

The twist-cofinite topology on the mapping class group of a surface

Ingrid Irmer

Abstract. A topology is defined on the mapping class group $\text{Mod}(\Sigma)$ of a compact, connected, orientable surface Σ . It is shown that a notion of “genericity” on subsets of $\text{Mod}(\Sigma)$ arises from this definition. Many plausible results follow from this notion easily; for example, the set of pseudo-Anosov maps is shown to be generic and can be assumed to have arbitrarily large stretch factor, generically. Let M be a 3-manifold obtained from a Heegaard splitting of fixed genus g and generic gluing map. It is shown that for such manifolds, generically $b_1(M) = 0$, M is hyperbolic and M has Heegaard genus exactly g .

1. Introduction

Let Σ be a closed, compact, connected, orientable surface of genus $g \geq 2$. The mapping class group of Σ is the group $\text{Mod}(\Sigma) = \pi_0(\text{Diff}^+(\Sigma))$ of isotopy classes of positively oriented diffeomorphisms of Σ . For any property $\mathcal{P}(f)$ depending on a mapping class $f \in \text{Mod}(\Sigma)$, one may ask whether the property $\mathcal{P}(f)$ is true for a generic element of $\text{Mod}(\Sigma)$.

Such questions have been investigated by various authors, for example, [3, 12, 19, 20]. Although there are many possible notions of genericity of a mapping class group element, in most of the above works, genericity is defined using the following measure theoretic definition: Choose a finite generating set S of $\text{Mod}(\Sigma)$, and let W_n be a random variable in $\text{Mod}(\Sigma)$ which is obtained by an n -step random walk in the Cayley graph of $\text{Mod}(\Sigma)$, starting at the identity. The property \mathcal{P} is said to be generic if, in the limit as $n \rightarrow \infty$, the probability that \mathcal{P} is true for an element W_n approaches 1. In the papers [4, 19], several versions of the notion of random elements of $\text{Mod}(\Sigma)$ are developed, but these notions of genericity stay probabilistic in nature.

For subsets of \mathbb{R} , there are two alternative notions of being generic:

- A measure theoretic one, stating that a subset $S \subset \mathbb{R}$ is generic (in the sense of Lebesgue) if $\mathbb{R} \setminus S$ has Lebesgue measure 0.
- A topological one, stating that S is generic (in the sense of Baire) if S contains a countable intersection of open dense subsets of \mathbb{R} .

Mathematics Subject Classification 2020: 20F65 (primary); 57M50 (secondary).

Keywords: mapping class group, generic group element, Heegaard splitting, mapping torus.

The aim of this paper is to develop a topological notion of genericity of a subset of the mapping class group. In order to achieve this goal, a topology on $\text{Mod}(\Sigma)$ called the *twist-cofinite topology* was defined.¹ For γ a non-trivial simple closed curve on Σ , let D_γ be the Dehn twist around γ .

Definition 1 (Twist-cofinite topology). A set of open sets is given by

$$\mathcal{O} = \{U \subset \text{Mod}(\Sigma) \mid \forall f \in U, \forall \gamma, f \circ D_\gamma^n \in U \text{ for all except at most finitely many } n \in \mathbb{Z}\}.$$

Theorem 2. *The open sets \mathcal{O} define a topology on $\text{Mod}(\Sigma)$, and $\text{Mod}(\Sigma)$ is a connected topological group for this topology.*

A fundamental property of the twist-cofinite topology, similar, for example, to the Zariski topology, is that open sets are quite large. Specifically, a subset S of $\text{Mod}(\Sigma)$ will be called *thick* if S has non-empty interior. The next proposition suggests that thickness of a subset is a good notion of genericity.

Proposition 3. *In the topology \mathcal{O} ,*

- *the intersection $U_1 \cap U_2$ of two non-empty open sets U_1 and U_2 is non-empty,*
- *any thick subset is dense,*
- *the intersection $S_1 \cap S_2$ of two thick subsets is thick.*

With this notion of genericity, generic properties of subsets of $\text{Mod}(\Sigma)$ will be studied. For example, see the following.

Theorem 4. *The following subsets of $\text{Mod}(\Sigma)$ are non-empty and open (i.e., thick):*

- *The set of pseudo-Anosov maps $f \in \text{Mod}(\Sigma)$.*
- *The sets $\{f \in \text{Mod}(\Sigma) \text{ pseudo-Anosov} \mid \text{Vol}(M_f) > C\}$, where M_f is the mapping torus of f and $C > 0$.*
- *The sets $\{f \in \text{Mod}(\Sigma) \text{ pseudo-Anosov} \mid \lambda(f) > C\}$, where $\lambda(f)$ is the stretch factor of f and $C > 0$.*
- *The set of pseudo-Anosovs that do not leave any integer homology class of curves on Σ invariant.*

In [3], Dunfield and Thurston utilised the measure theoretic notion of genericity in $\text{Mod}(\Sigma)$ and Heegaard splittings to study the genericity of various properties of 3-manifolds. The same notion of random 3-manifolds was studied in subsequent papers [12, 19, 20] by various authors.

Similar results can be obtained with the notion of thickness/genericity defined above.

¹R. Detcherry, unpublished personal communication (email), 2019.

Let H_A and H_B be a pair of handlebodies, with $\partial H_A = \partial H_B = \Sigma$. For $f \in \text{Mod}(\Sigma)$, let N_f be the 3-manifold:

$$N_f = H_A \cup_f H_B,$$

where f is called the gluing map. The properties of 3-manifold invariants of N_f for a generic f can be studied. Some of the simplest such invariants are the first Betti numbers. Denote by $b_1(N_f, \mathbb{Z})$ the first Betti number of N_f and by $|H_1(N_f; \mathbb{Z})|$ the number of elements in the group $H_1(N_f; \mathbb{Z})$.

Theorem 5. *The following are all thick:*

- (1) *The set $\{f \in \text{Mod}(\Sigma) \mid b_1(N_f, \mathbb{Z}) = 0\}$.*
- (2) *The set $\{f \in \text{Mod}(\Sigma) \mid |H_1(N_f; \mathbb{Z})| \geq k\}$ for any integer $k \geq 0$.*

Suppose p is prime and $0 \leq i \leq g$. In contrast, the set $\{f \in \text{Mod}(\Sigma) \mid b_1(N_f, \mathbb{F}_p) = i\}$ is dense but not thick.

The following will also be proven using somewhat different techniques.

Theorem 6. *The sets $\{f \in \text{Mod}(\Sigma) \mid g(N_f) = g(\Sigma)\}$ and $\{f \in \text{Mod}(\Sigma) \mid g(N_f) \geq n\}$ for every $n \geq 2$ are thick, where $g(N_f)$ is the Heegaard genus of N_f .*

A natural question to ask is whether, with this notion of genericity, a generic Heegaard splitting gives an irreducible manifold and whether it gives a hyperbolic manifold. Note that due to the geometrisation theorem, being hyperbolic is equivalent to being irreducible and atoroidal. A more general question is therefore whether N_f admits an embedded incompressible surface of genus less than or equal to g .

Theorem 7. *For any $g \geq 0$, the set*

$$\{f \in \text{Mod}(\Sigma) \mid N_f \text{ has no embedded incompressible surface of genus } \leq g\}$$

is thick.

This theorem will be reduced to studying a property of Hempel distance in Harvey's complex of curves. The Hempel distance also gives a lower bound on the genus of an incompressible surface. Theorem 7 will be shown to follow from Theorem 8.

Theorem 8. *For any $d \geq 3$, the set*

$$\{f \in \text{Mod}(\Sigma) \mid \text{the Heegaard splitting induced by } f \text{ has distance } \geq d\}$$

is an open dense set.

While most of the results in this paper follow almost immediately from the definitions or from classical results, the proof of a lemma needed for these last two theorems is quite technical.

On page 3 of [20], a list of properties of elements of subgroups of mapping class groups represented by random long words in a symmetric generating set is given. On page 4 of the same paper, a list of properties of random Heegaard splittings according to the Dunfield–Thurston model is given. The properties investigated for genericity in this paper are based on these two lists. Results in this paper should be compared with theorems in [20] and the extensive list of references given therein. Given the amount of literature available on the subject, the author apologises that no attempt is made here to do justice to related results in different frameworks.

As a final comment, the notion of genericity developed here has advantages and disadvantages. The measure theoretic definition of genericity a priori depends on the choice of a finite generating set of $\text{Mod}(\Sigma)$. Genericity of a property in that setting is often shown for one generating set only, such as the set of Humphries generators. It follows that the topological notion of genericity is more canonical, as it is less dependent on such choices. In addition, it is suited to formulating problems about subgroups of mapping class groups. These properties can be understood in terms of subset topologies. An example of such a problem is the well-known Ivanov conjecture [9]. When Σ has genus at least 3, the Ivanov conjecture predicts that any finite index subgroup Γ of $\text{Mod}(\Sigma)$ satisfies $H_1(\Gamma; \mathbb{R}) = 0$.

On the other hand, the measure theoretic version is better suited to asking questions about the asymptotic growth of invariants (stretch factor, volume of the mapping torus, etc.) of “generic mapping classes”. It is possible to simply study the average value of the invariant of the n -step random walk W_n .

The paper is organised as follows: In Section 2, topologies $\mathcal{O}_{G,S}$ on groups G equipped with a generating set S are defined. The general properties of those topologies are studied, and it is shown that the twist-cofinite topology is obtained as an example of such a topology, taking $G = \text{Mod}(\Sigma)$ and S to be the set of all Dehn twists. Theorem 2 and Proposition 3 are corollaries of this more general theory. In Section 3, generic properties of elements of $\text{Mod}(\Sigma)$ are discussed, and Theorem 4 is proven. Sections 4 and 5 deal with properties of 3-manifolds obtained from generic Heegaard splittings. The proof of Theorem 8 requires concepts such as subsurface projections, due to Masur and Minsky, and properties of handlebody sets, due to Masur and Schleimer. This background material is given in Subsections 5.2 and 5.3, as well as a crucial property of bounded and unbounded diameter subsurface projections of handlebody sets. The remainder of Section 5 deals mainly with showing the existence of certain large subsurface projections needed to obtain distance bounds in the proof of Theorem 8.

2. A topology on groups with generating sets

Let G be a group and S be a generating set for G . For $g \in G$, let L_g be the left multiplication

$$\begin{aligned} L_g : G &\longrightarrow G \\ h &\longrightarrow gh \end{aligned}$$

and R_g be the right multiplication

$$\begin{aligned} R_g : G &\longrightarrow G \\ h &\longrightarrow hg. \end{aligned}$$

Let also $s : G \rightarrow G$ be the map such that $s(g) = g^{-1}$. Two topologies on G are defined.

Definition 9. Let

$$\mathcal{O}_{G,S}^l = \{U \subset G \mid \forall g \in U, \forall x \in S \text{ for all except at most finitely many } n \in \mathbb{Z}, gx^n \in U\}$$

and

$$\mathcal{O}_{G,S}^r = \{U \subset G \mid \forall g \in U, \forall x \in S \text{ for all except at most finitely many } n \in \mathbb{Z}, x^n g \in U\}.$$

$$\text{Let } \mathcal{O}_{G,S} = \mathcal{O}_{G,S}^r \cap \mathcal{O}_{G,S}^l.$$

Theorem 10. *The sets $\mathcal{O}_{G,S}^l$ (resp. $\mathcal{O}_{G,S}^r$) determine a topology on G for which the operator L_g (resp. R_g) is continuous for any $g \in G$. In addition, G is a topological group for the topology $\mathcal{O}_{G,S}$, and G is connected for all three topologies.*

The left (resp. right) cyclically cofinite topology on G relative to the generating set S refers to the set of open sets $\mathcal{O}_{G,S}^l$ (resp. $\mathcal{O}_{G,S}^r$). The set of open sets $\mathcal{O}_{G,S}$ determines the cyclically cofinite topology on G relative to S .

