msp

Geometry & Topology 19 (2015) 2407–2464

Factorizations of diffeomorphisms of compact surfaces with boundary

ANDY WAND

We study diffeomorphisms of compact, oriented surfaces, developing methods of distinguishing those which have positive factorizations into Dehn twists from those which satisfy the weaker condition of being right-veering. We use these to construct open book decompositions of Stein-fillable 3–manifolds whose monodromies have no positive factorization.

57R17

1 Introduction

Let Σ be a compact, orientable surface with nonempty boundary. The (restricted) mapping class group of Σ, denoted $\mathrm{MCG}(\Sigma)$, is the group of isotopy classes of orientation preserving diffeomorphisms of Σ which restrict to the identity on $\partial\Sigma$. The goal of this paper is to study the monoid $\mathrm{Dehn}^+(\Sigma) \subset \mathrm{MCG}(\Sigma)$ of products of right Dehn twists, and apply this study to the question of the extent to which properties of a contact three-manifold are captured by the data associated to an arbitrary supporting (in the sense of Giroux) open book decomposition.

As was shown by Thurston in his proof of left-orderability of the braid group, and later by Goodman [5] and Honda, Kazez and Matic [6] in the context of contact topology, a neighborhood of the boundary of a properly embedded arc, in particular the information of whether the arc is mapped "to the right or left" at that endpoint (Example 2.1), carries important information about a given diffeomorphism. Indeed, as was shown in [6], existence of an arc in a surface Σ which is mapped to the left at either endpoint by $\varphi \in \mathrm{MCG}(\Sigma)$ is sufficient to conclude geometric flexibility, or *overtwistedness*, of the contact structure associated to φ by the Giroux correspondence theorem.

The main new idea of this paper is to provide a framework for generalization of these methods. The main tool introduced is that of a *right position* (Definition 3.1) for the images of a collection Γ of disjoint properly embedded arcs in Σ under a given diffeomorphism φ. Loosely speaking, this should be thought of as a collection of points of Γ which are "fixable" under φ (Definition 3.8), or, for the case that φ is

Published: 20 October 2015 DOI: 10.2140/gt.2015.19.2407

given as a product of Dehn twists, fixable under each successive twist. A right position thus gives a decomposition of $\Gamma \cap \varphi(\Gamma)$ into subarcs with fixed ends, analogous to the collection Γ with which we started.

We then show how comparing such subarcs, through means of rectangular regions they cut out of the surface, allows one to push the notions of "rightness" and "leftness" further into the interior of the surface. Our particular interest is in detecting whether a given mapping class lies in $\mathrm{Dehn}^+(\Sigma)$, as this implies a particular rigidity, *Stein fillability*, of the associated contact structure. We develop consistency conditions (Definition 3.6) which allow us to prove:

Theorem 1.1 *Let Σ be a compact surface with boundary, and $\varphi \in \mathrm{Dehn}^+(\Sigma)$. Then for any collection of disjoint, properly embedded arcs, there exists a consistent right position.*

We use Theorem 1.1 to develop necessary conditions for elements of the set of curves which can appear as twists in some positive factorization of a given $\varphi \in \mathrm{MCG}(\Sigma)$:

Theorem 1.2 *Let γ_1 and γ_2 be disjoint properly embedded arcs in a surface Σ, $\varphi \in \mathrm{MCG}(\Sigma)$, and $\varphi(\gamma_1)$ and $\varphi(\gamma_2)$ flat with respect to some D. Then τ_α is a Dehn twist in some positive factorization of φ only if α is nested with respect to D.*

(The non-standard terminology of Theorem 1.2 will be defined in Section 4.)

Finally, as an application of the above methods, we have:

Theorem 1.3 *There exist open book decompositions which support Stein-fillable contact structures but whose monodromies cannot be factorized into positive Dehn twists.*

Central to much of current research in contact geometry is the relation between the monodromy of an open book decomposition and geometric properties of the contact structure, such as overtwistedness and fillability. The starting point is the remarkable theorem of Giroux [4], demonstrating a one-to-one correspondence between oriented contact structures (up to isotopy) on a 3–manifold M and open book decompositions of M up to positive stabilization. It has been shown by Giroux [4], Loi and Piergallini [8], and Akbulut and Ozbagci [1], that any open book with monodromy which can be factorized into positive Dehn twists supports a Stein-fillable contact structure, and that every Stein-fillable contact structure is supported by *some* open book with monodromy which can be factored into positive twists. The question of whether *each* open book

which supports a Stein-fillable contact structure must be positive is then answered in the negative by our Theorem 1.3.

Section 2 introduces some conventions and definitions, and provides motivation for what is to come. Section 3 is devoted to the terminology and proof of Theorem 1.1. We also construct simple examples of open books whose monodromies, though right veering 2.2, have no positive factorization.

In Section 4 we use these results to address the question of under what conditions various isotopy classes of curves on a surface can be shown not to appear as Dehn twists in any positive factorization of a given monodromy. We obtain various necessary conditions under certain assumptions on the monodromy, culminating in Theorem 1.2.

Finally, in Section 5, we construct explicit examples of open book decompositions for an infinite family of 3–manifolds which support Stein-fillable contact structures yet whose monodromies have no positive factorizations. For the construction we demonstrate a method of modifying a certain mapping class group relation (the lantern relation) into an "immersed" configuration, which we then use to modify certain open books with positively factored monodromies (which therefore support Stein-fillable contact manifolds) into stabilization-equivalent open books (which support the same contact structures) whose monodromies now have non-trivial negative twisting. We then apply the methods developed in the previous sections to show that in fact this negative twisting is essential.

After a preliminary announcement of these results Baker, Etnyre, and Van Horn-Morris [2] were able to construct similar examples of non-positive open books supporting Stein-fillable contact structures, all of which use the same surface, $\Sigma_{2,1}$, involved in our construction. Their method for demonstrating non-positivity is on the one hand substantially shorter, but on the other is quite restrictive, being entirely specific to this surface, and also requiring complete knowledge of the Stein fillings of the contact manifold in question, which is known only for a small class of contact manifolds.

Acknowledgements We would like to thank Danny Calegari for many helpful comments on a previous version of this paper, Rob Kirby for his support and encouragement throughout the much extended period over which it was completed, and the referees for many helpful suggestions and comments. We would also like to acknowledge the support and hospitality of the Max Planck Institute for Mathematics.

2 Preliminaries

Throughout, Σ denotes a compact, orientable surface with nonempty boundary. The (restricted) mapping class group of Σ, denoted $\mathrm{MCG}(\Sigma)$, is the group of isotopy

classes of orientation preserving diffeomorphisms of Σ which restrict to the identity on $\partial\Sigma$. In general we will not distinguish between a diffeomorphism and its isotopy class.

Let $\mathrm{SCC}(\Sigma)$ be the set of simple closed curves on Σ. Given $\alpha \in \mathrm{SCC}(\Sigma)$ we may define a self-diffeomorphism D_α of Σ which is supported near α as follows. Let a neighborhood N of α be identified by oriented coordinate charts with the annulus $\{a \in \mathbb{C} \mid 1 \le \|a\| \le 2\}$. Then D_α is the map $a \mapsto e^{-i2\pi(\|a\|-1)}a$ on N, and the identity on $\Sigma \setminus N$. We call D_α the *positive Dehn twist* about α. The inverse operation, D_α^{-1}, is a *negative* Dehn twist. We denote the isotopy class of D_α by τ_α. It can easily be seen that τ_α depends only on the isotopy class of α.

We call $\varphi \in \mathrm{MCG}(\Sigma)$ *positive* if it can be factored as a product of positive Dehn twists, and denote the monoid of such mapping classes as $\mathrm{Dehn}^+(\Sigma)$.

An open book decomposition (Σ, φ), where $\varphi \in \mathrm{MCG}(\Sigma)$, for a 3–manifold M with binding K, is a homeomorphism between $((\Sigma \times [0,1])/\sim_\varphi, (\partial\Sigma \times [0,1])/\sim_\varphi)$ and (M, K). The equivalence relation $\sim_\varphi$ is generated by $(x, 0) \sim_\varphi (\varphi(x), 1)$ for $x \in \Sigma$ and $(y, t) \sim_\varphi (y, t')$ for $y \in \partial\Sigma$.

A recurring theme of this paper is to use properly embedded arcs in Σ to understand restrictions on possible positive factorizations of φ which can be derived from the images of the arcs (relative to the arcs themselves). Our general method is as follows. Let P be some property of pairs $(\varphi(\gamma), \gamma)$, which we abbreviate by referring to P as a property of the image $\varphi(\gamma)$. Suppose that P holds for the case that φ is the identity, and is preserved by positive Dehn twists. Suppose then that τ_α appears in some positive factorization of φ. Conjugating (ie using so-called *Hurwitz moves*), we can assume τ_α is the *final* twist, and so the monodromy given by $\tau_\alpha^{-1} \circ \varphi$ is also positive. It follows that P holds for $(\tau_\alpha^{-1} \circ \varphi)(\gamma)$ as well.

Example 2.1 As a motivating example, consider a pair of arcs $\gamma, \gamma' \colon [0,1] \hookrightarrow \Sigma$ which share an endpoint $\gamma(0) = \gamma'(0) = x \in \partial\Sigma$, isotoped to minimize intersection. Following [6], we say γ' is "to the right" of γ, denoted $\gamma' \ge \gamma$, if either the pair is isotopic, or if the tangent vectors $(\dot\gamma'(0), \dot\gamma(0))$ define the orientation of Σ at x. The property of being "to the right" of γ (at x) is then a property of the images $\varphi(\gamma)$ which satisfies the conditions of the previous paragraph. We conclude that if $\varphi(\gamma)$ is to the right of γ, then α can appear in a positive factorization of φ only if $(\tau_\alpha^{-1} \circ \varphi)(\gamma)$ is to the right of γ.

