_{i+1} are contained in the same component, say M_1, then one can replace σ_i by a nonseparating sphere σ_i' contained in M_2.

Choose reduced sphere systems Σ_i containing σ_i. Let $\Sigma_i^{(1)}, \ldots, \Sigma_i^{(N_i)}$ be a path in the reduced sphere system graph connecting Σ_i to Σ_{i+1} such that each $\Sigma_i^{(j)}$ contains σ_i for each $1 \leq j \leq N_i - 1$.

We now argue as in the proof of Theorem 5.3.2. Applying Lemma 5.3.8 inductively, we change the sphere systems $\Sigma_i^{(j)}$ by isotopy and obtain a sequence of embeddings $\varphi_i^{(j)}$ which intersect $\Sigma_i^{(j)}$ minimally. Let $A_i^{(j)}$ be the arc systems induced by $\varphi_i^{(j)}$ and $\Sigma_i^{(j)}$.

By construction, for $1 \leq j \leq N_i - 1$ the arc systems $A_i^{(j)}$ contain a common arc a_i. The sequence a_i defines an edge-path in the arc graph of length at most $2N$. Furthermore, by Lemma 5.3.5, the arc a_N is contained in $f(A_0)$ and thus is adjacent to a'. This proves the theorem. □

5.A Stabilizers of free factors and free splittings

In this appendix we identify the stabilizers of conjugacy classes of free splittings and corank one free factors of a free group topologically. To this end, let M_n be the connected sum of n copies of $S^1 \times S^2$. As explained in Section 5.2, the mapping class group of M_n projects onto $\mathrm{Out}(F_n)$ with a finite kernel. Our goal is to give an elementary topological proof of the following

Lemma 5.A.1. *i) Let σ be an essential separating sphere in M_n. Then the stabilizer of σ in $\mathrm{Map}(M_n)$ projects onto the stabilizer of the conjugacy class of a free splitting in $\mathrm{Out}(F_n)$. Furthermore, every stabilizer of a conjugacy class of a free splitting arises in this way.*

ii) Let σ be an essential nonseparating sphere in M_n. Then the stabilizer of σ in $\mathrm{Map}(M_n)$ projects onto the stabilizer of the conjugacy class of a corank 1 free factor in $\mathrm{Out}(F_n)$. Furthermore, every stabilizer of a conjugacy class of a corank 1 free factor arises in this way.

Proof. Let σ be as in *i)*, and denote by M^1 and M^2 the two complementary components of σ in M_n. We let $N^i = M^i \cup \sigma$. Since σ is simply connected, the van-Kampen theorem yields that the fundamental group of M_n can be written as a free product $\pi_1(M_n) = \pi_1(N^1) * \pi_1(N^2)$. The fundamental groups of N^1 and N^2 are thus free groups of rank $n-i$ and i, respectively. A mapping class of M_n that stabilizes σ (up to homotopy) induces an outer automorphism of $\pi_1(M_n)$ that stabilizes the free splitting $\pi_1(M_n, x) = \pi_1(N^1, x) * \pi_1(N^2, x)$ up to conjugation (here, x is an arbitrary basepoint on σ).

Conversely, let $[\varphi] \in \mathrm{Out}(F_n)$ be an outer automorphism fixing the conjugacy class of the free splitting $F_n = \pi_1(N^1, x) * \pi_1(N^2, x)$. We can choose a representative φ which fixes the free splitting itself. Such an automorphism φ induces automorphisms of the groups $\pi_1(N^1, x)$ and $\pi_1(N^2, x)$. By the pointed version of Theorem 5.2.2 ([L74, Théorème 4.3, part 1)]), there are homeomorphisms f_i of N_i which induce $\varphi|_{\pi_1(N_i,x)}$ on the respective fundamental groups. By gluing f_1 and f_2 across σ we obtain a homeomorphism of M_n which fixes S and which induces $[\varphi]$ as desired. This shows that the stabilizer of σ maps onto the stabilizer of the conjugacy class of the free splitting $\pi_1(M_n, x) = \pi_1(N_1, x) * \pi_1(N_2, x)$.