The motivation for those names is that the intersection of an open set in $\mathcal{O}_{G,S}^l$ with a left coset of a cyclic subgroup of G generated by an element of S is either empty or cofinite.

In the proof of Theorem 10, the assumption that S is a generating set is only used to show connectivity.

Proof. It will be shown that $\mathcal{O}_{G,S}^l$ is a topology; the case of $\mathcal{O}_{G,S}$ is analogous.

The empty set and G are both in $\mathcal{O}_{G,S}^l$. Suppose $U_1, U_2 \in \mathcal{O}_{G,S}^l$, $g \in U_1 \cap U_2$ and let x be in S . It follows that $gx^n \in U_1$ except for at most finitely many $n \in \mathbb{Z}$, and $gx^n \in U_2$ except for at most finitely many $n \in \mathbb{Z}$. Consequently, $gx^n \in U_1 \cap U_2$ except for at most finitely many $n \in \mathbb{Z}$. Therefore, $U_1 \cap U_2$ is in $\mathcal{O}_{G,S}^l$.

Now it will be shown that an arbitrary union of open sets is open. Let $(U_i)_{i \in I}$ be a family of elements of $\mathcal{O}_{G,S}^l$, and $U = \bigcup_{i \in I} U_i$. If $g \in U$ and $x \in S$, then $g \in U_i$ for some i , and $gx^n \in U_i$ for all $n \in \mathbb{Z}$ except at most finitely many. In particular, $gx^n \in U$ for all but finitely many n . Therefore, $\mathcal{O}_{G,S}^l$ is a topology.

It will now be shown that for any $g \in G$, the operator L_g is continuous (in the topology $\mathcal{O}_{G,S}^l$), or equivalently, for any open set U , that $g^{-1}U$ is also open. Let $g^{-1}h \in g^{-1}U$. As U is open, for any $x \in S$, all but finitely many hx^n are in U and all but finitely many $g^{-1}hx^n$ are in $g^{-1}U$. Therefore, $g^{-1}U$ is open.

Similarly, R_g is continuous on G for the topology $\mathcal{O}_{G,S}^r$.

To show connectivity, assume that U is a non-empty open and closed subset of G for the topology $\mathcal{O}_{G,S}^l$. If $g \in U$ and $x \in S$, then since U is open, $gx^n \in U$ for all but at most finitely many n . In fact, $gx^n \in U$ for all $n \in \mathbb{Z}$. To see why, note that if $gx^i \in G \setminus U$ for some i , since $G \setminus U$ is also open, $gx^n \in G \setminus U$ for all but finitely many $n \in \mathbb{Z}$, hence one can find an integer n such that $gx^n \in U \cap (G \setminus U) = \emptyset$, giving a contradiction. As this is true for any x in a generating set S , U must be all of G . It follows that G is connected for the topology $\mathcal{O}_{G,S}^l$ and similarly also for the topology $\mathcal{O}_{G,S}^r$.

For the intersection topology $\mathcal{O}_{G,S}$, G will also be connected, and for any $g \in G$, both L_g and R_g are continuous. To show that G with the topology $\mathcal{O}_{G,S}$ is a topological group, it remains to show that the inversion map $s : G \rightarrow G$ is continuous, or equivalently, that for any open $U \in \mathcal{O}_{G,S}$, $U^{-1} \in \mathcal{O}_{G,S}$.

If $U \in \mathcal{O}_{G,S}$, $g^{-1} \in U^{-1}$ and $x \in S$, then $gx^n \in U$ and $x^n g \in U$ for all but finitely many $n \in \mathbb{Z}$. It follows that $x^{-n} g^{-1} \in U^{-1}$ and $g^{-1} x^{-n} \in U^{-1}$ for all but finitely many $n \in \mathbb{Z}$, hence $U^{-1} \in \mathcal{O}_{G,S}$. Therefore, s is continuous and G is a topological group for the topology $\mathcal{O}_{G,S}$. ■

Having constructed those topologies, it remains to be seen whether they are interesting topologies. At the very least, one would want them to be non-trivial. It will be shown in Sections 3 and 4 that in the example in which $G = \text{Mod}(\Sigma)$ for some compact connected oriented surface Σ with S being the set of all Dehn twists, there are many interesting open sets. However, at this level of generality, the non-triviality of the topology $\mathcal{O}_{G,S}$ depends on the properties of the generating set S , in particular, the orders of elements in S .

Proposition 11. *Let G be a group and S a generating set for G .*

- (1) *If the generating set consists only of elements of infinite order, then $\mathcal{O}_{G,S}^l, \mathcal{O}_{G,S}^r$ and $\mathcal{O}_{G,S}$ are all finer than the cofinite topology on G .*
- (2) *In addition, if all pairs $x \neq y \in S$ satisfy $\langle x \rangle \cap \langle y \rangle = \{1\}$, then $\mathcal{O}_{G,S}^l, \mathcal{O}_{G,S}^r$ and $\mathcal{O}_{G,S}$ are strictly finer.*
- (3) *If, on the other hand, the generating set consists only of elements of finite order, $\mathcal{O}_{G,S}^l, \mathcal{O}_{G,S}^r$ and $\mathcal{O}_{G,S}$ are all the trivial topology on G .*

Proof. Both cases will only be proven for $\mathcal{O}_{G,S}^l$. This is because the proof of $\mathcal{O}_{G,S}^r$ is analogous and the case of $\mathcal{O}_{G,S}$ can be easily seen to be a consequence of the first two by taking intersections.

Assume first that S contains only elements of infinite order. It is necessary to show that for any finite subset $V \subset G$, the set $G \setminus V$ is open. Let $h \in G \setminus V$ and $x \in S$. As x has infinite order in G , the elements hx^n are all different; therefore, all but at most finitely many are in $G \setminus V$. Hence $G \setminus V$ is open, and $\mathcal{O}_{G,S}^l$ is finer than the cofinite topology on G .

Now assume also that $\langle x \rangle \cap \langle y \rangle = \{1\}$ for any two distinct elements of S . It will now be shown that for any generator $x \in S$, $G \setminus \langle x \rangle \in \mathcal{O}_{G,S}^l$.

If $g \in G \setminus \langle x \rangle$, then $gx^n \notin \langle x \rangle$ for any $n \in \mathbb{Z}$. Suppose $y \in S$ is another generator. It follows that $gy^n \in \langle x \rangle$ for at most one integer $n \in \mathbb{Z}$. If $gy^i, gy^j \in \langle x \rangle$ for $i \neq j$, then $y^{j-i} \in \langle x \rangle \cap \langle y \rangle$. But $y^{j-i} \neq 1$ as y has infinite order which contradicts the fact that $\langle x \rangle \cap \langle y \rangle = \{1\}$.

Finally, assume that S consists only of finite order elements. Let $U \in \mathcal{O}_{G,S}^l$ be a non-empty open set, and let $g \in U$ and $x \in S$. Assume that $x^d = 1$. It follows that all but at most finitely many elements gx^{1+kd} where $k \in \mathbb{Z}$ are in U . But $gx^{1+kd} = gx$, so $gx \in U$. Similarly, one shows that $gx^{-1} \in U$.

Therefore, U is closed under multiplication on the right by any element in S or any inverse of an element in S . As the set S is a generating set of G , this implies that $U = G$. Consequently, $\mathcal{O}_{G,S}^l$ is the trivial topology in this case. ■

The cyclically cofinite topologies relative to a generating set are compatible with morphisms of groups, as the following proposition shows.

Proposition 12. *Let $\phi : G \rightarrow H$ be a surjective morphism of groups and S be a generating set of G . It follows that $\phi(S)$ is a generating set for H , and for the topologies $\mathcal{O}_{G,S}^l$ and $\mathcal{O}_{H,\phi(S)}^l$ (or with $\mathcal{O}_{G,S}^r$ and $\mathcal{O}_{H,\phi(S)}^r$, or $\mathcal{O}_{G,S}$ and $\mathcal{O}_{H,\phi(S)}$), the map ϕ is continuous.*

Proof. The set $\phi(S)$ is a generating set for H as ϕ is a surjective morphism. Once again the statement will be proven for the left cyclically cofinite topologies only.

Let $U \in \mathcal{O}_{H,\phi(S)}^l$, and let $g \in \phi^{-1}(U)$. For any $x \in S$, it holds that $\phi(gx^n) = \phi(g)\phi(x)^n \in U$ for all but finitely many $n \in \mathbb{Z}$, as U is open. Therefore, $\phi^{-1}(U) \in \mathcal{O}_{G,S}^l$, and ϕ is continuous. ■

When the generating set S is closed under conjugation by an arbitrary element of G , the above topologies are better behaved.

Proposition 13. *Assume that S is a generating set for a group G that is closed under conjugation by any element of G . In this case, the topologies $\mathcal{O}_{G,S}^r, \mathcal{O}_{G,S}^l$ and $\mathcal{O}_{G,S}$ all coincide.*

Also, the intersection of any two non-empty open subsets of G is a non-empty open subset of G , that is, the topologies $\mathcal{O}_{G,S}^r, \mathcal{O}_{G,S}^l$ and $\mathcal{O}_{G,S}$ are not Hausdorff.

Proof. Let $U \in \mathcal{O}_{G,S}^l$, and let $g \in U$. It holds that for any $x \in S$, the element $g^{-1}xg \in S$ as S is closed under conjugation. Therefore, $x^n g = g(g^{-1}xg)^n \in U$ for all but at most finitely many $n \in \mathbb{Z}$. This shows that $U \in \mathcal{O}_{G,S}^r$ and $\mathcal{O}_{G,S}^l \subset \mathcal{O}_{G,S}^r$.

Similarly, $\mathcal{O}_{G,S}^r \subset \mathcal{O}_{G,S}^l$, and hence $\mathcal{O}_{G,S}^l = \mathcal{O}_{G,S}^r = \mathcal{O}_{G,S}$.

Now let U_1 and U_2 be two non-empty open sets of G . For $g \in U_1$ and $h \in U_2$, as S is a generating set, one can write

$$h = gx_1^{i_1} x_2^{i_2} \dots x_k^{i_k}$$

for some $k \geq 0$, some $x_j \in S$ and $i_j \in \mathbb{Z}$. It will be shown that one can find $g \in U_1$ and $h \in U_2$ such that k can be taken to be zero which implies that $g = h \in U_1 \cap U_2$.

Let $g \in U_1$ and $h \in U_2$. Since U_2 is open, $hx_k^n \in U_2$ for all but finitely many $n \in \mathbb{Z}$. Also, $gx_k^m \in U_1$ for all but finitely many $m \in \mathbb{Z}$. In particular, one can choose n and m such that $hx_k^n \in U_2$, $gx_k^m \in U_1$ and $m = i_k + n$. In this case,

$$hx_k^n = (gx_k^m)x_k^{-m}x_1^{i_1} \cdots x_k^{i_k+n} = (gx_k^m)x_k^{-m}x_1^{i_1} \cdots x_k^m = (gx_k^m)y_1^{i_1} \cdots y_{k-1}^{i_{k-1}},$$

where $y_j = x_k^{-m}x_jx_k^m \in S$ because S is closed under conjugation. Therefore, it is possible to inductively decrease k by 1 until $k = 0$ is obtained, showing that $U_1 \cap U_2 \neq \emptyset$. ■

Proposition 13 implies that if S is a generating set with the property of being closed under conjugation by G , then any non-empty open set in $\mathcal{O}_{G,S}$ is dense.