The main motivation for the use of properly embedded arcs to characterize properties of mapping classes is the well-known result that a mapping class of a surface with boundary is completely determined by the images of a set of properly embedded disjoint

arcs whose complement is a disc (an application of the "Alexander method", see eg [3]). Perhaps the most well-known example of such a property is what is now generally known as *right veering*:

Definition 2.2 [6] Let φ be a mapping class in $\mathrm{MCG}(\Sigma)$, and $\gamma \hookrightarrow \Sigma$ a properly embedded arc with endpoint $x \in \partial\Sigma$. Then φ is *right-veering* if for each such γ and x, the image $\varphi(\gamma)$ is to the right of γ at x.

We denote the set of isotopy classes of right-veering diffeomorphisms as $\mathrm{Veer}(\Sigma) \subset \mathrm{MCG}(\Sigma)$.

The right-veering property can perhaps best be thought of as a necessary condition of positivity, and (by [6]), using the Giroux correspondence, also a necessary condition for tightness of the supported contact structure. It was however shown in that paper that right veering is far from a sufficient condition, indeed, that *any* open book decomposition may be stabilized to a right-veering one.

The failure of right veering to capture either tightness or positivity is at least partially due to two obvious shortcomings. Firstly, we observe that right veering "only sees one arc at a time". Referring back to the original motivation of the Alexander method, we would rather have a property which can see arbitrary collections of (disjoint) arcs. Secondly, right veering is "localized" to the boundary; in particular negative twisting in the interior of a surface may be hidden by positivity nearer to the boundary. The goal of Section 3 is to introduce a substantial refinement of right veering which takes into account each of these issues.

3 Right position

This section is devoted to the proof of Theorem 1.1. We will begin with an introduction to the rather non-standard terminology, and then turn to the various ideas involved in the proof.

3.1 Definitions and examples

Let Σ be a closed surface with non-empty boundary. Throughout the paper, an *arc* in Σ will refer to a properly embedded arc. We will refer to a set Γ of disjoint arcs as an *arc collection*. A central object of study will be *augmented* open book decompositions $(\Sigma, \varphi, \Gamma)$. Given such a triple, and $\gamma \in \Gamma$, a point p in $\gamma \cap \varphi(\gamma)$ may be given a sign as follows: orienting γ arbitrarily, and giving $\varphi(\gamma)$ the opposite orientation, p is *positive* if the pair of tangent vectors at p along γ and $\varphi(\gamma)$ in that order define the

orientation of Σ at p, *negative* otherwise (Figure 1). The resulting sign is of course independent of the choice of orientation.

Our main definition is the following.

Definition 3.1 Suppose $(\Sigma, \varphi, \Gamma)$ is an augmented open book decomposition and $\varphi \in \mathrm{Veer}(\Sigma)$, with $\varphi(\Gamma)$ isotoped to minimize $\Gamma \cap \varphi(\Gamma)$. Let A be a subset of the positively oriented interior intersections in $\bigcup_{\gamma\in\Gamma}(\gamma \cap \varphi(\gamma))$. Then the set $\partial\Gamma \cup A$ is a *right position* $\mathcal{P} = \mathcal{P}(\Sigma, \varphi, \Gamma)$ for the augmented open book.

Note that we are thinking of the set I of positively oriented interior intersections in $\bigcup_{\gamma\in\Gamma}(\gamma \cap \varphi(\gamma))$ as depending only on the data of Γ and the isotopy class φ, and thus independent of isotopy of Γ and $\varphi(\Gamma)$ as long as intersection minimality holds. In particular, the integer $|I|$ clearly depends only on this data. Labeling $\Gamma = \{\gamma_i\}_{i=1}^n$ we may label elements of I with indices i and j, where j increases along γ_i from the distinguished endpoint c_i. Right positions for the triple $(\Sigma, \varphi, \Gamma)$ are thus in 1-1 correspondence with subsets of the set of these pairs of indices.

Figure 1: p_1 is a positive, p_2 a negative, point of $\gamma \cap \varphi(\gamma)$. This figure introduces the conventions, to be followed throughout the paper, that elements of a given arc collection Γ are drawn as straight lines, their images under some diffeomorphism curved, while in any figure with multiple line weights, the thickest is reserved for the boundary of Σ.

Associated to a right position is the set $H(\mathcal{P}) := \{[v, v'] \subset \varphi(\gamma) \mid v, v' \in \mathcal{P}, \gamma \in \Gamma\}$. We denote these segments by $h_{v,v'}$, or, if only a single endpoint is required, simply use h_v to denote a subarc starting from v and extending as long as the context requires. In this case the direction of extension along $\varphi(\gamma)$ will be clear from context.

A right position can thus be thought of as a way of decomposing $\varphi(\Gamma)$ into "horizontal" segments $H(\mathcal{P})$, separated by the points $\mathcal{P}$ (Figure 2). Note that for any right-veering φ, and any arc collection Γ, $(\Sigma, \varphi, \Gamma)$ has a trivial right position consisting of the points $\partial(\Gamma)$, and so there is a single horizontal segment $h_{c,c'} = \varphi(\gamma)$ for each $\gamma \in \Gamma$. Of course if I is nonempty, there are $2^{|I|} - 1$ non-trivial right positions. Intuitively, the points in $\mathcal{P}$ play a role analogous to the boundary points of the arcs, allowing

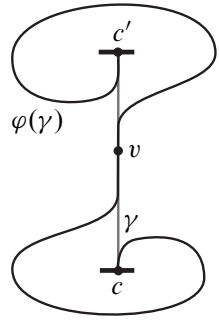


Figure 2: $\mathcal{P} = \{c, v, c'\}$ is a right position of $(\Sigma, \varphi, \{\gamma\})$. There are three distinct horizontal segments: $h_{c,v}, h_{v,c'}$, and $h_{c,c'}$.

us to localize the global "right"-ness of an arc image though the decomposition into horizontal segments.

We are interested in using a right position to compare *pairs* of arcs, the idea being that if φ is positive, we expect to be able to view any pair of horizontal segments as either unrelated, or as belonging to the "same" horizontal segment.

To get started, we need to be able to compare horizontal segments. To that end we have:

Definition 3.2 Let $(\Sigma, \varphi, \Gamma)$ be an augmented open book decomposition. A *rectangular region* of $(\Sigma, \varphi, \Gamma)$ is the image of an immersion of the standard disc $[0,1] \times [0,1]$ such that the edges are embedded alternately in Γ and $\varphi(\Gamma)$ and vertices map to intersections $\Gamma \cap \varphi(\Gamma)$.

Definition 3.3 Let $(\Sigma, \varphi, \Gamma)$ be an augmented open book decomposition, and $\mathcal{P} = \mathcal{P}(\Sigma, \varphi, \Gamma)$ a right position. We let $\mathcal{R}(\mathcal{P})$ denote the set of rectangular regions of $(\Sigma, \varphi, \Gamma)$ such that for each region, (1) there exist orientations of Γ which induce an orientation of the boundary of the region, and (2) exactly two vertices lie in $\mathcal{P}$ (see Figure 3). We refer to this collection as the *initially parallel* regions associated to $\mathcal{P}$.

Given an initially parallel region, we refer to the pair of vertices in $\mathcal{P}$ as $\bullet$*–points*, $\circ$*–points* otherwise (so $\circ$–points and $\bullet$–points alternate along the boundary of a region). Throughout the paper, we will draw $\bullet$–points as solid dots, $\circ$–points as circles.

Definition 3.4 Let $B_1, B_2 \in \mathcal{R}(\mathcal{P})$ (as in the previous definition), and orient Γ so that the induced orientation of $\partial(B_1)$ agrees with the standard (counterclockwise)

orientation of the boundary of a region in an oriented surface. If these orientations then induce the standard orientation of $\partial(B_2)$, we say that the pair have the same orientation, while if the orientations induce the opposite orientation on $\partial(B_2)$, we say the regions have the opposite orientations.

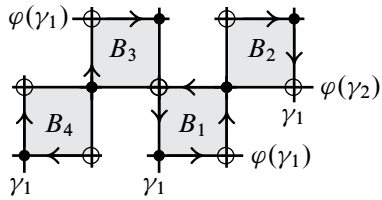


Figure 3: Various initially parallel regions in a collection $\mathcal{R}(\mathcal{P})$, where we have given the arcs the orientations which induce the standard orientation on B_1. Region B_1 has the opposite orientation of B_2 and B_3 (so B_2 and B_3 have the same orientation), while the given orientations do not orient B_4.

Remark 3.5 It is helpful to think of a rectangular region as the image under the covering map $\rho\colon \widetilde{\Sigma} \to \Sigma$ of an *embedded* disc in the universal cover. As such, we will draw regions as embedded discs whenever doing so does not result in any loss of essential information.

Definition 3.6 Let $\mathcal{P}$ be a right position for an augmented open book $(\Sigma, \varphi, \Gamma)$. We say that $\mathcal{P}$ is *consistent* if for each $B \in \mathcal{R}(\mathcal{P})$ there is a $B' \in \mathcal{R}(\mathcal{P})$ with the same $\circ$–points, and opposite orientation, as B. We refer to B' as a *completing region* for B.

Example 3.7 The right position indicated in Figure 4(a) is consistent, with two completed pairs of regions, as indicated in (b). It is straightforward to verify that there are no other regions in $\mathcal{R}(\mathcal{P})$.

The configuration in Figure 5, however, admits no consistent right position: supposing otherwise, then h_{c_1} and h_{c_2} are edges of an initially parallel region B, so there must be $v \in \mathcal{P}|_{\gamma_1}$ and $w \in \mathcal{P}|_{\gamma_2}$ satisfying the compatibility conditions. The only candidates are c_1' and c_2', and these do not give a completing region. By Theorem 1.1, this is sufficient to conclude that the mapping class has no positive factorization.