Let now $F_n = G * H$ be an arbitrary free splitting, where G has rank i and H has rank $n - i$. Choose a sphere σ_i separating M_n into N^1 and N^2 as above, such that the rank of $\pi_1(N^1, x)$ is i (and thus the rank of $\pi_1(N^2, x)$ is $n - i$). Since the automorphism group of F_n acts transitively on the set of free splittings with fixed ranks, the last sentence of part *i)* follows from Theorem 5.2.2.

To prove part *ii)*, let σ be a non-separating sphere. Choose a basepoint $p \in M \setminus \sigma$. Then the subgroup $G < \pi_1(M, p) = F_n$ of all homotopy classes of loops which do not intersect σ is a free factor of corank one. Any diffeomorphism of M which preserves σ also preserves the conjugacy class of

G. Therefore the stabilizer of σ in $\mathrm{Out}(F_n)$ injects into the stabilizer of the conjugacy class of G.

To show that it is equal to this stabilizer, let $\varphi \in \mathrm{Out}(F_n)$ be an outer automorphism which preserves the conjugacy class of G. We may choose a diffeomorphism f of M which fixes p and such that the induced isomorphism f_* of the fundamental group is contained in the conjugacy class defined by φ and fixes G.

Let $\sigma' = f(\sigma)$ be the image of σ under f. Since f_* preserves the group G, the subgroup of all homotopy classes of loops which do not intersect σ' is equal to G.

By Lemma 2.2 of [HV98] the group G is thus the subgroup of $\pi_1(M, p)$ defined by all homotopy classes of loops which do not intersect both σ and σ' simultaneously. We now argue by contradiction, supposing that σ and σ' are not homotopic.

Suppose first that σ' and σ are disjoint up to homotopy. Then the fundamental group of the complement of $\sigma \cup \sigma'$ has rank at most $n-2$. This is a contradiction since G has rank $n-1$.

If σ and σ' intersect, we argue similarly. Namely, at least one connected component of $\sigma' \setminus \sigma$ is an open disk. Let D be the closure of this component in σ'. The surface D is a closed disk whose boundary curve ∂D is contained in σ. Let D' be a complementary component of ∂D on σ. The union $S = D \cup D'$ is an essential sphere which, up to homotopy, is disjoint from σ.

Furthermore, every loop which is disjoint from both σ and σ' is also disjoint from S'. Thus G can be identified with the subgroup of $\pi_1(M, p)$ of those loops which are disjoint from σ, σ' and S.

Since σ and S are disjoint, the fundamental group of the complement of $\sigma \cup S$ has rank at most $n-2$. Since removing σ' as well decreases the rank of the fundamental group further, this again contradicts the fact that G has rank $n-1$.

Thus, f preserves the homotopy class of σ. $\square$

Bibliography

[A02] E. Alibegović, *Translation Lengths in* $\mathrm{Out}(F_n)$, Geom. Dedicata 92 (2002), 87–93.

[AS09] J. Aramayona, J. Souto, *Automorphisms of the graph of free splittings*, arXiv:0909.3660.

[BH92] M. Bestvina, M. Handel, *Train tracks and automorphisms of free groups*, Ann. of Math. 135(1) (1992), 1–51.

[BFH97] M. Bestvina, M. Feighn, M. Handel, *Laminations, trees, and irreducible automorphisms of free groups*, Geom. Funct. Anal. 7(2) (1997), 215–244 (Erratum: Geom. Funct. Anal. 7(6) (1997), 1143).

[Bi74] J. S. Birman, *Braids, links and mapping class groups*, Ann. Math. Studies 82, Princeton University Press, Princeton, N.J. 1974.

[BS00] M. Bonk, O. Schramm, *Embeddings of Gromov hyperbolic spaces* Geom. Funct. Anal. 10 (2000) no. 2, 266–306

[Bo85] P. L. Bowers, *Dense embeddings of sigma-compact, nowhere locally compact metric spaces*, Proc. Amer. Math. Soc. 95 (1985) no. 1, 123–130

[B06] B. Bowditch, *Intersection numbers and the hyperbolicity of the curve complex*, J. Reine Angew. Math. 598 (2006), 105–129.