As in the introduction, call a subset $V \subset G$ *thick* if it has non-empty interior and *thin* if its complement is thick. Any thick set is dense because its interior is dense, and any intersection of two thick sets is thick as the intersection of their interiors is a non-empty open set.

It will sometimes be necessary to consider sets that are dense but not thick. An example will be given in Theorem 5.

Lemma 14. *Let G be a group and S be a generating set for G that is closed under conjugation. Assume in addition that V is a non-empty subset of G such that for any $g \in V$, and any $x \in S$, infinitely many gx^n are in V . Under these assumptions, V is dense in G for the topology $\mathcal{O}_{G,S}$.*

Proof. As in the proof of Proposition 13, consider $g \in U$, where U is a non-empty open set and $h \in V$ and V satisfies the hypothesis of the proposition. In this case,

$$h = gx_1^{i_1}x_2^{i_2} \cdots x_k^{i_k}$$

for some $k \geq 0$ and $x_j \in S$. Since $h \in V$, the set $T = \{m \in \mathbb{Z} \mid hx_k^m \in V\}$ is infinite. Moreover, $gx_k^n \in U$ for all but finitely many $n \in \mathbb{Z}$, and hence there is some n in $\{m - i_k \mid m \in T\}$ such that $gx_k^n \in U$. As in the previous proof, it follows that one could find another pair $g' \in U$, $h' \in V$ with $h' = g'y_1^{j_1} \cdots y_{k-1}^{j_{k-1}}$ with all y_l s in S . By induction, it follows that $U \cap V \neq \emptyset$. Hence V has non-empty intersection with any open set, that is, V is dense. ■

The twist-cofinite topology defined in Theorem 2 is an example of such a topology, where $G = \text{Mod}(\Sigma)$ for some compact connected oriented surface Σ , and S is the set of all Dehn twists. The set of all Dehn twists generates $\text{Mod}(\Sigma)$ and is closed under conjugation; if $f \in \text{Mod}(\Sigma)$ and D_γ is a Dehn twist around a curve γ on Σ , then $f \circ D_\gamma \circ f^{-1} = D_{f(\gamma)}$. Therefore, Theorem 2 and Proposition 3 are consequences of the more general Definition 9 and Proposition 13.

3. Generic properties of pseudo-Anosovs

In this section, Theorem 4 will be obtained as a corollary of classical results.

Theorem (Theorem 4 of the introduction). *The following subsets of $\text{Mod}(\Sigma)$ are non-empty and open and hence thick:*

- (1) *The set of pseudo-Anosov maps $f \in \text{Mod}(\Sigma)$.*
- (2) *The sets $\{f \in \text{Mod}(\Sigma) \text{ pseudo-Anosov} \mid \text{Vol}(M_f) > C\}$, where M_f is the mapping torus of f and $C > 0$.*
- (3) *The sets $\{f \in \text{Mod}(\Sigma) \text{ pseudo-Anosov} \mid \lambda(f) > C\}$, where $\lambda(f)$ is the stretch factor of f and $C > 0$.*
- (4) *The set of pseudo-Anosovs that does not leave any integer homology class of curves on Σ invariant.*

Proof. The first part of the theorem follows directly from the following.

Theorem 15 ([11, Theorem A, part i]). Let $f : \Sigma \rightarrow \Sigma$ be a pseudo-Anosov map and c a simple curve on Σ . It follows that for all but at most finitely many n , the composition $D_c^n f$ is also pseudo-Anosov. In [5, Theorem 0.1], this theorem was strengthened to the statement that $D_c^n f$ is pseudo-Anosov for all but at most 7 consecutive values of n .

A Dehn surgery on a 3-manifold M is a construction whereby the interior $\overset{\circ}{T}_1$ of a solid torus T_1 in the interior $\overset{\circ}{M}$ of M is removed, and another solid torus T_2 is glued in via a homeomorphism $h : \partial T_2 \rightarrow \partial T_1$. The 3-manifold $(M \setminus \overset{\circ}{T}_1) \cup_h T_2$ is said to be obtained from M by a Dehn surgery.

To prove the second part of the theorem, note that $M_{fD_c^n}$ is obtained from M_f by a Dehn surgery. The behaviour of volume under Dehn surgeries has been studied in [17], where it follows from Theorem 1A that for sufficiently large $|n|$,

$$\text{Vol}(M_f) < \text{Vol}(M_{fD_c^n}).$$

Similarly, for sufficiently large $|n|$,

$$\text{Vol}(M_f) < \text{Vol}(M_{D_c^n f}).$$

The third part of the theorem also follows directly from a theorem of [11]. It is a consequence of [11, Theorem C] that there are positive constants K , L_1 and L_2 such that

$$|n|L_1 - K \leq \lambda(D_c^n f) \leq L_2|n| + K.$$

Denote by $[c]$ the integer homology class in $H_1(\Sigma; \mathbb{Z})$ with representative c . If $[c] = 0$, the action of the mapping class f on homology is identical to that of fD_c^n or $D_c^n f$. If c is a nonseparating curve, choose a set of curves containing c that determine a symplectic basis for $H_1(\Sigma; \mathbb{Z})$, and denote by d the unique curve in the set intersecting c . Since the

action of the mapping class group on $\pi_1(\Sigma)$ defines an action on $H_1(\Sigma; \mathbb{Z})$, the action of a mapping class on $H_1(\Sigma; \mathbb{Z})$ is determined by its action on the curves making up the chosen basis. It follows immediately that there is at most one value of n for which fD_c^n acts trivially on homology. Replacing the set of curves making up the basis by their images under f^{-1} , the same argument shows that the same is true for $D_c^n f$. Therefore, the set of pseudo-Anosovs that do not leave any integer homology class of curves on Σ invariant is open. This set is known to be non-empty, so the theorem follows. ■

Since Dehn twists are not of finite order in the mapping class group, it follows immediately from Proposition 11 that the subgroups of the mapping class group in Theorem 4 are also generic in finite index subgroups of mapping class groups. In [13], it was shown that there is a sense in which pseudo-Anosovs are generic in every finitely generated subgroup of $\text{Mod}(\Sigma)$ containing a pseudo-Anosov element. Apart from this, little seems to be known about subset topologies.

4. Heegaard splittings of 3-manifolds with generic gluing maps

In this section, the homology of 3-manifolds described as Heegaard splittings with generic gluing maps will be studied. A reference for any facts about Heegaard splittings used here is [10].

Suppose a 3-manifold M is decomposed as a Heegaard splitting $N_f = H_A \cup H_B$ along a surface Σ . The homology of M can be computed from the Heegaard splitting, using Mayer–Vietoris. If g is the genus of Σ , then

$$H_2(N_f; \mathbb{Z}) \rightarrow H_1(\Sigma; \mathbb{Z}) = \mathbb{Z}^{2g} \rightarrow H_1(H_A; \mathbb{Z}) \oplus H_1(H_B; \mathbb{Z}) = \mathbb{Z}^g \oplus \mathbb{Z}^g \rightarrow H_1(N_f; \mathbb{Z}).$$

To compute $H_2(N_f; \mathbb{Z})$ and $H_1(N_f; \mathbb{Z})$, it is therefore necessary to calculate the $2g \times 2g$ matrix $K(f)$ describing the map $H_1(\Sigma; \mathbb{Z}) \rightarrow H_1(H_A; \mathbb{Z}) \oplus H_1(H_B; \mathbb{Z})$.

Let $\{x_1, \dots, x_{2g}\}$ be a generating set for $\pi_1(\Sigma)$. For H_A , let a_1, \dots, a_g be a set of curves on Σ that bound a set of disks in H_A such that when H_A is cut along these disks, a ball is obtained. The curves $\{b_1, \dots, b_g\}$ are defined similarly for H_B . Each curve a_i or b_i for $i = 1, \dots, g$ is described by a word in the generators $\{x_1, \dots, x_{2g}\}$. The first row of $K(f)$ is the vector representing $[a_1]$ in the basis $\{[x_1], \dots, [x_{2g}]\}$, and the second row is the vector representing $[a_2]$ followed by vectors representing $[a_3], \dots, [a_g], [b_1], \dots, [b_g]$ in the given order.

Lemma 16 ([10, Lemma 3.31]). *The homology $H_1(N_f; \mathbb{Z})$ is finite iff the determinant of $K(f)$ is nonzero. In addition, if $H_1(N_f; \mathbb{Z})$ is finite, then the order of $H_1(N_f; \mathbb{Z})$ is equal to the absolute value of the determinant of $K(f)$.*

Theorem 17 (Theorem 5 of the introduction). *Suppose $b_1(N_f, \mathbb{Z})$ is the first Betti number of N_f , and $|H_1(N_f; \mathbb{Z})|$ is the number of elements in the group $H_1(N_f; \mathbb{Z})$. In this case,*

- (1) the set $\{f \in \text{Mod}(\Sigma) \mid b_1(N_f, \mathbb{Z}) = 0\}$ is a thick set,
- (2) the set $\{f \in \text{Mod}(\Sigma) \mid |H_1(N_f; \mathbb{Z})| \geq k\}$ for any integer $k \geq 0$ is a thick set,
- (3) if p is prime and $0 \leq i \leq g$, the set $\{f \in \text{Mod}(\Sigma) \mid b_1(N_f, \mathbb{F}_p) = i\}$ is dense but not thick.

Proof. By Lemma 16, proving the first part of the theorem is the same as showing that the set $\{f \in \text{Mod}(\Sigma) \mid \det(A(f)) \neq 0\}$ is thick. Since this set is clearly non-empty, it suffices to show that the set is open. Consider the map from $K(f)$ to $K(fD_c^n)$. This map fixes the last g rows, and adds the vector $n\hat{i}(a_i, c)[c]$ to row i , for $1 \leq i \leq g$. Here \hat{i} denotes algebraic intersection number, and $[c]$ is assumed to be a vector representing the homology class with representative c in the basis $\{[x_1], \dots, [x_{2g}]\}$.

Since the determinant of $K(f)$ is nonzero, $[a_1], \dots, [a_g]$ and $[b_1], \dots, [b_g]$ can be chosen such that $[c]$ is in the span of $[a_g]$ and $[b_1]$. It follows that the determinant of $K(fD_c^n)$ can only be zero if the projection of $n\hat{i}(a_i, c)[c]$ onto the unit vector parallel to $[a_g]$ is given by $-[a_g]$. This can happen for at most one value of n .

Similarly, consider the map from $K(f)$ to $K(D_c^n f)$. This map fixes the first g rows, and adds the vector $n\hat{i}(b_i, c)[c]$ to row $i + g$, for $1 \leq i \leq g$. The same argument as before shows that the determinant of $K(D_c^n f)$ can only be zero for finitely many n . This proves the first part of the theorem.

From the expressions for $K(fD_c^n)$ and $K(D_c^n f)$ just given, it follows that for n^* sufficiently large, $n > n^*$ implies that $|\det(K(fD_c^n))| \geq |\det(K)|$ and $|\det(K(D_c^n f))| \geq |\det(K)|$. The second part of the theorem is then a consequence of Lemma 16.

To prove the last part of the theorem, note that $b_1(N_f, \mathbb{F}_p)$ is equal to the dimension of the kernel of $K(f) \pmod p$. Similarly, this holds for $b_1(N_f D_c^n, \mathbb{F}_p)$ and $b_1(N_{D_c^n f}, \mathbb{F}_p)$. Since $K(fD_c^{np}) = K(D_c^{np} f) = K(f) \pmod p$, the theorem follows from Lemma 14. ■

5. Hempel distance and thick sets

This final section uses the notion of Hempel distance to prove further theorems about 3-manifolds obtained as Heegaard splittings with generic gluing maps. Subsection 5.1 defines Hempel distance and shows how it relates to the study of properties of 3-manifolds describable as Heegaard splittings with generic gluing maps. Subsection 5.2 covers some background on handlebody sets. In Subsection 5.3, subsurface projections are introduced and subsurfaces to which handlebody sets have unbounded diameters in subsurface projections are studied. The final subsection uses these results to partition the handlebody sets into 3 subsets. This partition is used to give a proof of Lemma 19 and hence of Theorems 6–8.