Finally, consider the open book decomposition indicated in Figure 6, where the monodromy φ is $\tau_\alpha^{-2}\tau_{\beta_1}^{n_1}\tau_{\beta_2}^{n_2}\tau_{\beta_3}\tau_{\beta_4}$ (we have drawn the case $n_1 = 1$). Using the shaded initially parallel region, observe that $(\Sigma, \varphi, \{\gamma_1, \gamma_2\})$ admits no consistent right position: indeed, for any $n_1, n_2 > 0$ this will clearly remain true, as the picture will only be modified by twists about the boundary component parallel to β_2. In particular, φ has

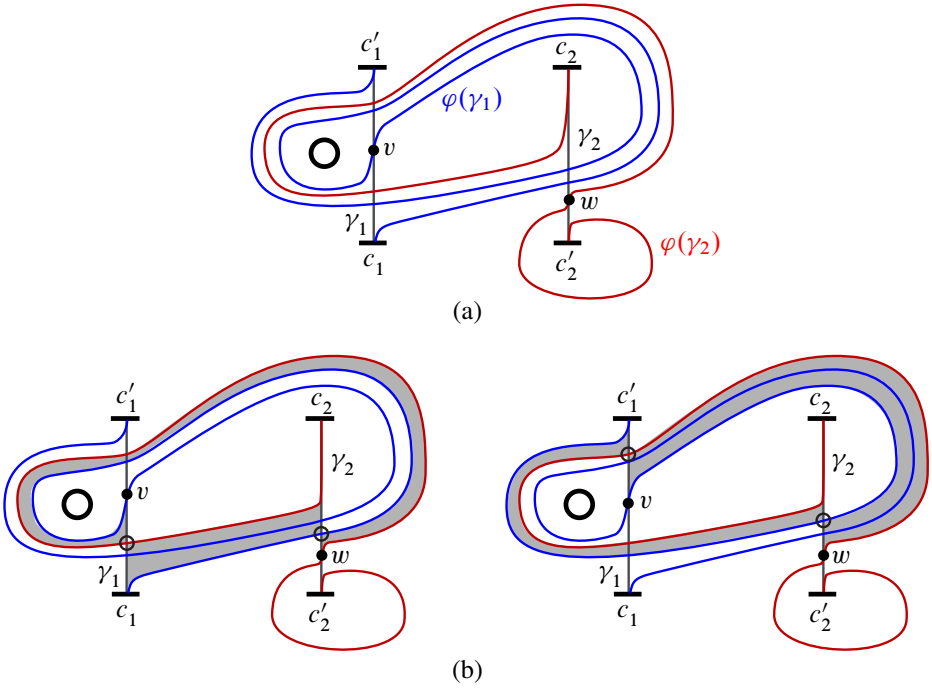


Figure 4: (a): The triple $(\Sigma, \varphi, \{\gamma_1, \gamma_2\})$ admits a consistent right position. The pairs of completing discs are shaded in (b).

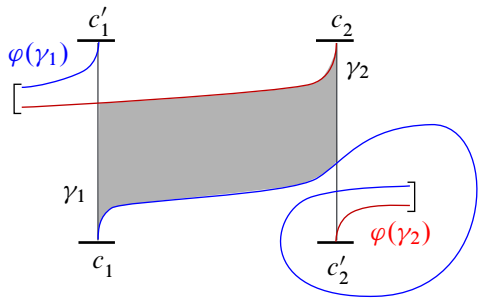


Figure 5: The augmented open book decomposition $(\Sigma, \varphi, \{\gamma_1, \gamma_2\})$ does *not* admit a consistent right position. The pair of brackets is meant to signify an identification, ie what goes in one comes out the other. This convention is followed throughout the paper.

no positive factorization for any n_1 and n_2. On the other hand, it is a simple exercise to show that for any positive n_1 and n_2 each boundary component is "protected", ie each properly embedded arc with endpoint on the boundary component under

consideration maps to the right at that point (it suffices to check it for either of the boundary components parallel to β_3 or β_4). In particular, each such monodromy is right-veering. See eg [7] for a more in-depth examination of this example.

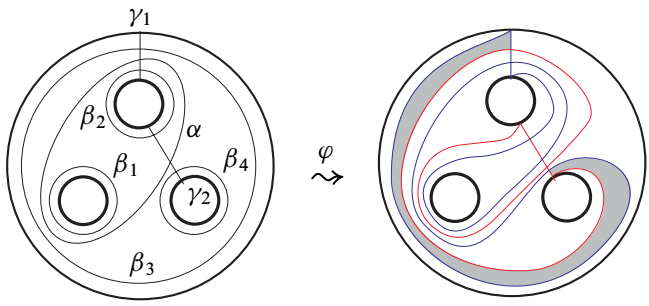


Figure 6: (a) Surface and curves for the open book decomposition denoted $(\Sigma_{0,4}, \tau_\alpha^{-2}\tau_{\beta_1}^{n_1}\tau_{\beta_2}^{n_2}\tau_{\beta_3}\tau_{\beta_4})$. (b) The triple $(\Sigma, \varphi, \{\gamma_1, \gamma_2\})$ does not admit any consistent right position, but φ is right-veering for all $m, n \geq 1$.

3.2 The right position associated to a factorization

Let $(\Sigma, \varphi, \Gamma)$ be an augmented open book decomposition, and φ admit a factorization ω into positive Dehn twists. The goal of this subsection is to associate to ω a consistent right position $\mathcal{P}_\omega(\Gamma)$ of $(\Sigma, \varphi, \Gamma)$. As mentioned in the introduction, we expect $\mathcal{P}_\omega(\Gamma)$ to play a role analogous to $\partial\Gamma$; the particular property we are interested in is that of being "away from" the support of (a representative of) φ. We are however working with a particular factorization of φ, so require a somewhat more refined notion. Recall that D_α refers to the Dehn twist about fixed $\alpha \in [\alpha]$, while τ_α is its isotopy class.

Definition 3.8 Let $(\Sigma, \varphi, \Gamma)$ be an augmented open book, $p \in (\Gamma \cap \varphi(\Gamma))$, and $\alpha \in \mathrm{SCC}(\Sigma)$. If there is a representative $\alpha' \in \mathrm{SCC}(\Sigma)$ of the isotopy class $[\alpha]$ such that the image $D_{\alpha'}(\varphi(\Gamma))$ can be isotoped to minimally intersect Γ while fixing a neighborhood of p in $D_{\alpha'}(\varphi(\Gamma))$, we say p is *fixable under* the mapping class τ_α. We will label the image $p^{\alpha'}$. Of course the identification of $p^{\alpha'} \in (\Gamma \cap (\tau_\alpha \circ \varphi)(\Gamma))$ depends on $\alpha' \in [\alpha]$.

Our goal then is to prove Theorem 1.1 by showing that the right position given by the set of all points of $\Gamma \cap \varphi(\Gamma)$ which are fixable under each successive twist in ω is consistent.

As a simple example, Figure 7 illustrates distinct right positions of a pair (φ, γ) associated to distinct factorizations of a mapping class φ.

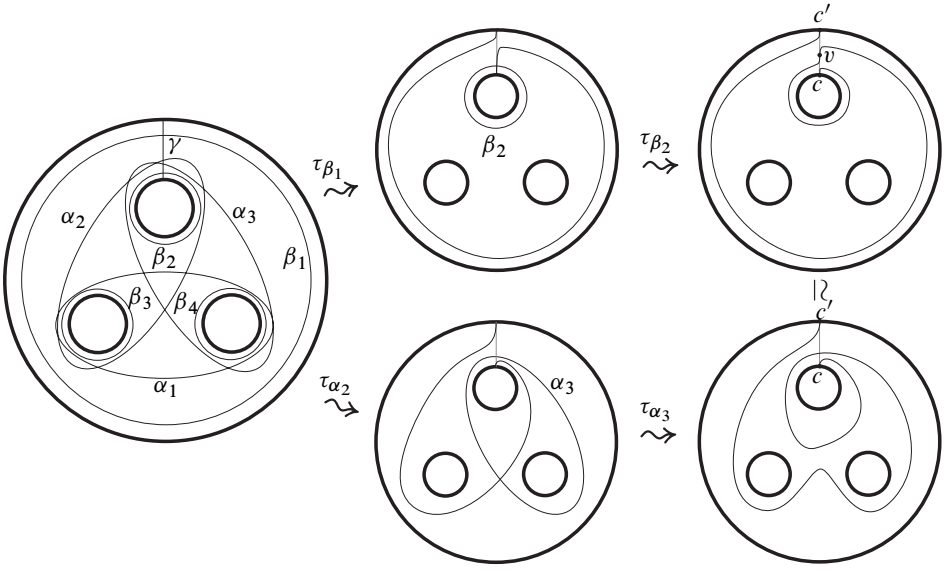


Figure 7: On the left are the curves of the well-known lantern relation on $\Sigma_{0,4}$. Setting $\omega_1 = \tau_{\beta_4}\tau_{\beta_3}\tau_{\beta_2}\tau_{\beta_1}, \omega_2 = \tau_{\alpha_3}\tau_{\alpha_2}\tau_{\alpha_1}$, the lantern relation tells us that ω_1, ω_2 are factorizations of a common $\varphi \in \mathrm{MCG}(\Sigma)$. Clearly, for the arc γ indicated, the intersection point $v \in \varphi(\gamma) \cap \gamma$ in the upper right figure is fixable under ω_1. However, as is clear from the lower sequence of figures, v is *not* fixable under the factorization ω_2. Thus the right position $\mathcal{P}_{\omega_1}$ contains the point v as well as the endpoints of γ, while $\mathcal{P}_{\omega_2}$ contains only the endpoints of γ.

3.3 Tools

The goal of this subsection is to develop the tools necessary for keeping track of fixable points through a given factorization.