[BH99] M. R. Bridson, A. Haefliger, *Metric spaces of non-positive curvature*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Volume 319, Springer-Verlag, Berlin, 1999

[BV95] M. Bridson, K. Vogtmann, *On the geometry of the group of automorphisms of a free group*, Bull. London Math. Soc. 27 (1995), 544–525.

[BV10] M. Bridson, K. Vogtmann, *The Dehn functions of* $\mathrm{Out}(F_n)$ *and* $\mathrm{Aut}(F_n)$, arXiv:1011.1506

[BFP07] N. Broaddus, B. Farb, A. Putman, *Irreducible Sp-representations and subgroup distortion in the mapping class group*, arXiv:0707.2262, to appear in Comm. Math. Helv.

[CV86] M. Culler, K. Vogtmann, *Moduli of graphs and automorphisms of free groups*, Invent. Math. 84 (1) (1986), 91–119.

[Cu85] D. W. Curtis, *Boundary sets in the Hilbert cube*, Topology Appl. 20 (1985) no. 3, 201–221

[Du85] J. Dugundji, *Absolute neighborhood retracts and local connectedness in arbitrary metric spaces*, Compositio Math. 13 (1958), 229–246

[En78] R. Engelking, *Dimension theory*, Translated from the Polish and revised by the author; North-Holland Mathematical Library 19, North-Holland Publishing Co., Amsterdam, 1978

[ECHLPT] D. Epstein, J. Cannon, D. Holt, S. Levy, M. Paterson and W. Thurston, *Word Processing in Groups*, London, Jones and Bartlett (1992).

[FLM01] B. Farb, A. Lubotzky, Y. Minsky, *Rank one phenomena for mapping class groups*, Duke Math. J. 106 (2001), 581–597.

[FM11] B. Farb, D. Margalit, *A primer on mapping class groups*, Princeton Mathematical Series, Princeton University Press, Princeton N.J. 2011.

[G09] D. Gabai, *Almost filling laminations and the connectivity of ending lamination space*, Geom. Topol. 13 (2009) no. 2, 1017–1041

[G11] D. Gabai, *On the topology of ending lamination space*, arXiv:1105.3648

[Gr87] M. Gromov, *Hyperbolic Groups*, in *Essays in Group Theory*, ed. by S. M. Gersten, M.S.R.I., Springer-Verlag, 8 (1987), 75–263.

[Ha06a] U. Hamenstädt, *Train tracks and the Gromov boundary of the complex of curves*, in *Spaces of Kleinian groups*, London Math. Soc. Lecture Note Ser. 329, Cambridge Univ. Press, Cambridge, 2006, 187–207

[Ha06b] U. Hamenstädt, *Geometric properties of the mapping class group*, in *Problems on mapping class groups and related topics*, Proc. Sympos. Pure Math. 74, Amer. Math. Soc., Providence, RI, 2006, 215–232

[Ha09a] U. Hamenstädt, *Geometry of the mapping class group I: Boundary amenability*, Invent. Math. 175 (2009), 545–609.

[Ha09b] U. Hamenstädt, *Geometry of the mapping class group II: A biautomatic structure*, arXiv:0912.0137.

[HH11a] U. Hamenstädt, S. Hensel, *The geometry of the handlebody groups I: Distortion*, arXiv:1101.1838.

[HH11b] U. Hamenstädt, S. Hensel, *Sphere systems, intersections and the geometry of* $\mathrm{Out}(F_n)$, arXiv:1109.2687.

[HH11c] U. Hamenstädt, S. Hensel, *Isoperimetric inequalities for the handlebody groups*, arXiv:1109.5255.

[HM10] M. Handel, L. Mosher, *Lipschitz retraction and distortion for subgroups of* $\mathrm{Out}(F_n)$, arXiv:1009.5018.

[Ha95] A. Hatcher, *Homological stability for automorphism groups of free groups*, Comm. Math. Helv. 70 (1995), 39–62.

[HV96] A. Hatcher, K. Vogtmann, *Isoperimetric inequalities for automorphism groups of free groups*, Pacific Journal of Mathematics, Vol. 173, No. 2, 1996.