5.1. Incompressible surfaces and Hempel distance

Hempel distance is defined in terms of Harvey’s curve complex $\mathcal{C}(\Sigma)$. This is the simplicial complex with vertices corresponding to non-trivial homotopy classes of simple

closed curves on Σ and simplices corresponding to sets of curves with pairwise geometric intersection number zero. There is a combinatorial distance defined on $\mathcal{C}(\Sigma)$, where the distance between two vertices is equal to the smallest number of edges of any path connecting the vertices. A Heegaard splitting $N_f = H_A \cup H_B$ determines two sets in $\mathcal{C}(\Sigma)$. One of the sets consists of the set of simple curves on $\Sigma = \partial H_A = \partial H_B$ that bound disks in H_A , and the other set consists of the set of simple curves on Σ that bound disks in H_B . The former set will be called A and the latter B . The Hempel distance [7] is defined to be the distance $d(A, B)$ in $\mathcal{C}(\Sigma)$ between the two sets A and B , where $d(A, B) := \inf_{a \in A, b \in B} d(a, b)$.

It is argued that the Hempel distance is an interesting measure of the complexity of N_f , for example, because reducibility properties of N_f can be elegantly formulated in terms of Hempel distance [2]. In the context of this paper, Hempel distance will be used to bound from below the genus of an incompressible surface embedded in N_f .

Lemma 18 ([6, Theorem 1.2] or consequence of [1, Theorem 3.1]). *The genus of a closed, incompressible surface F embedded in N_f is greater than or equal to half the Hempel distance.*

An important special case of Lemma 18 in [7] is that if the Hempel distance of N_f is greater than or equal to 3, N_f cannot contain an incompressible sphere or torus, from which it is then shown that N_f is hyperbolic.

This final section will be primarily devoted to proving the following lemma.

Lemma 19. *Suppose that $d(A, B) \geq 3$, and let c be a simple curve. It follows that for all but finitely many $n \in \mathbb{Z}$,*

$$d(D_c^n(A), B) \geq d(A, B).$$

This lemma implies Theorem 8. Together with Lemma 18, this proves Theorem 7.

It will now be shown that Theorem 6 also follows from Lemma 19.

Theorem 20. *Assuming Lemma 19, the sets $\{f \in \text{Mod}(\Sigma) \mid g(N_f) = g(\Sigma)\}$ and $\{f \in \text{Mod}(\Sigma) \mid g(N_f) \geq n\}$ for $n \geq 2$ are thick.*

Proof. When $g(N_f) < g(\Sigma)$, the Heegaard splitting is said to be stabilised. By [10, Lemma 5.5], this happens iff there is an embedded disk d_a in H_A and an embedded disk d_b in H_B such that the boundaries of the disks are curves on Σ in general position that intersect once only.

A necessary condition for the Heegaard splitting to be stabilised is therefore that $d(A, B) \leq 2$. By Lemma 19, the set

$$\{f \in \text{Mod}(\Sigma) \mid \text{the Heegaard splitting induced by } f \text{ has Hempel distance } \geq 3\}$$

is thick. The intersection of the set

$$\{f \in \text{Mod}(\Sigma) \mid g(N_f) = g(\Sigma)\}$$

with the thick set

$$\{f \in \text{Mod}(\Sigma) \mid \text{the Heegaard splitting induced by } f \text{ has Hempel distance } \geq 3\}$$

is the entire set

$$\{f \in \text{Mod}(\Sigma) \mid \text{the Heegaard splitting induced by } f \text{ has Hempel distance } \geq 3\}$$

and hence thick. That $\{f \in \text{Mod}(\Sigma) \mid g(N_f) = g(\Sigma)\}$ is thick then follows from Proposition 3.

For compact, orientable 3-manifolds with Heegaard splitting surfaces P and Q , in [21] it was shown that either Q is isotopic to a possibly stabilised copy of P or the Hempel distance of the splitting with splitting surface P is bounded from above by twice the genus of Q . It follows from the argument just given, Proposition 3, Lemma 19 and the first statement of this theorem that the set $\{f \in \text{Mod}(\Sigma) \mid g(N_f) \geq n, \text{ for } n \geq 2\}$ is thick. ■

5.2. Handlebody sets

Recall that H_A and H_B are handlebodies, with ∂H_A and ∂H_B both given by the surface Σ . The set of curves on ∂H_A that bound embedded disks in H_A will be denoted by A , and the set of curves on ∂H_B that bound embedded disks in H_B by B . Sets of vertices in $\mathcal{C}(\Sigma)$ that arise in this way are called handlebody sets. This subsection gives some background on handlebody sets that will be needed in the proof of Lemma 19.

Basic assumptions and notations. A curve is a non-trivial isotopy class of unpointed embeddings of S^1 into Σ . Arcs in a subsurface of Σ are isotopy classes, where isotopies are required to keep the endpoints on the boundary of the subsurface. Where convenient, a curve or arc will sometimes be confused with the image of a particular representative of the isotopy class or the corresponding vertex of $\mathcal{C}(\Sigma)$. For example, when cutting a (sub)surface along a curve or arc, it will always be assumed that “curve” or “arc” refers to the image in the surface of a particular representative of the isotopy class, such as a geodesic. A curve or arc will be said to be contained in a subsurface if there is a representative of the homotopy class with image contained in the subsurface.

The geometric intersection number of two curves, c_1 and c_2 , on Σ will be denoted by $i(c_1, c_2)$. The geometric intersection number of two arcs c_1 and c_2 will also be denoted by $i(c_1, c_2)$. Recall that $\hat{i}(c_1, c_2)$ denotes algebraic intersection number.

The most difficult case in Lemma 19 occurs when the curve c is distance 1 from A or B . The reason for this will become apparent later. From now on, unless explicitly stated otherwise, it will be assumed that the curve c in the statement of Lemma 19 is distance 1 from A . The general case, in which c could be any distance from A , will be dealt with in Lemma 19.

The set A contains an unoriented set of $3g - 3$ curves $\{a_i\}$ that determine a pants decomposition of Σ . It will be assumed that the curves $\{a_1, \dots, a_{m-1}\}$ are disjoint from c , and the curves $\{a_m, \dots, a_n\}$ are not. Since c is distance 1 from A , $m - 1 \neq n$ and the curves can be chosen such that $m > 1$.

The embedded disk bounded by a curve $a \in A$ will be denoted by \mathcal{D}_a .

Curves in a handlebody set can be conveniently described in terms of a certain decomposition, which will now be defined.

Band sums. Suppose d and e are two simple, disjoint curves on Σ and h is an embedded arc in $\Sigma \setminus (d \cup e)$ with one endpoint on d and the other on e . The band sum of d and e is the connected component of the boundary of a regular neighbourhood of $d \cup h \cup e$ that is not isotopic to either d or e . This is illustrated in Figure 1.

The band sum of two curves in A is also in A . The band sum of d and e bounds a disk obtained by taking the (necessary disjoint) disks \mathcal{D}_d and \mathcal{D}_e and attaching a long, narrow rectangle r , with one “short side” on $\partial\mathcal{D}_d$ and the other “short side” on \mathcal{D}_e . The rectangles are attached in such a way that the orientations match up to give an oriented disk in the closure of the handlebody. Push the interior of the oriented disk into the interior of H_A to obtain a disk with boundary the band sum of d and e .

Bigon surgery. A pair of disks \mathcal{D}_{d_1} and \mathcal{D}_{d_2} embedded in H_A with boundary curves d_1 and d_2 on ∂H_A can be put in general and minimal position. The intersections consist of a finite number of embedded arcs with endpoints on the crossings of d_1 and d_2 . These arcs cut each of the disks into pieces, as shown in Figure 2. Due to the fact that the sets of arcs are embedded in each of the disks, cutting each disk along the intersections must give at least two bigons for each disk.

Let x be a connected component of the intersection of \mathcal{D}_{d_1} with \mathcal{D}_{d_2} , with the property that x makes up one side of a bigon in \mathcal{D}_{d_2} . Cut \mathcal{D}_{d_1} along x , and glue in two copies of the bigon from \mathcal{D}_{d_2} , each with a different orientation. This gives two new embedded disks in H_A with boundaries in ∂H_A . These disks are each disjoint from \mathcal{D}_{d_1} and together have fewer crossings with \mathcal{D}_{d_2} than \mathcal{D}_{d_1} . These two disks are obtained from the disk \mathcal{D}_{d_1} by a bigon surgery.

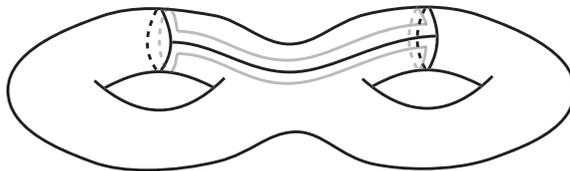


Figure 1. A band sum of two curves on Σ , shown in grey.

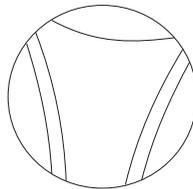


Figure 2. The intersections of the disk \mathcal{D}_{d_i} with \mathcal{D}_{d_j} .

Remark 21. Any pair of curves d and e in A has algebraic intersection number zero. This is because they bound disks that intersect along arcs in the interior of the handlebody. Each of these arcs has a pair of endpoints that are points of intersection of d and e with opposite handedness.

For sufficiently small ε , each boundary component of the ε -neighbourhood of the union of geodesics $d_1 \cup d_2$ is either homotopically trivial or a curve in A . Assuming that the disks \mathcal{D}_{d_1} and \mathcal{D}_{d_2} are in minimal position, the disk bounded by a boundary component is a boundary component of the ε -neighbourhood of $\mathcal{D}_{d_1} \cup \mathcal{D}_{d_2}$ in H_A .

5.3. Subsurface projections of handlebody sets

This subsection is dedicated to the study of subsurfaces to which projections of handlebody sets have unbounded diameter.

An embedded subsurface of Σ will be called essential if all its boundary curves are homotopically non-trivial in Σ . Subsurfaces will be assumed to be compact, connected, essential, embedded subsurfaces of Σ .

In order to facilitate inductive arguments about distances and coarse geometry of $\mathcal{C}(\Sigma)$, Masur and Minsky [15] defined the distance between two curves in a subsurface projection. Distances in subsurface projections to annuli are a special case of this definition. Distances between two curves c_1 and c_2 in the subsurface projection to a subsurface Y will be denoted by $d_Y(c_1, c_2)$. When discussing subsurface projections, in addition to the assumptions listed above, it will also be assumed that the subsurface Y to which the projection is made is not a 3-holed sphere. This is because it was shown that large subsurface projections to 3-holed spheres never arise; their arc and curve complexes have bounded diameter. This paper will not give a detailed exposition of subsurface projections, as the topological nature of the arguments only requires qualitative statements, rather than the strong numerical and/or uniform bounds for which the full force of [15] is required. Interested readers not already familiar with the concept of subsurface projections are referred to [15].