Definition 3.9 Let $(\Sigma, \varphi, \Gamma)$ be an augmented open book, $\alpha \in \mathrm{SCC}(\Sigma)$, and $p \in \Gamma \cap \varphi(\Gamma)$ fixable under τ_α. The *image set* of p under τ_α is

$$i_{[\alpha]}(p) := \{q \in \Gamma \cap (\tau_\alpha \circ \varphi)(\Gamma) \mid \exists \alpha \in [\alpha] \text{ such that } q = p^\alpha\} = \bigcup_{\alpha \in [\alpha]} p^\alpha.$$

We begin with a brief description, with an illustrative example in Figure 8. Let $(\Sigma, \varphi, \Gamma)$ be an augmented open book, $\gamma \in \Gamma$, and α a representative of $[\alpha]$ which minimizes $\alpha \cap \gamma$ and $\alpha \cap \varphi(\gamma)$. Furthermore, choose support(D_α) so as not to intersect any point of $\gamma \cap \varphi(\gamma)$ (see Figure 8(a)). Consider then the image $D_\alpha(\varphi(\gamma))$, which differs from

$\varphi(\gamma)$ only in support(D_α) (see Figure 8(b)). Now, each bigon bounded by the arcs $D_\alpha(\varphi(\gamma))$ and γ contains vertices which were in some "triangular region" bounded by $\gamma \cup \varphi(\gamma) \cup \alpha$ in the original configuration (T_1 in the figure). In particular, there is an isotopy of $D_\alpha(\varphi(\gamma))$ over (possibly some subset of) these bigons which minimizes $\gamma \cap \tau_\alpha(\varphi(\gamma))$ (the issue of exactly when a proper subset of the bigons suffices is taken up below, in Section 3.4.2). Thus each point of $\mathcal{P}$ which is *not* a vertex of such a triangular region is fixable (see Figure 8(c)).

Note however that this description fails to see the full image set, and that moreover this failure may again be expressed in terms of triangular regions. Indeed, if T is any triangular region which does *not* become a bigon (eg T_2 in the figure), we may isotope the given α over T (eg go from Figure 8(a) to Figure 8(d), with $T = T_2$) to obtain $\alpha' \in [\alpha]$ such that $\alpha' \cap \gamma$ and $\alpha' \cap \varphi(\gamma)$ are also minimal. This isotopy has then the effect of changing the identification of our fixable point in the image (see Figure 8(f)). In particular, our image set must include each of these points (see Figure 8(g)).

3.3.1 Fixable points and triangular regions As indicated in the previous subsection, fixability may be characterized in an isotopy-invariant way by means of "triangular regions". We begin by setting out a precise framework for their study.

Definition 3.10 Suppose γ and γ' are properly embedded arcs in a surface Σ, isotoped to minimize intersection. Then for $\alpha \in \mathrm{SCC}(\Sigma)$, a *triangular region* T (of the triple $(\alpha, \gamma, \gamma')$) is the image of an immersion $f\colon \Delta \looparrowright \Sigma$, where Δ is a 2–simplex and the image of each edge is contained in one of α, γ or γ'.

Definition 3.11 A triangular region T for the ordered triple $(\alpha, \gamma, \gamma')$ is:

- *Essential* if α can be isotoped relative to $\partial(T) \cap \alpha$ so as to intersect γ and γ' in a minimal number of points.
- *Upward* (*downward*) if bounded by α, γ and γ' in clockwise (counterclockwise) order.

We continue with some technical results concerning triangular regions. Let γ, γ' be properly embedded, non-isotopic arcs in Σ, with $\partial(\gamma) = \partial(\gamma')$, isotoped to minimize intersection. Our motivating example, of course, is the case $\gamma' = \varphi(\gamma)$. Let $\alpha \in \mathrm{SCC}(\Sigma)$ be isotoped to be disjoint from $\gamma \cap \gamma'$. A *bigon* in this setup is an immersed disc B bounded by the pair (α, γ) or by (α, γ'). Similarly, a *bigon chain* is a set $\{B_i\}_{i=1}^n$ of bigons bounded exclusively by one of the pairs $(\alpha, \gamma), (\alpha, \gamma')$, for which the union $\bigcup_{i=1}^n \{\alpha \cap \partial(B_i)\}$ is a connected segment of α (see Figure 10(a)). Finally, we say points $p, p' \in \alpha \cap (\gamma \cup \gamma')$ are *bigon-related* if they are vertices in a common bigon chain.

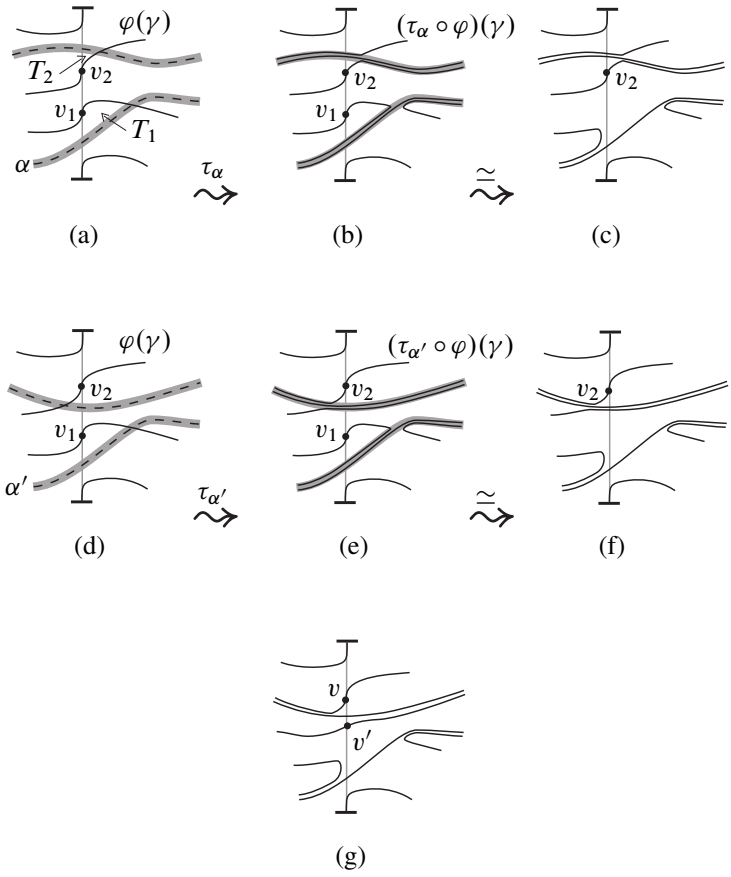


Figure 8: Top row: (a) The setup for the above discussion. The support of the twist D_α is shaded. (b) The result of D_α. There is a single bigon, in which v_1 is a vertex. (c) The result of isotoping $D_\alpha(\varphi(\gamma))$ over this bigon so as to minimize intersection with γ. Bottom row: Same as the top, but with α'. Note that while the images $D_\alpha(\varphi(\gamma))$ and $D_{\alpha'}(\varphi(\gamma))$ are of course isotopic, the fixed points of (c) and (f) differ. Finally, (g) indicates the isotopy-independent image set (ie the points v and v', along with $\partial(\gamma)$).

Suppose then that T is a triangular region in this setup, with vertex $v \in \gamma \cap \gamma'$. We define the *bigon collection associated to* T, denoted $\mathcal{B}_T$, as the set of points $\{p \in \alpha \cap (\gamma \cup \gamma') \mid p \text{ is bigon-related to a vertex of } T\}$ (see Figure 10(b)). Note that if a vertex of T along α is not a vertex in any bigon, the vertex is still included in $\mathcal{B}_T$, so that, for example, the bigon collection associated to an essential triangle T is just the vertices of T along α.

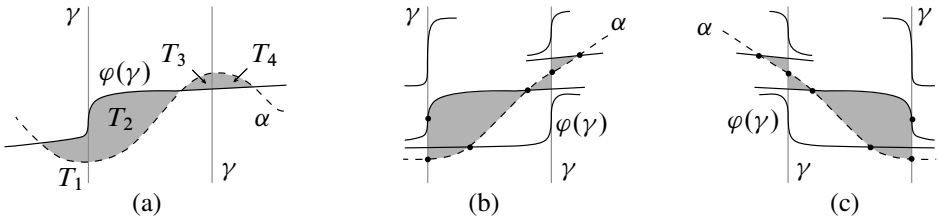


Figure 9: (a) If the shaded region is a maximal bigon chain, then T_1 and T_4 are essential triangular regions for $(\alpha, \gamma, \varphi(\gamma))$, T_2 and T_3 are not. (b) Upward triangular regions. (c) Downward triangular regions.

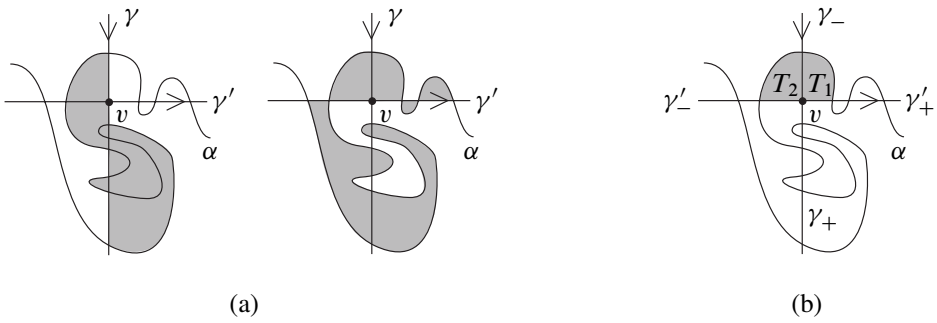


Figure 10: (a) The shaded regions are distinct maximal bigon chains. (b) As each chain from (a) has a vertex in each of the triangular regions T_1, T_2, the bigon collection $\mathcal{B}_{T_1} = \mathcal{B}_{T_2}$ includes each intersection point of each chain. In this example, $\sigma_v(\mathcal{B}_{T_1}) = (0, 1, 1, 0)$, so the collection is downward. Note that T_1 is an essential downward triangular region (Definition 3.11), while T_2 is non-essential.