[HV98] A. Hatcher, K. Vogtmann, *The Complex of Free Factors of a Free Group*, Quart. J. Math. Oxford Ser. (2) 49 (1998), no. 196, 459–468.

[HP11] S. Hensel, P. Przytycki, *The ending lamination space of the five-punctured sphere is the Nöbeling curve*, J. London Math. Soc. (2) 84 (2011), 103–119.

[I02] N. Ivanov, *Mapping class groups*, Handbook of geometric topology, 523–633, North-Holland, Amsterdam, 2002.

[KLT97] K. Kawamura, M. Levin, E. D. Tymchatyn, *A characterization of 1-dimensional Nöbeling spaces*, in *Proceedings of the 12th Summer Conference on General Topology and its Applications (North Bay, ON, 1997)*, Topology Proc. 22 (Summer 1997), 155–174

[K99] E. Klarreich, *The boundary at inifinity of the curve complex and the relative Teichmüller space*, electronic preprint at `http://www.msri.org/people/members/klarreic/curvecomplex.ps` (1999)

[L74] F. Laudenbach, *Topologie de la dimension trois. Homotopie et isotopie*, Asterisque 12 (1974)

[LS08] C. Leininger, S. Schleimer, *Connectivity of the space of ending laminations*, arXiv:0801.3058

[LMS08] C. Leininger, M. Mj, S. Schleimer, *Universal Cannon–Thurston maps and the boundary of the curve complex*, arXiv:0808.3521

[Lu00] F. Luo, *Automorphisms of the complex of curves*, Topology 39 (2000) no. 2, 283–298

[M86] H. Masur, *Measured foliations and handlebodies*, Erg. Th. & Dynam. Sys. 6 (1986), 99–116.

[MM99] H. Masur, Y. Minsky, *Geometry of the complex of curves I: Hyperbolicity*, Invent. Math. 138 (1999), 103–149.

[MM00] H. Masur, Y. Minsky, *Geometry of the complex of curves II: Hierarchical structure*, Geom. Funct. Anal. 10 (2000), 902–974.

[MS10] H. Masur, S. Schleimer, *The geometry of the disk complex*, arXiv:1010.3174

[MMS10] H. Masur, L. Mosher, S. Schleimer, *On train track splitting sequences*, arXiv:1004.4564

[Mo03] L. Mosher, *Train track expansions of measured foliations*, unpublished manuscript (2003), availible at `http://andromeda.rutgers.edu/~mosher/arationality_03_12_28.pdf`

[McC85] D. McCullough, *Twist groups of compact 33-manifolds*, Topology 24 (1985), no. 4, 461–474.

[Mo95] L. Mosher, *Mapping class groups are automatic*, Ann. of Math. (2) 142 (1995), no. 2, 303–384.

[N06] A. Nagórko, *Characterization and topological rigidity of Nöbeling manifolds*, arXiv:0602574

[PH92] R. C. Penner, J. L. Harer, *Combinatorics of train tracks*, Annals of Mathematics Studies 125, Princeton University Press, Princeton, NJ, 1992

[R07] K. Rafi, *A combinatorial model for the Teichmüller metric*, Geom. Funct. Anal. 17 (2007), no. 3, 936–959.

[St99] J. Stallings, *Whitehead graphs on handlebodies*, Geometric group theory down under (Canberra, 1996), 317–330, de Gruyter, Berlin, 1999.

[S77] S. Suzuki, *On homeomorphisms of a 3-dimensional handlebody*, Canad. J. Math. 29 (1977), no. 1, 111-124.

[T80] W. P. Thurston, *The geometry and topology of three-manifolds*, Princeton Univ. Math. Depr. Lecture Notes, 1980, electronic version availible at `http:msri.org/publications/books/gt3m/`

[V02] K. Vogtmann, *Automorphisms of Free Groups and Outer Space*, Geom. Dedicata 94 (2002), 1–31.

[Wa98] B. Wajnryb, *Mapping class group of a handlebody*, Fund. Math. 158 (1998), no. 3, 195–228.

[W79] S. Wolpert, *The length spectra as moduli for compact Riemann surfaces*, Ann. of Math. (2) 109 (1979), no. 2, 323–351.