Let \mathcal{A}_c be an annulus with core curve c embedded in Σ . Distances in the subsurface projection to \mathcal{A}_c are a measure of the number of times one curve has been twisted around c relative to the other. For example, as stated in [15, equation (2.6)]

$$d_{\mathcal{A}_c}(a, D_c^n(a)) = \begin{cases} 2 + |n|, & \text{if } d(a, c) > 1 \text{ and } n \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Distances in subsurface projections are used to obtain information about geodesic paths in the curve complex. The simplest way of doing this is based on the idea that if $d_Y(a, b)$ is sufficiently large relative to $d(a, b)$, any geodesic in $\mathcal{C}(\Sigma)$ from a to b must pass through a vertex corresponding to a curve disjoint from Y . An even stronger form of this statement is given in the following theorem.

Theorem 22 ([15, Theorem 3.1]). *Suppose Y is a subsurface of Σ . Let γ be a geodesic segment, ray or infinite line in $\mathcal{C}(\Sigma)$ such that no vertex of γ is disjoint from Y . There is a constant $\beta(\Sigma)$ depending only on Σ such that any two vertices v_1 and v_2 of γ satisfy $d_Y(v_1, v_2) \leq \beta(\Sigma)$.*

The next proposition is a consequence of the previous theorem and is included here to illustrate a “hinge-like” construction that will be used repeatedly in the proof of Lemma 19. The proposition is also a consequence of [23, Theorem 1.4].

Proposition 23. *Let a , b and c be curves in Σ . It follows that for all but finitely many $n \in \mathbb{Z}$,*

$$d(D_c^n(a), b) \geq d(a, b).$$

Proof. Two basic but crucial observations are the following: Firstly, the mapping class group (in this case, specifically powers of Dehn twists around c) acts by isometry on $\mathcal{C}(\Sigma)$. Secondly, D_c fixes a vertex iff the vertex is distance at most one from c .

If $d(a, c) \leq 1$, then $D_c^n(a) = a$, so $d(D_c^n(a), b) = d(a, b)$. If $d(b, c) \leq 1$, then $D_c^n(b) = b$, so $d(a, b) = d(D_c^n(a), D_c^n(b)) = d(D_c^n(a), b)$, and the lemma is again trivially true. It is therefore only necessary to deal with the case in which c has essential intersections with both a and b .

The construction used in the rest of the proof is illustrated in Figure 3. By Theorem 22, there is an $n^* \in \mathbb{N}$ such that for $|n| \geq n^*$, every geodesic path in $\mathcal{C}(\Sigma)$ from $D_c^n(a)$ to b passes through a vertex (this vertex depends on the path) corresponding to a curve disjoint from c . When $n = n^*$, denote this vertex by c' . In Figure 3, γ_{n^*} is a geodesic from $D_c^{n^*}(a)$ to b . As shown in Figure 3, $\gamma_{n^*,1}$ is used to denote the restriction of γ_{n^*} from $D_c^{n^*}(a)$ to c' , and $\gamma_{n^*,2}$ the restriction of γ_{n^*} from c' to b . Since $D_c^{-n^*}$ acts on $\mathcal{C}(\Sigma)$ by isometry, fixing c' , a path from a to b with the same length as γ_{n^*} is obtained by concatenating $D_c^{-n^*}(\gamma_{n^*,1})$ with $\gamma_{n^*,2}$. This path cannot be shorter than a geodesic from a to b , showing that $d(D_c^{n^*}(a), b) \geq d(a, b)$.

An analogous construction works for any $n \in \mathbb{Z}$ with $|n| > n^*$, proving the proposition. ■

Remark 24. The construction in the proof of Proposition 23 implies that for all sufficiently large $|n|$, there exists a geodesic from $D_c^n(a)$ to b passing through c' . This is proved simply by contradiction as follows: Choose n such that $|n|$ is large enough to ensure that every geodesic from $D_c^n(a)$ to b passes through a vertex representing a curve disjoint from c , and suppose that no geodesic from $D_c^n(a)$ to b passes through the vertex c' . This means that the path from $D_c^n(a)$ to b obtained by concatenating $D_c^{(n-n^*)}(\gamma_{n^*,1})$ with $\gamma_{n^*,2}$ is longer than a geodesic. However, as this path has the same length as γ_{n^*} , a geodesic γ_n from $D_c^n(a)$ to b can then be used to construct a path from $D_c^{n^*}(a)$ to b shorter than the geodesic γ_{n^*} , which is the required contradiction.

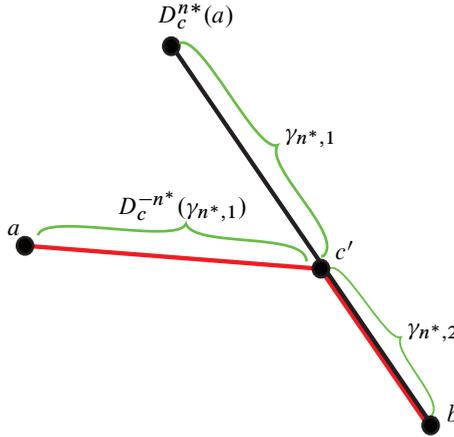


Figure 3. Path construction in the proof of Proposition 23. The path γ_{n^*} is a geodesic from $D_c^{n^*}(a)$ to b , and c' is a curve disjoint from c . The geodesic γ_{n^*} is used to construct a path from a to b as shown in red.

Large subsurface projections in handlebody sets. Subsurfaces in which handlebody sets have unbounded diameter in subsurface projections will now be characterised. A tool that will be used to do this is an I -bundle. An I -bundle is a fibre bundle with fibre given by an interval I . Suppose the I -bundle is contained in H_A and the base space Y is an orientable surface with boundary. The boundary of the I -bundle consists of a horizontal boundary and a vertical boundary. The horizontal boundary is the union of the boundaries of the fibres. It consists of two disjoint, embedded subsurfaces Y_1 and Y_2 of Σ both of which are homeomorphic to Y . The vertical boundary consists of boundary points in the interiors of fibres and is a union of open annuli, one for each closed curve on ∂Y . The I -bundle will be called essential if Y_1 and Y_2 are essential subsurfaces of Σ and each connected component of the vertical boundary is either contained in Σ or the two curves on its boundary are not homotopic in Σ .

Theorem 25 (Corollary of Theorem 1.1 of [14], as stated in [16]). *The diameter of the subsurface projection of A to a subsurface Y of Σ is bounded from above by a number $\delta(g)$ depending on the genus g of Σ , unless*

- (1) *there is an element of A in the complement of Y ,*
- (2) *there is an element of A in Y ,*
- (3) *there is an essential I -bundle I in H_A with Y a component of its horizontal boundary and at least one vertical annulus of I lying in Σ .*

The reason for the assumption that $d(A, B) \geq 3$ in Lemma 19 comes from the next lemma.

Lemma 26. *Let A , B and Σ be as defined above, where $d(A, B) \geq 3$. Let Y be a subsurface of Σ to which A has unbounded diameter in the subsurface projection. In this case, B has diameter bounded by $\delta(g)$ in the subsurface projection to Y . The same is true with A and B interchanged.*

Proof. The lemma will first be proven for subsurface projections to annuli and then for other subsurfaces.

When Y is an annulus, condition (3) of Theorem 25 does not apply, because the assumption that at least one vertical annulus of I is in Σ would prevent I from being essential. In this case, the lemma can be seen to follow directly, because if there is both a curve $a \in A$ contained in Y or $\Sigma \setminus Y$ and a curve $b \in B$ contained in Y or $\Sigma \setminus Y$, this would imply that $d(a, b) \leq 2$, contradicting the assumption that $d(A, B) \geq 3$. Similarly, if Y is not an annulus, whenever the third item of Theorem 25 is ruled out for both H_A and H_B , the lemma holds for the subsurface Y .

Now suppose Y is not an annulus, and Y is a component of the horizontal boundary of an essential I -bundle in H_A , as in condition (3) of Theorem 25. It will first be shown that Y cannot also be a component of the horizontal boundary of an essential I -bundle in H_B , as this would imply $d(A, B) \leq 2$. Denote by y_1 a boundary component of Y on the boundary of a vertical annulus of I lying in Σ . Since Y is not an annulus, there is an essential arc γ_1 contained in Y with endpoints on y_1 . A disk \mathcal{D}_a embedded in H_A is obtained by taking the union of the fibres of I that each has an endpoint on γ_1 . It follows that $\partial\mathcal{D}_a$ is a curve a in A . The curve a is disjoint in Σ from a curve γ in Y obtained as a concatenation of γ_1 with a subarc of y_1 joining the endpoints on y_1 of γ_1 . If Y is also a component of the horizontal boundary of an essential I -bundle in H_B , a curve b in B disjoint from γ will now be constructed. If y_1 is also a boundary component of a vertical annulus in Σ of an essential I -bundle in H_B , b is the boundary of the disk obtained by taking the union of the fibres of the I -bundle in H_B with an endpoint on γ_1 . Otherwise, denote by y_2 a boundary component of a vertical annulus in Σ of the I -bundle in H_B , and let γ_2 be an essential arc in Y disjoint from γ_1 with its endpoints on y_2 . The existence of γ_2 is guaranteed by the assumption that Y is an essential subsurface that is not an annulus or 3-holed sphere. The curve $b \in B$ is disjoint from γ , implying that $d(A, B) \leq 2$ as required.

It has now been shown that if B also has an unbounded diameter in the subsurface projection to Y , there must be a $b \in B$ contained in Y or $\Sigma \setminus Y$. Since $d(A, B) \geq 3$, this means that every element of A must intersect every boundary component of Y . If Y has more than one boundary component, the construction in the previous paragraph gives an element of A that does not intersect every boundary component of Y . It follows that Y can have only one boundary component. However, if Y has only one boundary component, it must have genus at least 1. The construction in the previous paragraph shows how to construct an element of A intersecting Y along an arc disjoint from any fixed curve in the interior of Y . If the curve b is in Y , this gives a contradiction to the assumption that

$d(A, B) \geq 3$. It follows that b must be in $\Sigma \setminus Y$. However, this also gives a contradiction, because there is a curve in Y disjoint from both a and b . ■

5.4. Three subsets of a handlebody set

This subsection describes how to obtain annular subsurfaces to which A has large diameter in the subsurface projection. The construction will be used to partition A into three subsets. Different techniques will then be used to prove Lemma 19 for each of these subsets.

The disjoint sum subset s_1 of A . Recall the assumption that A is distance 1 from c . Define a subset s_1 of A , the “disjoint sum subset”, consisting of all the curves in A that are disjoint from c , as well as their band sums. To see that s_1 does not have bounded diameter in the subsurface projection to \mathcal{A}_c , note that an arc in the definition of band sum of two curves disjoint from c can be chosen to wrap around c any number of times. However, for this reason, s_1 is invariant under the action of D_c .

Remark 27. Although the subset $s_1 \subset A$ is invariant under D_c , [18, Theorem 1.11] (see Lemma 29) implies that $A \setminus s_1$ is not.

The disjoint sum subset is not the only one of the 3 subsets with unbounded diameter in the subsurface projection to \mathcal{A}_c , as illustrated in the next example.

Example 28. Let c_1 be a curve in A that intersects c and d_1 be a curve in A disjoint from c and c_1 . As illustrated in Figure 4 (a), it is possible to take a band sum of c_1 with d_1 in such a way that a connected component of $c_1 \cap \mathcal{A}_c$ twists once more around c than previously. Call the resulting curve c_2 . As shown in Figure 4 (b), it is then possible to take a band sum of c_2 with a second copy of d_1 , in such a way as to increase by one the number of times a connected component of $c_2 \cap \mathcal{A}_c$ twists around c . Call the resulting curve c_3 . A sequence of curves $\{c_i\}$ is obtained by iteration.

By construction, the set $\{c_i\}$ has infinite diameter in the subsurface projection to \mathcal{A}_c , but it also has infinite diameter in the subsurface projection to another annulus. This annulus has core curve t_1 disjoint from both c and d_1 .