Now, orienting γ and γ' so that the point v is positive, v separates each of γ and γ' into two segments, which we label $\gamma_+, \gamma_-, \gamma'_+$ and γ'_- in accordance with the orientation. For each bigon collection $\mathcal{B}$ associated to a triangle with vertex v, we define $\sigma_v(\mathcal{B}) := (|\mathcal{B} \cap \gamma_+|, |\mathcal{B} \cap \gamma_-|, |\mathcal{B} \cap \gamma'_+|, |\mathcal{B} \cap \gamma'_-|) \in (\mathbb{Z}_2)^4$, where intersection numbers are taken mod 2 (see Figure 10(b)).

Finally, we label a bigon collection $\mathcal{B}$ as

- *upward* (with respect to v) if $\sigma_v(\mathcal{B}) \in \{(1, 0, 1, 0), (0, 1, 0, 1)\}$,
- *downward* if $\sigma_v(\mathcal{B}) \in \{(1, 0, 0, 1), (0, 1, 1, 0)\}$,
- *non-essential* otherwise (see Figure 10(b)).

Thus the bigon collection associated to an upward (downward) essential triangular region T is itself upward (downward).

Lemma 3.12 *Let $\alpha \in [\alpha]$ be such that $\alpha \cap \gamma$ and $\alpha \cap \gamma'$ are minimal, and $v \in \gamma' \cap \gamma$. Then the number of essential triangles with vertex v in the triple $(\alpha, \gamma, \gamma')$ depends only on the isotopy class $[\alpha]$.*

Proof Let α be an arbitrary representative of $[\alpha]$ (in particular $\alpha \cap \gamma$ and $\alpha \cap \gamma'$ are not necessarily minimal), and v a vertex. We define an equivalence relation on the set of triangular regions with vertex v by $T_1 \sim T_2 \iff \mathcal{B}_{T_1} = \mathcal{B}_{T_2}$. In particular there is a 1-1 correspondence between $\{\text{triangular regions with vertex } v\}/\sim$ and the set of bigon collections associated to triangular regions with vertex v.

Now, if α' is another representative of the isotopy class $[\alpha]$, we may break the isotopy into a sequence $\alpha = \alpha_1 \simeq \cdots \simeq \alpha_n = \alpha'$, where each isotopy $\alpha_i \simeq \alpha_{i+1}$ is either a bigon birth, a bigon death, or does not affect either of $|\alpha \cap \gamma|$ and $|\alpha \cap \gamma'|$. Note that if $\mathcal{B}$ is upward or downward, an isotopy $\alpha_i \simeq \alpha_{i+1}$ which does not cross v does not affect $\sigma_v(\mathcal{B})$, while an isotopy $\alpha_i \simeq \alpha_{i+1}$ which crosses v changes $\sigma_v(\mathcal{B})$ by addition with $(1,1,1,1)$. In either case, the type of the triangle is preserved, ie we can keep track of an essential upward (downward) bigon collection through each isotopy. Furthermore, any two distinct bigon collections will have distinct images under each such isotopy (a bigon birth/death cannot cause two maximal bigon chains to merge). We therefore have an integer $a(v)$, defined as the number of essential bigon collections $\mathcal{B}$ associated to triangular regions with vertex v, which depends only on $[\alpha]$.

Finally, if we take α' to be a representative of $[\alpha]$ which intersects γ and γ' minimally, then $\{\text{triangular regions with vertex } v\}/\sim$ is by definition just the set of essential triangular regions with vertex v, and has size $a(v)$. □

Lemma 3.13 *Let α be a fixed representative of the isotopy class $[\alpha]$ which has minimal intersection with γ and γ', and $v \in \gamma \cap \gamma'$ a vertex of an upward triangle T. Then v is not a vertex of any downward triangular region T' with edge along α.*

Proof Consider a neighborhood of v, labeled as in Figure 11. As T is upward, it must be T_1 or T_3, while any downward T' must be T_2 or T_4. Without loss of generality then suppose $T = T_1$. But then neither of T_2, T_4 can be triangular regions without creating a bigon, violating minimality. □

It follows from Lemmas 3.12 and 3.13 that we may unambiguously refer to a vertex of an essential triangular region as downward, upward, or neither, in accordance with

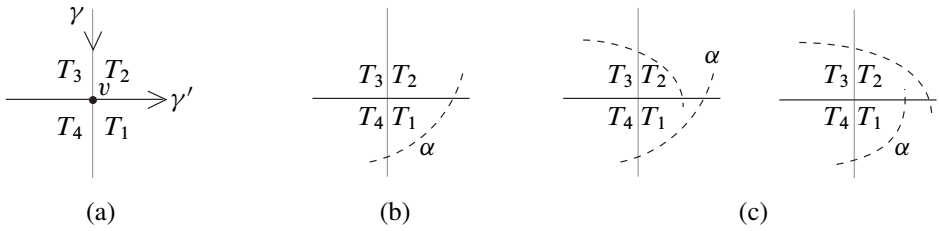


Figure 11: (a) A neighborhood of v. (b) The upward triangle T_1. (c) As α has no self-intersection, T_2 cannot be a triangular region without creating a bigon.

any essential triangular region of which it is a vertex. In particular, if α has minimal intersection with both γ and γ', then all triangular regions are essential. Henceforth we will drop the adjective "essential".

Definition 3.14 Suppose T is a triangular region with vertex $v \in \gamma \cap \gamma'$. Consider then an isotopy of α, supported in a neighborhood of T, which crosses v, thus giving another triangular region T' of the same type (upward/downward) as T with vertex v, and such that in some neighborhood of v, $T \cap T' = v$ (see Figure 12). Note that if T contains sub-regions with vertex v then the process involves isotoping the innermost region first, and proceeding to T itself. We call such an isotopy a *shift over* v.

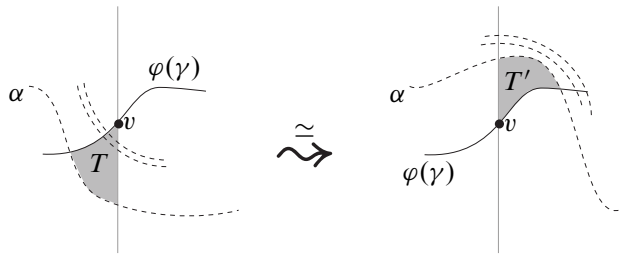


Figure 12: A shift over a vertex v

Observation 3.15 The relation between fixability and triangular regions is given by the following readily verifiable facts:

- A downward point is always fixable (see Figure 13(a)).
- An upward point is *not* fixable (see Figure 13(b)), unless there is another upward triangle sharing the other two vertices, corresponding to a canceling bigon in the image (see Figure 13(c)). Note that, in this case, while either of the points v, v' is individually fixable, the pair is not necessarily simultaneously fixable.

- The reason for demanding that a *neighborhood* of the point be fixed is illustrated in Figure 13(b): though the intersection-minimizing isotopy may be done without removing the intersection point p, there is no fixable neighborhood, and so p is not fixable.

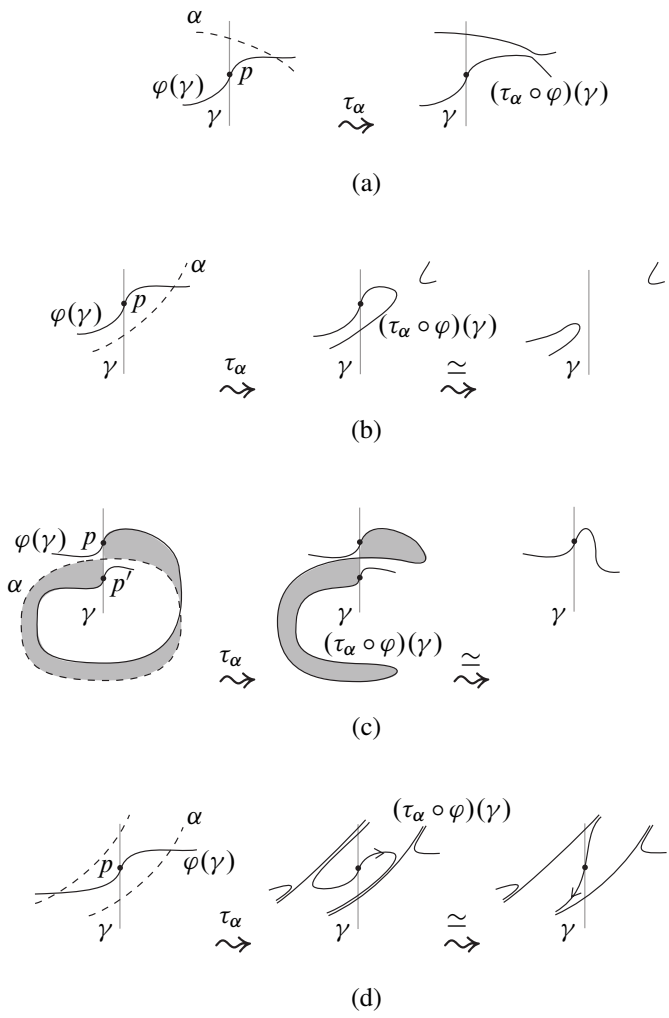


Figure 13

3.4 Image sets

Let $(\Sigma, \varphi, \Gamma)$ be an augmented open book decomposition. We will use the results of the previous subsection to describe the image set of an arbitrary point $p \in \Gamma \cap \varphi(\Gamma)$ under a

Dehn twist τ_α. We begin by choosing a representative α of the isotopy class $[\alpha]$ which has minimal intersection with $\Gamma \cup \varphi(\Gamma)$. Choose support$(D_\alpha)$ to be disjoint from p. We will find it helpful to explicitly notate the inclusion $i_\alpha\colon \Gamma \cap \varphi(\Gamma) \hookrightarrow \Gamma \cap D_\alpha\varphi(\Gamma)$. We may then refer to fixability of $i_\alpha(p)$, which simply means that there is a isotopy of $D_\alpha\varphi(\Gamma)$, supported away from a neighborhood of $i_\alpha(p)$, which minimizes intersection with Γ.