In Example 28, each band sum alters c_i in such a way that it is twisted around c , then backwards around a curve t_1 that cobounds an annulus with c in H_A . Outside of the annuli \mathcal{A}_c and \mathcal{A}_{t_1} , c_i is unchanged. The letter t (sometimes with a subscript) will be used to denote a simple curve in Σ , disjoint from and not homotopic to c in Σ , such that $c \cup -t$ is the boundary of an annulus in H_A . The set of all such curves in Σ will be denoted by T .

Lemma 29 ([18, Theorem 1.11]). *Suppose $\phi : H_A \rightarrow H_A$ is an automorphism with the property that $\phi|_\Sigma$ is a composition of Dehn twists around disjoint curves. Then ϕ is a composition of Dehn twists around the boundaries of a disjoint collection of properly embedded disks and annuli.*

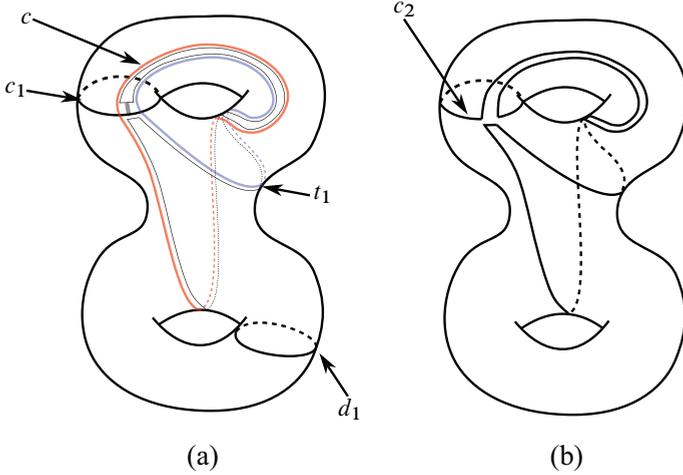


Figure 4. Part (a) shows an example of curves c_1 , d_1 , t_1 and c . The curves c is red, t_1 is blue and the thin line is d_1 . The thick grey arc is the arc along which the band sum is performed. The curve c_2 is shown in part (b).

An important special case of Lemma 29 is that when t and c are homotopically non-trivial curves in Σ that bound an embedded annulus $L(t, c)$ in the interior of H_A , the mapping class $D_t^{-1}D_c$ maps A to itself.

The c -bounded and c -unbounded subsets s_2 and s_3 of A . Suppose $\{a_1, \dots, a_{3g-3}\}$ is a choice of pairwise disjoint curves in A as stated under “Basic assumptions and notations” in Subsection 5.2. The c -bounded subset s_2 of $A \setminus s_1$ is defined to be the set of curves within radius R of the curves $Z = \{a_m, \dots, a_{3g-3}\}$ in the subsurface projection to \mathcal{A}_c , where R should be thought of as a sufficiently large natural number whose size will be estimated later. The choice of Z will only affect the size of R , not the existence of an R with the desired properties. A specific choice for Z will be given below. The c -unbounded subset s_3 is defined to be $A \setminus (s_1 \cup s_2)$.

The strategy is to show that R can be chosen large enough such that s_2 contains every curve s in $A \setminus s_1$ with the property that Theorem 22 allows the existence of a geodesic $\gamma(s, b)$ from s to a point $b \in B$ such that every vertex of $\gamma(s, b)$ intersects every $t \in T$. Suppose Z contains a curve z in $A \setminus s_1$ with the smallest possible intersection number with c . Fix $b \in B$, and choose z in such a way that

$$K_1 := \max_{Y \in W_T} d_Y(z, b) \tag{5.1}$$

is minimised, where W_T is the subset of compact, connected, essential, embedded subsurfaces of $\Sigma \setminus c$, each of which contains a curve in T . The constant K_1 is finite because it is less than the length of a hierarchy path (as defined in [15]) between two points.

The purpose of the next corollary is to understand how to construct curves in s_3 .

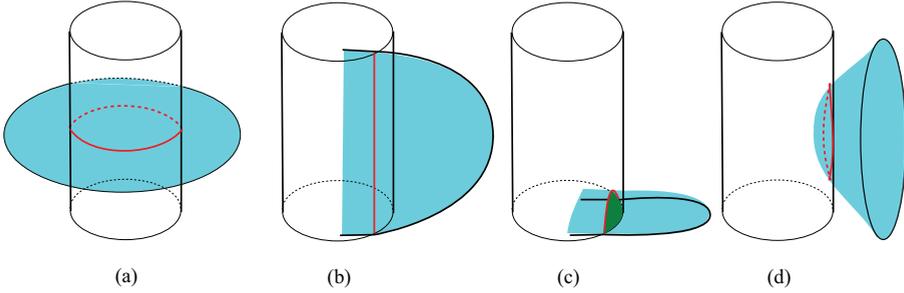


Figure 5. The different ways an annulus $L(t, c)$ embedded in H_A with boundary curves t and c in Σ can intersect a disk \mathcal{D}_a for a in A . A portion of \mathcal{D}_a is shown in blue. If $L(t, c)$ and \mathcal{D}_a are embedded and homotoped to intersect transversely, connected components of the intersection are either homotopic to the core curve of $L(t, c)$ (a), arcs with one endpoint on t and one on c (b), arcs with both endpoints on the same boundary curve of $L(t, c)$ (c) or curves bounding a disk in both $L(t, c)$ and \mathcal{D}_a (d). In part (a), the curve shown in red bounds a disk in \mathcal{D}_a . This type of intersection is possible only if $c \in A$. There are intersections of the type shown in part (b) iff $a \in A \setminus s_1$. The intersection in part (d) can be removed by a homotopy of \mathcal{D}_a .

Lemma 30. *For any $t \in T$, every curve in s_3 can be obtained as $D_c^k D_t^{-k}(a)$ for some $a \in s_2$ and some $k \in \mathbb{Z}$.*

Proof. Choose and fix $s \in s_3$ and $t \in T$. The integer $k(s)$ can then be chosen such that $d_{\mathcal{A}_c}(D_c^{-k} D_t^k(s), Z) \leq R$. By Lemma 29, $D_c^{-k} D_t^k(s) \in A$. Since $D_c^{-k} D_t^k(s)$ is not in s_3 , it remains to show that $D_c^{-k} D_t^k(s)$ is not in s_1 . This follows from Remark 27 and the following construction showing that s_1 is also invariant under Dehn twists around t .

The remainder of this proof involves understanding the different ways a disk embedded in H_A with boundary in A can intersect an annulus $L(t, c)$ embedded in H_A with boundary curves t and c in Σ . Examples of the different possibilities are illustrated in Figure 5.

A disk with boundary e in s_1 is obtained as a band sum of curves disjoint from c . Cut the bands intersecting c to obtain a set of curves $\{e_1, \dots, e_i\}$ in s_1 that are disjoint from c . This corresponds to performing bigon surgeries on \mathcal{D}_e , where an example of one such bigon is shown in green in Figure 5 (c). It will now be shown that this gives a set of curves $\{e_1, \dots, e_i\}$ in A that are all disjoint from both c and t . Suppose otherwise; then there is an $e_j \in \{e_1, \dots, e_i\}$ that has nonzero geometric intersection number with t . The intersection of the disk \mathcal{D}_{e_j} with $L(t, c)$ cannot contain a loop homotopic to a nonzero multiple of the core curve of $L(t, c)$, as shown in Figure 5 (a), because then the Jordan curve theorem in \mathcal{D}_{e_j} would imply that c and t are in A . Since e_j does not intersect c , ignoring intersections of \mathcal{D}_{e_j} with $L(t, c)$ that can be removed by a homotopy of \mathcal{D}_{e_j} , the only type of intersection between \mathcal{D}_{e_j} and $L(t, c)$ is as shown in Figure 5 (c). It follows that the intersection of \mathcal{D}_{e_j} with $L(t, c)$ determines bigon surgeries that can be performed on \mathcal{D}_{e_j} to obtain a set of disks with boundary disjoint from $L(t, c)$. Since e_j has nonzero intersection number

with t , the boundaries of these disks cannot all be homotopically trivial in Σ . This construction shows that elements of s_1 can be obtained as band sums of curves disjoint from both c and t . Consequently, s_1 is invariant under D_c and D_t , for the same reason that s_1 is invariant under D_c (see, e.g., the comment preceding Remark 27). ■

Recall that $z \in Z$ is the curve in the definition of s_2 in equation (5.1).

Lemma 31. *Let d be a curve in s_3 for which $d_{\mathcal{A}_c}(z, d) = |k|$ for $k \in \mathbb{Z}$, in particular, $|k| \geq R$. Suppose the sign of k is chosen such that $i(z, D_c^{-k}(d)) < i(z, d)$. It holds that $i(z, D_c^{-k}(d)) \geq |k| - 1$. Moreover, there is a homotopy class of arcs d^a of $d \cap (\Sigma \setminus c)$ and z^a of $z \cap (\Sigma \setminus c)$ such that d^a and z^a both intersect every curve $t \in T$.*

Proof. Since $|k| > 1$, the assumption that c is not in A implies that the homology class $[D_c^{-k}(d)] - [z]$ in $H_1(H_A; \mathbb{Z})$ is nonprimitive. The surgery construction in the proof of Lemma 30 showed that every curve in $A \setminus s_1$, in particular z , must intersect every primitive homology class of which $[D_c^{-k}(d)] - [z]$ is a multiple, that is, there exist arcs of $\mathcal{D}_z \cap L(t, c)$, as shown in Figure 5 (b). Since $[D_c^{-k}(d)] - [z]$ is at least a $|k|$ multiple of a primitive homology class, the bound $i(z, D_c^{-k}(d)) \geq |k| - 1$ follows immediately.

To prove the statement about the arcs d^a and z^a , recall from the proof of Lemma 30 that by homotoping the disk if necessary, the intersections of the disks \mathcal{D}_d and \mathcal{D}_z with $L(c, t_i)$ can all be made to be of the type shown in Figure 5 (b) or (c). This will be assumed from now on. Since d and z are not in s_1 , it was also shown that not all intersections can be of the type shown in Figure 5 (c). Then d^a and z^a can be chosen to be arcs containing one side of a bigon of the type shown in Figure 5 (b). For the curve $t \in T$ on the boundary of the annulus $L(t, c)$, it is clear that this arc intersects t . To show that such an arc intersects every curve in the set T , there are two cases to consider, as illustrated in Figure 6. In the first case, the existence of a $t' \in T$ for which d^a does not intersect t' implies that the intersections of \mathcal{D}_d with $L(c, t')$ could all be taken to be of the type shown in Figure 5 (c). This contradicts the assumption that $d \in (A \setminus s_1)$. In the second case, if d^a does not intersect t' , d has a different algebraic intersection number with t' , as with c , contradicting Remark 21 and Lemma 29. ■

Remark 32. The first part of Lemma 31 could alternatively have been proven by constructing a sequence γ of multicurves in $A \setminus s_1$ representing a path from z to d in an appropriate analogue of $\mathcal{C}(\Sigma)$. Here the i th vertex of γ is obtained from the $(i - 1)$ st vertex by a bigon surgery, using a bigon in a connected component of $\mathcal{D}_d \cap (H_A \setminus \mathcal{D}_{\gamma_i})$. As a result of the bigon surgery, $i(\gamma_i, d) \leq i(z, \gamma_{i-1}) - 2$. Lemma 31 implies that a minimum number, $|k| - 1$, of intersections between z and d resulting from the constraint $d_{\mathcal{A}_c}(z, d) = |k|$ are “paired” with at least the same number of intersections between z and $D_c^{-k}(d)$. This inspires the next definition.