A general observation of which we will make extensive use is that for a fixable point p, the identification of p^α in the image is unchanged by isotopies of α which do not cross p. To be precise:

Lemma 3.16 *Let $p \in (\Gamma \cap \varphi(\Gamma))$ and $\alpha \in \mathrm{SCC}(\Sigma)$ be such that p is fixable under the Dehn twist τ_α. If $h_t\colon \Sigma \to \Sigma$ is an isotopy such that $h_0 = \mathrm{Id}$, and $h_t\alpha$ does not cross p, then $i_{h_1\alpha}(p)$ is also fixable, and $p^\alpha = p^{h_1\alpha}$.*

Proof Label Γ so that $p \in \gamma_1 \cap \varphi(\gamma_2)$. By assumption there is an isotopy g_t, supported away from $i_\alpha(p)$, such that $g_0 = \mathrm{Id}$ and $g_1 D_\alpha(\varphi(\gamma_2))$ has minimal intersection with γ_1. Thus $g_t \circ D_{h_{1-t}\alpha}(\varphi(\gamma_2))$ is an isotopy from $D_{h_1\alpha}(\varphi(\gamma_2))$ to $g_1 D_\alpha(\varphi(\gamma_2))$ which fixes $i_{h_1\alpha}(p)$. The second statement, that the image is unchanged, is immediate. □

Now, if $p \in \gamma_1 \cap \varphi(\gamma_2)$ is neither downward nor upward in $(\alpha, \gamma_1, \varphi(\gamma_2))$, then p is clearly fixable, with a unique image point. Our interest then lies in downward/upward points.

3.4.1 The image set of a downward point Let $p \in \Gamma \cap \varphi(\Gamma)$ be downward with respect to α. As p is a vertex of no bigon in $\Gamma \cup D_\alpha((\varphi(\Gamma))$, we see that p^α is a unique point whose identification within the image set of p will (by Lemma 3.16) hold for any isotopy of α which does not involve a shift over p. Moreover, each such shift will change the identification, so that if p is the vertex of m distinct downward triangles, $i_{[\alpha]}(p)$ will contain exactly $m+1$ points (see Figure 14).

3.4.2 The image set of an upward point For the case that $p \in \Gamma \cap \varphi(\Gamma)$ is upward with respect to α, assume for the moment that α has been isotoped such that the triangular region T_p is embedded. Note then that T_p corresponds to a bigon R in $\Gamma \cup D_\alpha((\varphi(\Gamma))$ in which $i_\alpha(p)$ is a vertex (see Figure 15(a), (b)). In particular, $i_\alpha(p)$ is fixable only if there is an second bigon R', which intersects R in the vertex which is not $i_\alpha(p)$ (see Figure 15(c)). Observe then that R' also corresponds to an upward triangle, $T_{p'}$, for a point $p' \in \Gamma \cap \varphi(\Gamma)$, such that $T_p \cap T_{p'}$ (for the given α in its isotopy class) consists of the two vertices along α (see Figure 15(d)). We refer to such $T_{p'}$ as the *pair* of T_p, and say that each of T_p and T_p' is *paired*. Note that each

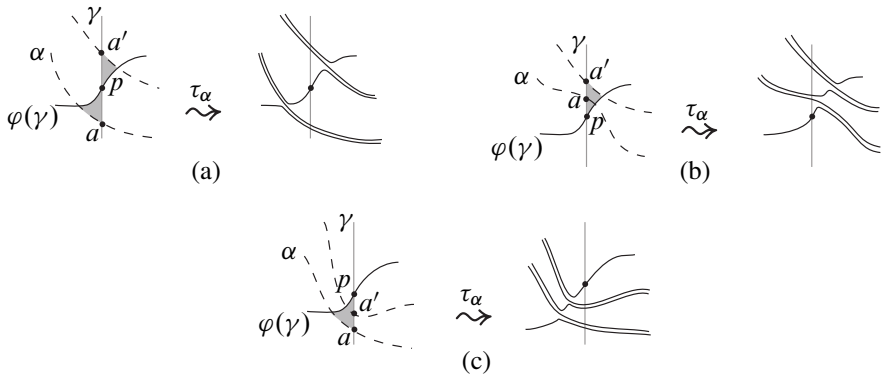


Figure 14: The point p is downward in 2 distinct triangles. (a), (b), and (c) show representatives of the possible shift-isotopy classes, each of which gives a distinct point $i_\alpha(p)$, while $i_{[\alpha]}(p)$ is the same for each.

of $i_\alpha(p)$ and $i_\alpha(p')$ are thus fixable, though not simultaneously, and have the same image. If $i_\alpha(p)$ is *not* fixable, we refer to T_p as *unpaired*. Note further that, while the property of being paired depends on the representative α, using Lemma 3.16 a paired region on a point p will remain paired for any isotopy of α which does not cross p (though the regions may cease to be embedded).

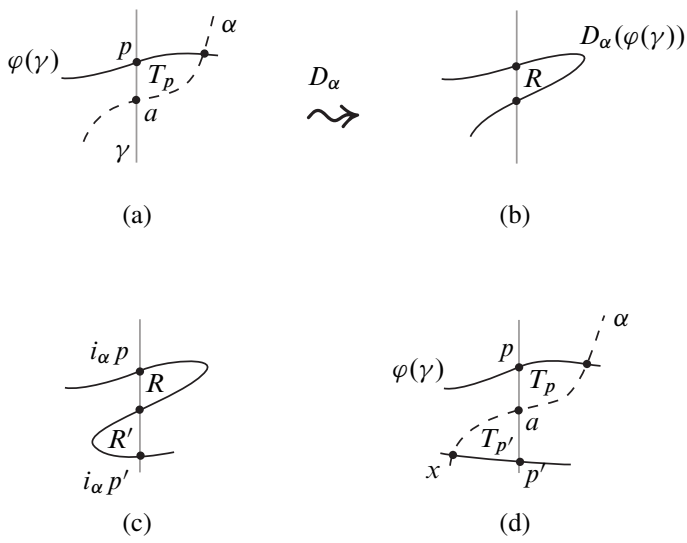


Figure 15: (a) An upward point, and (b) a corresponding bigon in the image. (c) Canceling bigons, and (d) corresponding paired regions.

So, we see that for given α, and fixable point p, $i_\alpha(p)$ will be fixable if and only if each upward triangle on p is paired. This motivates the following definition.

Definition 3.17 Let $(\Sigma, \varphi, \Gamma)$ be an augmented open book decomposition, and $\alpha \in \mathrm{SCC}(\Sigma)$. Then α is *nice* (with respect to $\Gamma \cup \varphi(\Gamma)$) if, whenever $p \in \Gamma \cap \varphi(\Gamma)$ is fixable, then $i_\alpha(p)$ is fixable under some isotopy of $D_\alpha(\varphi(\Gamma))$ which eliminates all bigons with Γ (see Figure 16).

Figure 16: (a) The curve α is not nice with respect to γ_1 and $\varphi(\gamma_2)$, as $i_\alpha(p_1)$ is not fixable. (b) The situation is remedied by a shift over p_1.

So, if α is nice, then for each fixable p, $i_\alpha(p)$ is fixable. As it turns out, niceness is not a particularly demanding requirement. Indeed:

Lemma 3.18 *Let $(\Sigma, \varphi, \Gamma)$ be an augmented open book decomposition, and $\alpha \in \mathrm{SCC}(\Sigma)$. Suppose $P \subset (\Gamma \cap \varphi(\Gamma))$ is a collection of upward, fixable points such that for each $p \in P$, $i_\alpha(p)$ is fixable. Then there is a nice $\alpha' \in [\alpha]$ such that for each $p \in P$, $i_{\alpha'}(p) = i_\alpha(p)$; ie all identifications are preserved.*

Proof If α is not nice, then there is some upward fixable point p_1 such that the associated triangle T_{p_1} is not paired. We may assume T_{p_1} is an innermost such region. Using Lemma 3.16, we want to show that T_{p_1} contains no upward fixable point, so that the shift over T_{p_1} changes no identifications. Suppose otherwise, so there is upward fixable $p_2 \in T_{p_1}$, with associated $T_{p_2} \subset T_{p_1}$ (see Figure 17(a)); as we are assuming T_{p_1} is innermost among unpaired regions, T_{p_2} is paired (see Figure 17(b)). To keep track of things, we index the involved arcs so that $p_1 \in \gamma_1 \cap \varphi(\gamma_2)$ and $p_2 \in \gamma_3 \cap \varphi(\gamma_4)$.

Let $T_{p_2'}$ be the pair of T_{p_2}, so $p_2' \in \gamma_3 \cap \varphi(\gamma_4)$. Now, T_{p_2} is such that each of the 2 vertices along α of T_{p_1} lies in the interior of the edge along α of $T_{p_2'}$ (see Figure 17 (b)). Thus following along γ_1 and $\varphi(\gamma_2)$ away from p_1 into $T_{p_2'}$, they must intersect in some point p_1', thus giving a region $T_{p_1'} \subset T_{p_2'}$ which is paired with T_{p_1} (see Figure 17(c)), a contradiction.

We may then shift α over T_{p_1}, and repeat the process with any remaining innermost unpaired regions to obtain a nice α' as desired. □

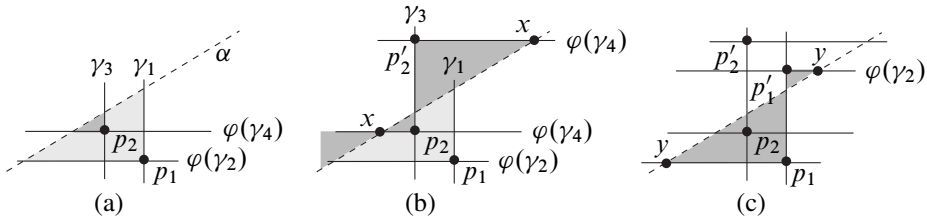


Figure 17: (a) Region T_{p_1} is lightly shaded, T_{p_2} is dark. (b) The pair of regions T_{p_2} and $T_{p_2'}$ are darkly shaded. As usual, the figures are in the universal cover, with all lifts of a given object given the same label as the object. (c) The region pair T_{p_1} and $T_{p_1'}$ is shaded.

Definition 3.19 Motivated by the above, given nice α, and not-necessarily-distinct arcs γ_1 and γ_2, we introduce an equivalence relation $\overset{\alpha}{\sim}$ on intersection points $\gamma_1 \cap \varphi(\gamma_2)$, where $p \overset{\alpha}{\sim} p' \iff p^\alpha = p'^\alpha$ (see Figure 18). Associated to each such equivalence class $[p]_\alpha$ are then the unique minimal connected sub-arcs of γ_1 and $\varphi(\gamma_2)$ which contain each point of the class; we refer to these as the *vertical* (along γ_1) and *horizontal* (along $\varphi(\gamma_2)$) spans of the class. Observe that the image under D_α of the horizontal span is entirely contained in the boundary of a bigon chain of even length in $D_\alpha(\varphi(\gamma_2))$ and γ_1.

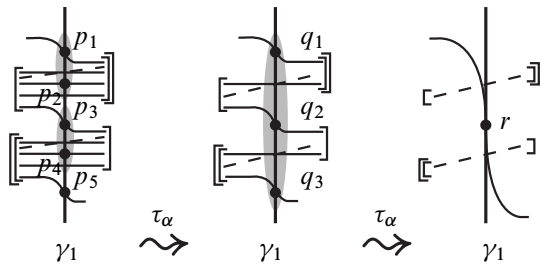


Figure 18: Equivalence classes are shaded. In the left-most configuration, we have $i_\alpha(p_1) = i_\alpha(p_2) = q_1$, $i_\alpha(p_3) = i_\alpha(p_4) = q_2$, and $i_\alpha(p_5) = q_3$, while in the second twist $i_\alpha(q_i) = r$ for each i.

Lemma 3.20 *For an arc σ, and points $a, b \in \sigma$, let $[a, b]_\sigma$ denote the (closed) segment of σ from a to b. Suppose $p, p' \in \gamma_1 \cap \varphi(\gamma_2)$ are consecutive (along γ_1) elements of the equivalence class $[p]_\alpha$. Then any nice α intersects each of $[p, p']_{\gamma_1}$ and $[p, p']_{\varphi(\gamma_2)}$ exactly once.*

Proof As p is fixable, we may isotope α to α', so that there are paired, in particular embedded, triangles T_p and $T_{p'}$ on the given points. Thus α' satisfies the desired

intersection conditions, and by Lemma 3.18 may be assumed nice. In particular, $p \overset{\alpha'}{\sim} p'$. Now, the isotopy $\alpha \rightsquigarrow \alpha'$ can involve a shift over at most one of T_p and $T_{p'}$, but either of these will change the image under i_α of exactly one of the points, so that p and p' are no longer in the same equivalence class, a contradiction. So the isotopy crosses neither point, and so α satisfies the same intersection conditions. □

Now, Lemma 3.20 implies in particular that if $p, p' \in \gamma_1 \cap \varphi(\gamma_2)$ are consecutive (along γ_1) elements of the equivalence class $[p]_\alpha$, then the closed curve obtained as the union of $[p, p']_{\gamma_1}$ and $[p, p']_{\varphi(\gamma_2)}$ along the endpoints is isotopic to α through an isotopy which crosses neither p nor p'. Suppose then that there is a further point p'' such that the union of $[p', p'']_{\gamma_1}$ and $[p', p'']_{\varphi(\gamma_2)}$ along the endpoints is a (nice) curve α' again isotopic to α. Then either $\alpha' \cap [p, p']_{\gamma_1} \neq \varnothing$, so that $p'' \in [p]_\alpha$, or otherwise, so there is an isotopy $\alpha \rightsquigarrow \alpha'$ given by a shift over a triangular region with vertex p'. Using this observation, we have:

Lemma 3.21 *If $p \in \gamma_1 \cap \varphi(\gamma_2)$, and α a nice representative of $[\alpha]$, are such that $[p]_\alpha$ contains greater than 2 elements, then any nice $\alpha' \in [\alpha]$ is isotopic to α through an isotopy which intersects no element of $[p]_\alpha$, ie $[p]_\alpha = [p]_{\alpha'}$.* □

As an example of Lemma 3.21, consider again Figure 18. Observe that α may be isotoped over p_2 and p_4 to a nice representative α', thus changing the equivalence relation to $p_2 \sim p_3$ and $p_4 \sim p_5$. In the second picture (from either side), however, there can be no nice representative of α which does not preserve $q_1 \sim q_2 \sim q_3$.

3.5 Consistency

Having collected our tools, we return to Theorem 1.1, so consistency of the right position $\mathcal{P}_\omega$ given by the union of all fixable points under a given positive factorization of a mapping class. Our proof will be by induction on the number of twists in our given factorization ω, so we will consider a more general question: given a consistent right position $\mathcal{P} = \mathcal{P}(\Sigma, \varphi, \Gamma)$, and $\alpha \in \mathrm{SCC}(\Sigma)$, is the right position for $(\Sigma, \tau_\alpha \circ \varphi, \Gamma)$ given by $\bigcup_{p \in \mathcal{P}} i_{[\alpha]}(p)$, which we will denote simply by $\mathcal{P}^\alpha$, again consistent?

We will for the remainder of the section consider only the case that our arc collection Γ consists of exactly two arcs, which we label γ_1 and γ_2. We will focus attention on a region $B \in R(\mathcal{P})$. As such, B has an edge along each of γ_i and $\varphi(\gamma_i)$, $i = 1, 2$. We will find it convenient to label the edges of B as e_i, $i = 1, 2, 3, 4$, where e_1 denotes the edge along $\varphi(\gamma_2)$, e_1 the edge along $\varphi(\gamma_2)$, and for $i = 3, 4$, e_i is the edge along γ_{i-2}. Finally, we let v_i and y_i respectively denote the $\bullet$–point and $\circ$–point of B along γ_i (see Figure 19).

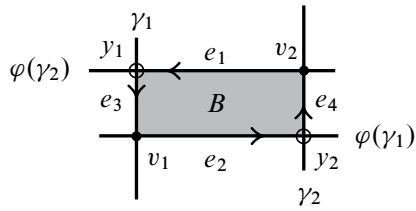


Figure 19

With an eye toward somewhat reducing the combinatorial aspects of the lemmas to come, we would like to exploit the symmetry between regions and their completions, and also that between $\bullet$–points and $\circ$–points.

Remark 3.22 Given $(\Sigma, \varphi, \Gamma)$ such that φ has factorization $\tau_{\alpha_n} \cdots \tau_{\alpha_1}$, consider $(\overline{\Sigma}, \overline{\varphi}, \Gamma)$, where $\overline{\Sigma}$ is gotten by reversing the orientation of Σ, and $\overline{\varphi}$ has factorization $\tau_{\alpha_n}^{-1} \cdots \tau_{\alpha_1}^{-1}$. We refer to $(\overline{\Sigma}, \overline{\varphi}, \Gamma)$ as the *reflection* of $(\Sigma, \varphi, \Gamma)$. Note that, eg if $p \in \Gamma \cap \varphi(\Gamma)$ is upward (for some α), then its reflection is downward. Furthermore, given $p \in \Gamma \cap \varphi(\Gamma)$ and $q \in \Gamma \cap \tau_\alpha(\varphi(\Gamma))$, we have $q \in i_{[\alpha]}(p)$ if and only if the reflection of p is in the $\alpha-$image set of the reflection of q (see Figure 20(a)). Finally, if we have a right position $\mathcal{P}$, and associated initially parallel regions $\mathcal{R}(\mathcal{P})$, we obtain a collection of regions in the reflection by simply switching $\circ$–points and $\bullet$–points; ie if p is an $\bullet$–point ($\circ$–point) of a region, its reflection is a $\circ$–point ($\bullet$–point). We denote the reflection of a region B by $\overline{B}$, and decorate the vertices accordingly (see Figure 20(b)).

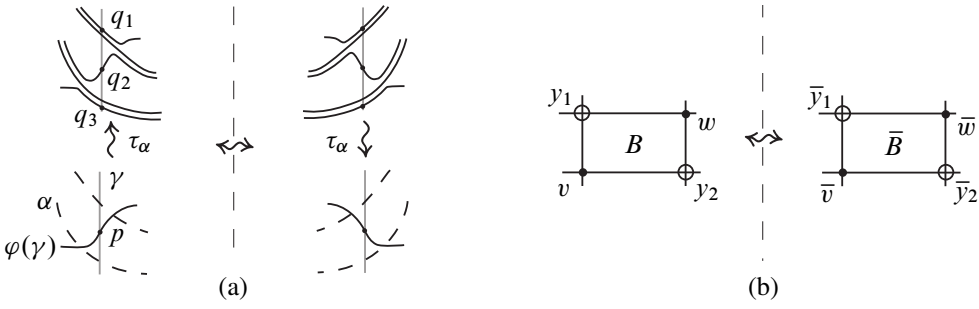


Figure 20: (a) A point $p \in \gamma \cap \varphi(\gamma)$, its image set $i_{[\alpha]}(p)$, and the reflections of each of these (note that the dots here are simply for labeling purposes, and not meant to indicate $\bullet$–points). (b) Reflecting a region.

Definition 3.23 Let $B \in \mathcal{R}(\mathcal{P})$ and let $\alpha \in \mathrm{SCC}(\Sigma)$ be a fixed, nice representative of its isotopy class. We say B is *collapsed* (by α) if (1) each edge is contained in the

span of the $\alpha-$equivalence class of its $\circ$–point vertex, and (2) $\alpha \cap B \neq \varnothing$. Similarly, a region $B^\alpha \in \mathcal{R}(\mathcal{P}^\alpha)$ is *created* (by α) if its reflection is collapsed (see Figure 21).