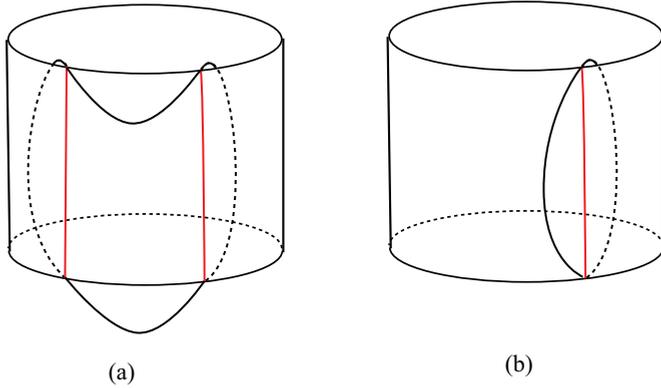


Figure 6. The annulus $L(t, c)$ in H_A has a tubular neighbourhood that can be oriented. This makes it possible to distinguish between bigons of the type shown in Figure 5 (b) that are to the right of $L(t, c)$ and to the left of $L(t, c)$. Each connected component of $d \cap (\Sigma \setminus c)$ that contains one edge of such a bigon contains an edge of two such bigons. These bigons are either both on the same side of $L(t, c)$ as shown in part (a) or both on opposite sides of $L(t, c)$ as shown in part (b).

Paired intersections. For a_1 and a_2 in A , two crossings of a_1 and a_2 will be said to be *paired* if the two crossings are the endpoints of an embedded arc in the intersection of \mathcal{D}_{a_1} with \mathcal{D}_{a_2} . Similarly, for $a_1 \in a$ and $t \in T$, two crossings of a_1 with the multicurve $t \cup c$ will be called *paired* if the crossings are the endpoints of an arc in the intersection of \mathcal{D}_{a_1} with $L(t, c)$. A crossing of a_1 with t will be said to be *paired with a crossing in \mathcal{A}_c* if it is paired with a crossing on the boundary curve c of $L(t, c)$.

Remark 33. For $d \in A \setminus s_1$, the intersection of \mathcal{D}_d with $L(t, c)$ gives at least two bigons in $\mathcal{D}_d \cap (H_A \setminus L(t, c))$. If these two bigons are on opposite sides of $L(t, c)$ as in Figure 6 (b), a curve $z \in A \setminus s_1$ with $i(c, z) = 1$ can be constructed by attaching the two bigons to form \mathcal{D}_z . If the two bigons are on the same side of $L(t, c)$, a curve $z \in A \setminus s_1$ with $i(c, z) = 2$ can be constructed by attaching the two bigons to a rectangle in H_A with a pair of opposite sides on $L(t, c)$, as shown in Figure 6 (a).

It will now be shown that, within the c -unbounded subset s_3 , moving out to infinity in the subsurface projection to \mathcal{A}_c necessarily also involves moving out to infinity in some other subsurface. This will be used to obtain lower bounds on $d(D_c^n(A), B)$ for large $|n|$. The next theorem will be used to relate the intersection number from Lemma 31 to a large distance in a subsurface projection. Some notation is needed: $c(\Sigma)$ is the set of curves on Σ , \mathcal{W} is the set of essential, connected, nonannular subsurfaces of Σ that are not 3-holed spheres and $[n]_k$ is the cutoff function defined as follows:

$$[n]_k = \begin{cases} n & \text{if } n > k, \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 34 ([22, Theorem 1.5]). *For every two vertices v_1 and v_2 of $\mathcal{C}(\Sigma)$ and all $k > 0$, we have*

$$\log_2 i(v_1, v_2) \leq V_g(k) \left(1 + \sum_{Y \subset W} [d_Y(v_1, v_2)]_k + \sum_{\mathcal{A}_l, l \subset c(\Sigma)} \log_2 [d_{\mathcal{A}_l}(v_1, v_2)]_k \right),$$

where $V_g(k) = (40000(2g - 2)(k + 200(3g - 3)))^{3g-1}$, and $\log_2(0) := 0$.

Lemma 35 ([8, Lemma 11]). *Suppose c_1 and c_2 are two curves in Σ in minimal position, for example, c_1 and c_2 are geodesics. Then the number of isotopy classes of arcs $c_1 \cap (\Sigma \setminus c_2)$ is uniformly bounded from above by $3g - 3$.*

Recall that $\beta(\Sigma)$ is the constant from Theorem 22, and z is the curve in the definition of s_2 , given in equation (5.1).

Lemma 36. *Choosing R in the definition of s_2 sufficiently large ensures that for every d in s_3 , there is a subsurface $Y(d)$ containing a curve $t(d) \in T$ for which $d_{Y(d)}(d, b) \geq \beta(\Sigma)$ for every $b \in B$. Here “sufficiently large” depends only on Σ .*

Proof. Recall that the surgery construction in the proof of Lemma 30 showed that for every $t \in T$, there is a curve in s_1 disjoint from t . The assumption that $d(A, b) \geq 3$ then implies that every $b \in B$ intersects every $t \in T$. If a subsurface of Σ contains a $t \in T$, it follows that every curve in B is either contained in this subsurface or has essential intersection with the subsurface. Choosing R in the definition of s_3 sufficiently large will now be shown to imply the existence of a $Y(d)$ containing a curve $t \in T$ with $d_Y(d, z)$ sufficiently large that the lemma then follows from Lemma 26.

Recall that K_1 is the constant from equation (5.1) and δ is the constant from Theorem 25. The lemma will first be proven for any curve $d \in s_3$ with $i(d, c) = i(z, c)$ and $\hat{i}(d, c) = \hat{i}(z, c)$, where, by Remark 33, $i(z, c) = 1$ or in the case that $i(z, c) = 2$, $\hat{i}(z, c) = 0$. In the special case that $d = D_c^k D_t^{-k}(z)$, note that choosing R larger than $\beta(\Sigma) + K_1 + \delta$ will ensure that $d_{\mathcal{A}_l}(d, b) \geq \beta$ for all $b \in B$.

Informally, the notation $I(z, d)$ will mean the “smallest” embedded, essential subsurface of $\Sigma \setminus c$ containing the arcs $z \cap (\Sigma \setminus c)$ and $d \cap (\Sigma \setminus c)$. The precise definition of $I(z, d)$ is given next. Let $\{z_1, \dots, z_k\}$ be the isotopy classes of arcs of $z \cap (\Sigma \setminus c)$.

These isotopy classes determine a set of pairwise disjoint rectangles, $\{R_1, \dots, R_k\}$ contained in $\Sigma \setminus c$, one for each isotopy class. The rectangle R_i corresponding to the i th isotopy class z_i has a pair of opposite sides on the boundary of the subsurface $\Sigma \setminus c$ and the other pair of opposite sides given by arcs in the same isotopy class as z_i . Denote by $\Sigma_{c,k}$ the subsurface of $\Sigma \setminus c$ obtained by removing all the rectangles $\{R_1, \dots, R_k\}$. In the subsurface $\Sigma_{c,k}$, next construct a set of rectangles $\{R'_1, \dots, R'_k\}$ corresponding to the isotopy classes of arcs of $d \cap \Sigma_{c,k}$, analogous to the construction of $\{R_1, \dots, R_k\}$. Glue the rectangles in the sets $\{R_1, \dots, R_k\}$ and $\{R'_1, \dots, R'_k\}$ along common boundary

points to obtain an embedded subsurface. For each boundary component of this embedded subsurface that bounds a polygon in $\Sigma \setminus c$, glue the polygon along the boundary. The resulting embedded subsurface is $I(z, d)$.

The difference $[z] - [D_c^{-k}(d)]$ is a nonprimitive homology class in $H_1(H_A; \mathbb{Z})$ given by $\lambda[t]$, where $|\lambda| \geq R$. It follows that not only does $I(z, d)$ contain a curve $t \in T$ but any disjoint union m of curves and properly embedded arcs in $I(z, d)$ that intersects every curve in T contained in $I(z, d)$ satisfies $i(z, m) + i(d, m) \geq R$.

The existence of a large subsurface projection to a subsurface $Y(d)$ of $I(z, d)$ is guaranteed by Lemmas 31, 34 and 35 whenever R is chosen sufficiently large. An algorithm for constructing a subsurface $Y(d)$ will now be derived.

If the distance between z and d in the subsurface projection to $I(z, d)$ is greater than or equal to $\beta + K_1 + \delta$, set $Y(d)$ equal to $I(z, d)$. Otherwise, a subsurface $I(z, d)_1$ will be constructed. The construction depends on whether an arc a_1 with properties given below exists.

Case 1. There exists an essential, properly embedded arc a_1 in $I(z, d)$ that has small intersection numbers with d and z . In this case, cut $I(z, d)$ along a_1 to obtain the subsurface $I(z, d)'_1$. The notion of “small” intersection numbers will be defined later; for the moment, it is enough to observe that an arc a_1 with sufficiently small intersection numbers with d and z will give a boundary component c_s of a tubular neighbourhood of $a_1 \cup \partial I(z, d)$ with $d_{I(z, d)}(d, c_s) + d_{I(z, d)}(c_s, z) \leq \beta + K_1 + \delta$, showing that $I(z, d)$ could not be $Y(d)$.

If the subsurface $I(z, d)'_1$ has any connected components consisting of n -holed spheres, replace each n -holed sphere in $I(z, d)'_1$ by the set of annuli with core curves in the n -holed sphere to which the subsurface projections of d and z have distance at least $\beta + K_1 + \delta$. Also, if any connected components of $I(z, d)'_1$ have boundary curves that define annuli to which the subsurface projections of z and d have distance at least $\beta + K_1 + \delta$, add these annuli to the set of subsurfaces. Call the resulting set of subsurfaces $I(z, d)_1$.

Case 2. When there does not exist an arc a_1 as defined above, since the distance between z and d in the subsurface projection to $I(z, d)$ is assumed to be less than $\beta + K_1 + \delta$, there exists a nonperipheral curve a_1 in $I(z, d)$ with small intersection numbers with d and z . In this case, $I(z, d)'_1$ is obtained by cutting $I(z, d)$ along a_1 . The set of subsurfaces $I(z, d)_1$ is then obtained from $I(z, d)'_1$ in the same way as in Case 1.

Either every element of $I(z, d)_1$ has the property that the distance between z and d in the subsurface projection to the embedded subsurface corresponding to that element is bounded from below by $\beta(\Sigma) + K_1 + \delta$ or the construction can be repeated on the elements not satisfying this condition to obtain $I(z, d)_2$. It is possible to iterate the construction j times to obtain $I(z, d)_j$, where, by Lemma 35, $j < 3g - 3$. At every step of

the iteration, the arc or curve $a_i, i = 1, \dots, j - 1$ is chosen such that

$$i(a_i, z) + i(a_i, d) \leq \frac{R}{6g - 6}. \quad (5.2)$$

The justification for inequality (5.2) is the following: Suppose there is an i with $1 \leq i \leq j - 1$ for which every possible choice of curve or arc a_i violates inequality (5.2). In this case, for sufficiently large R , Lemma 34 would give that the distance between z and d in the subsurface projection to every element of $I(z, d)_{i-1}$ is at least $\beta(\Sigma) + K_1 + \delta$. That being the case, set $j = i - 1$ and terminate the construction.

It remains to be shown that a subsurface in $I(z, d)_j$ contains a curve in T . When the choices a_1, \dots, a_{i-1} are all made in accordance with inequality (5.2), the curves and arcs $\{a_1, \dots, a_{i-1}\}$ cannot have large enough intersection numbers in total with d and z to cut through every curve in T contained in $I(z, d)$. It is therefore possible to find an element of $I(z, d)_j$ containing a curve in T . This subsurface is then called $Y(d)$. This concludes the proof of the lemma in the case that $i(d, c) = i(z, c)$.

To prove the lemma when $i(d, c) > i(z, c)$, note that the arcs z^a and d^a guaranteed by Lemma 31 are contained in the subsurface $I(z, d)_1$ and must have intersection number at least $R - 1$ in this subsurface. The rest of the argument is then identical to the proof in the case that $i(d, c) = i(z, c)$. ■

It is now possible to put all the pieces together to give a proof of Lemma 19.

Lemma 37 (Lemma 19 from Subsection 5.1). *Suppose that $d(A, B) \geq 3$, and let c be a curve, which is no longer assumed to be distance 1 from A . It follows that for all but finitely many $n \in \mathbb{Z}$,*

$$d(D_c^n(A), B) \geq d(A, B).$$

Proof. If A and B could each be shown to have finite diameter in the subsurface projection to \mathcal{A}_c , the theorem would follow from the same argument as for Proposition 23. By Theorem 25, this is the case, for example, when $d(c, A) \geq 2$ and $d(c, B) \geq 2$. This special case of the lemma was also proven in [23, Theorem 1.4]. Alternatively, if D_c takes A or B to itself, for example, $c \in B$ or $c \in A$, the lemma also follows. It remains to prove the lemma under the assumption that $d(c, A) = 1, d(c, B) \geq 2$ and $d(A, B) \geq 3$; the argument in the case that $d(c, B) = 1, d(c, A) \geq 2$ is symmetric.

The remaining case of the lemma will be proven by showing the following three statements. The first statement is that $d(D_c^n(s_1), B) = d(s_1, B)$ for every $n \in \mathbb{Z}$. The second statement is that for all but finitely many $n \in \mathbb{Z}$, $d(D_c^n(s_2), B) \geq d(s_2, B)$. The third statement is that for all but finitely many $n \in \mathbb{Z}$, $d(D_c^n(s_3), B) \geq d(s_3, B)$.

The first statement is a consequence of the first part of Remark 27. By Lemma 26, the second statement follows from an identical argument to that given in the proof of Proposition 23, because both s_2 and B have bounded diameters in the subsurface projection to \mathcal{A}_c . The third statement follows from a double iteration of the construction in the proof of Proposition 23, as will now be explained.

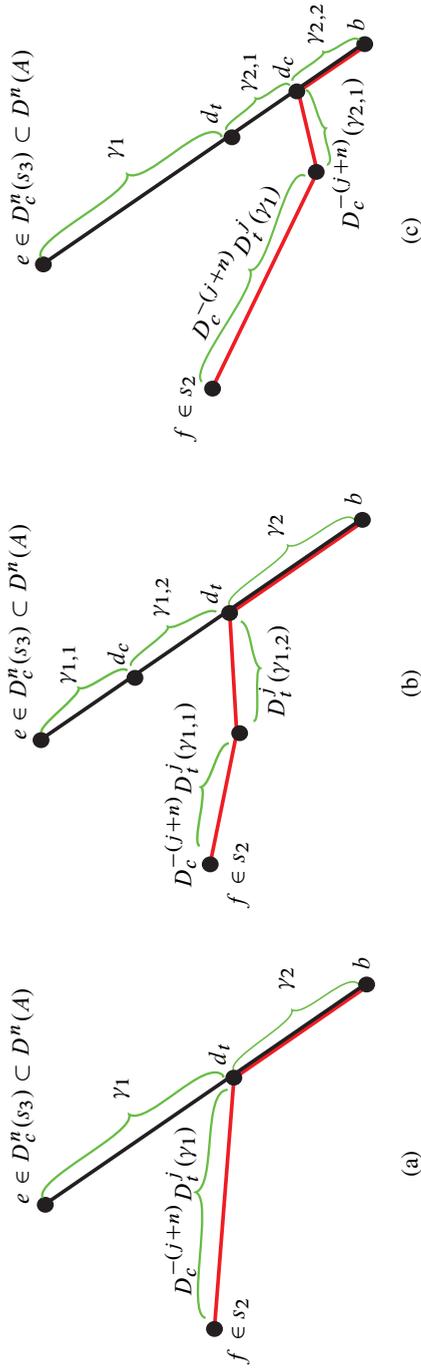


Figure 7. For sufficiently large n , a geodesic γ from $e \in D_c^n(s_3)$ to $b \in B$ determines a path of the same length from $f = D_c^{-(j+n)} D_t^j(e) \in s_2$ to $b \in B$. Part (a) shows the construction of the path when the vertex d_t is disjoint from c as well as t , part (b) shows the construction when there is a vertex d_c disjoint from c in the segment γ_1 of γ and part (c) shows the construction when there is a vertex d_c disjoint from c in the segment γ_2 of γ .

Denote by e a curve in $D_c^n(s_3)$, where $|n|$ is large enough to ensure that $d_{\mathcal{A}_c}(e, B) > \beta$. Theorem 22 together with Lemma 36 ensures that a geodesic γ from e to $b \in B$ must pass through a vertex d_t disjoint from a curve t in T and a vertex d_c disjoint from c . Choose $j \in \mathbb{Z}$ such that $f = D_c^{-j} D_t^j D_c^{-n}(e) \in s_2$ as guaranteed by Lemma 30. In the case that d_t is also disjoint from c , a path from f to b of length equal to the length of γ is constructed as in Figure 7 (a). When d_t intersects c , the vertex d_t will not be fixed by $D_c^{-(j+n)} D_t^j$. In this case, either the subpath γ_1 or γ_2 contains the vertex d_c . In the first case, a path from f to b with the same length as γ is constructed as in Figure 7 (b), and in the second case, a path from f to b with the same length as γ is constructed as in Figure 7 (c). This concludes the proof of the third statement and hence of the proof. ■

Acknowledgements. As noted above, the definition of the twist-cofinite topology comes from Renaud Detcherry, and the first two subsections of this paper were also based on R. Detcherry (unpublished personal communication), with the author’s permission. The author would also like to thank Saul Schleimer for helpful discussions and an anonymous referee for pointing out a number of useful references and devoting considerable time to writing a helpful commentary that led to improved clarity of exposition.

References

- [1] D. Bachman and S. Schleimer, [Surface bundles versus Heegaard splittings](#). *Comm. Anal. Geom.* **13** (2005), no. 5, 903–928 Zbl 1138.57026 MR 2216145
- [2] A. J. Casson and C. M. Gordon, [Reducing Heegaard splittings](#). *Topology Appl.* **27** (1987), no. 3, 275–283 Zbl 0632.57010 MR 0918537
- [3] N. M. Dunfield and W. P. Thurston, [Finite covers of random 3-manifolds](#). *Invent. Math.* **166** (2006), no. 3, 457–521 Zbl 1111.57013 MR 2257389
- [4] V. Erlandsson, J. Souto, and J. Tao, [Genericity of pseudo-Anosov mapping classes, when seen as mapping classes](#). *Enseign. Math.* **66** (2020), no. 3/4, 419–439 Zbl 1481.37044 MR 4254267
- [5] A. Fathi, [Dehn twists and pseudo-Anosov diffeomorphisms](#). *Invent. Math.* **87** (1987), no. 1, 129–151 Zbl 0618.58027 MR 0862715
- [6] K. Hartshorn, [Heegaard splittings of Haken manifolds have bounded distance](#). *Pacific J. Math.* **204** (2002), no. 1, 61–75 Zbl 1065.57021 MR 1905192
- [7] J. Hempel, [3-manifolds as viewed from the curve complex](#). *Topology* **40** (2001), no. 3, 631–657 Zbl 0985.57014 MR 1838999
- [8] I. Irmer, [Geometry of the homology curve complex](#). *J. Topol. Anal.* **4** (2012), no. 3, 335–359 Zbl 1270.57057 MR 2982447
- [9] N. V. Ivanov, [Fifteen problems about the mapping class groups](#). In B. Farb (ed.), *Problems on mapping class groups and related topics*, pp. 71–80, Proc. Sympos. Pure Math. 74, American Mathematical Society, Providence, RI, 2006 Zbl 1281.57011 MR 2264532
- [10] J. Johnson, [Notes on Heegaard splittings](#). <https://www.scribd.com/document/75183305/Jesse-Johnson-Notes-on-Heegaard-Splittings> visited on 11 July 2025
- [11] D. D. Long and H. R. Morton, [Hyperbolic 3-manifolds and surface automorphisms](#). *Topology* **25** (1986), no. 4, 575–583 Zbl 0607.57007 MR 0862441

- [12] A. Lubotzky, J. Maher, and C. Wu, [Random methods in 3-manifold theory](#) (in Russian). *Tr. Mat. Inst. Steklova* **292** (2016), 124–148. English translation: *Proc. Steklov Inst. Math.* **292** (2016), no. 1, 118–142 Zbl [1354.57026](#) MR [3628457](#)
- [13] J. Maher, [Random walks on the mapping class group](#). *Duke Math. J.* **156** (2011), no. 3, 429–468 Zbl [1213.37072](#) MR [2772067](#)
- [14] H. Masur and S. Schleimer, [The geometry of the disk complex](#). *J. Amer. Math. Soc.* **26** (2013), no. 1, 1–62 Zbl [1272.57015](#) MR [2983005](#)
- [15] H. A. Masur and Y. N. Minsky, [Geometry of the complex of curves II: hierarchical structure](#). *Geom. Funct. Anal.* **10** (2000), no. 4, 902–974 Zbl [0972.32011](#) MR [1791145](#)
- [16] Y. N. Minsky, [Curve complexes, surfaces and 3-manifolds](#). In M. Sanz-Solé, J. Soria, J. L. Varona, and J. Verdera (eds.), *Proceedings of the international congress of mathematicians*, pp. 1001–1033, Invited lectures II, European Mathematical Society, Zürich, 2006 Zbl [1099.57012](#) MR [2275633](#)
- [17] W. D. Neumann and D. Zagier, [Volumes of hyperbolic three-manifolds](#). *Topology* **24** (1985), no. 3, 307–332 Zbl [0589.57015](#) MR [0815482](#)
- [18] U. Oertel, [Automorphisms of three-dimensional handlebodies](#). *Topology* **41** (2002), no. 2, 363–410 Zbl [0991.57017](#) MR [1876895](#)
- [19] I. Rivin, [Generic phenomena in groups: some answers and many questions](#). In E. Breuillard and H. Oh (eds.), *Thin groups and superstrong approximation*, pp. 299–324, Math. Sci. Res. Inst. Publ. 61, Cambridge University Press, Cambridge, 2014 Zbl [1341.20035](#) MR [3220895](#)
- [20] I. Rivin, [Statistics of Random 3-Manifolds occasionally fibered over the circle](#). 2014, arXiv:[1401.5736v4](#)
- [21] M. Scharlemann and M. Tomova, [Alternate Heegaard genus bounds distance](#). *Geom. Topol.* **10** (2006), no. 1, 593–617 Zbl [1128.57022](#) MR [2224466](#)
- [22] Y. Watanabe, [Intersection numbers in the curve complex via subsurface projections](#). *J. Topol. Anal.* **9** (2017), no. 3, 419–439 Zbl [1366.05037](#) MR [3661650](#)
- [23] M. Yoshizawa, [High distance Heegaard splittings via Dehn twists](#). *Algebr. Geom. Topol.* **14** (2014), no. 2, 979–1004 Zbl [1295.57025](#) MR [3180825](#)

Received 17 October 2022.

Ingrid Irmer

International Center for Mathematics, Southern University of Science and Technology, 1088 Xueyuan Av, 51088 Shenzhen, P. R. China; ingridmary@sustech.edu.cn