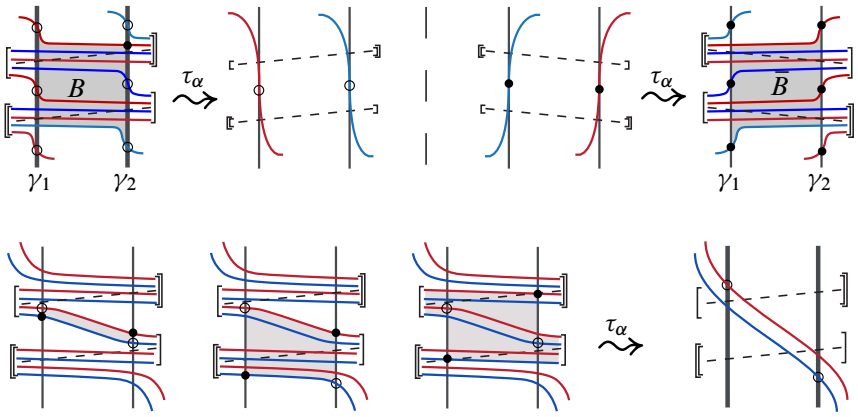


Figure 21: Above: To the left, a collapsed region B, and its (trivial) image. To the right, the reflection of B is a created region $\bar{B}$. Below: The three figures to the left are various regions in a common configuration, which is mapped by τ_α to the right-most figure. From left to right, these 3 regions are non-collapsed (satisfies condition (1) but not condition (2)), non-collapsed (satisfies condition (2) but not condition (1)), and collapsed (but not embedded).

We have:

Lemma 3.24 *Suppose $B \in \mathcal{R}(\mathcal{P})$ (not necessarily consistent) is collapsed by some α. Then there is a $B^c \in \mathcal{R}(\mathcal{P})$ with the same $\bullet$–points as B, and opposite orientation. Similarly, if $B^\alpha \in \mathcal{R}(\mathcal{P}^\alpha)$ is created by α then there is a $B^{\alpha c} \in \mathcal{R}(\mathcal{P}^\alpha)$ with the same $\circ$–points as B^α, and opposite orientation.*

Proof We begin with the first statement. We begin by assigning a *length* $l(B) = \#|[y_1]_\alpha \cap e_1|$ and *width* $w(B) = \#|[y_1]_\alpha \cap e_3|$ (so, eg, the region B in the upper left corner of Figure 21 has length 1, width 2, while the collapsed but not embedded example on the lower row has length = width = 1). Furthermore, we will order points along a given oriented arc in accordance with that orientation, so if p and p' are points along oriented arc σ, then $p' >_\sigma p$ means that the orientation of σ points from p to p'.

We first consider the case $l = w = 1$, and orient all arcs such that ∂B has the standard (counterclockwise) orientation (see Figure 22(a)). By the collapsed condition, $[y_1]_\alpha$

contains at least one element greater than v_1 along γ_1; let y_1' denote the minimal such element. As $w = l = 1$, we see that y_1' and y_1 are consecutive in the class, and so $v_1 \in [y_1, y_1']_{\gamma_1}$ and $v_2 \in [y_1, y_1']_{\varphi(\gamma_2)}$. In particular, y_2 is a vertex of an embedded upward triangular region T_{y_2}; by niceness of α there is then a y_2' and an embedded pair $T_{y_2'}$. Furthermore, $v_1 \in [y_2, y_2']_{\varphi(\gamma_1)}$ and $v_2 \in [y_2, y_2']_{\gamma_2}$. Now observe that by isotoping α to pass through each $\bullet$–point, we may realize B as $T_{y_1} \cup T_{y_2}$. But then the union of the complementary regions $T_{y_1'}$ and $T_{y_2'}$ gives a region B^c as desired.

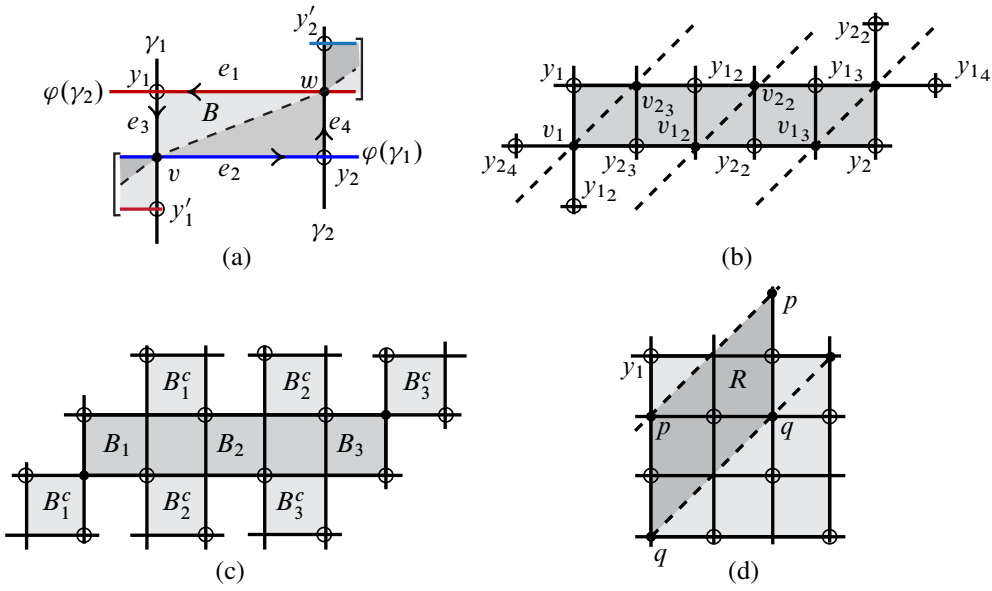


Figure 22

Consider then the case $l(B) > 1, w(B) = 1$ (see Figure 22(b)). Let $y_{1_i} \in [y_1]_\alpha$, $i = 1, \dots, l+1$, be consecutive elements of the class, with $y_{1_1} = y_1$ and $y_{1_2} \in e_1$. As $w = 1$, we have $y_{1_2} >_{\gamma_1} v_1$. Now, each y_{1_i}, $1 \le i \le l$, is the endpoint of a segment of γ_1 which cuts across B; let v_{1_i} denote the opposite (on e_2) endpoint of this segment (so $v_{1_1} = v_1$). Similarly, we have $y_{2_i} \in [y_1]_\alpha$, v_{2_i} along e_2. For notational convenience, we augment $\mathcal{P}$ to $\mathcal{P}'$ by including the v_{i_j}. There are then $B_i \in \mathcal{R}(\mathcal{P}')$, $i = 1, \dots, l$, where each B_i is a subregion of B with vertices $y_{1_i}, v_{1_i}, y_{2_{l+1-i}}$ and $v_{2_{l+1-i}}$ (see Figure 22(c)). In particular, each B_i is collapsed, and $l(B_i) = w(B_i) = 1$, so by the previous paragraph there are complementary regions B_i^c. But then the union of the regions B_i^c with the complement in B of the B_i is a region B^c as desired. Observe that $w(B^c) = l(B)$, and $w(B) = l(B^c)$; in particular, the argument for the case $w(B) > 1, l(B) = 1$ is identical, with the roles of the B_i and B_i^c swapped.

Finally, suppose both $l(B) < 1$ and $w(B) < 1$. Let p denote the (unique) intersection of α with $[y_1, y_{1_2}]_{\gamma_1}$, q the intersection with $[y_{1_2}, y_{1_3}]_{\gamma_1}$. Then, lifting to the universal cover (see Figure 22(d)), we find a rectangular region R bounded by $\rho^{-1}(\alpha)$ and $\rho^{-1}(\gamma_1)$, and whose vertices are consecutive (along $\rho^{-1}(\alpha)$) lifts of p and q (where $\rho\colon \widetilde{\Sigma} \mapsto \Sigma$ is the covering map). But then $\rho(R)$ is an annulus in Σ bounded by two copies of α, a contradiction.

The second statement then follows immediately by reflection: if B is created by α, it is the reflection of a collapsed (by α) region $\bar{B}$. There is thus $\bar{B}^c$ as above, whose reflection B^c therefore shares the $\circ$–points of B, and has the opposite orientation, as desired. □

Many of the arguments of this subsection will involve keeping track of equivalence classes of upward vertices. An observation which will prove particularly useful throughout is the following:

Lemma 3.25 *Let $(\Sigma, \varphi, \Gamma)$ be an augmented open book, and $\alpha \in \mathrm{SCC}(\Sigma)$. Suppose $p, p' \in (\Gamma \cap \varphi(\Gamma))$ are consecutive in $[p]_\alpha$, and further that $q, q' \in (\Gamma \cap \varphi(\Gamma))$ are such that p, p', q and q' are the vertices of a rectangular region. Then there is an $\alpha' \in [\alpha]$ such that q and q' are consecutive in $[q]_{\alpha'}$.*

Proof While this is obvious, the point of view which makes it so perhaps bears reinforcing. As such, consider the case that the edges $[p, p']$ and $[q, q']$ of the given rectangular region are along elements of Γ. Now, as p, p' are consecutive in $[p]_\alpha$, we have a pair of upward triangles T_p and $T_{p'}$, which have in common 2 vertices, one along an element of Γ, and one, which we label x, along an element of $\varphi(\Gamma)$. As usual, we picture the setup in the universal cover (see Figure 23(a)). Expanding our view of the figure in the universal cover to include 2 copies of our original figure (see Figure 23(b)), we find that q and q' are themselves vertices in upward triangular regions, and that using these regions α may be isotoped to some nice α' for which q and q' are consecutive in $[q]_{\alpha'}$ (see Figure 23(c)). The modifications necessary for the case of edges along elements of $\varphi(\Gamma)$ are clear (see Figure 23(d)). □

Now, recall that, as in the discussion at the beginning of the previous subsection, if some intersection $p \in (\Gamma \cap \varphi(\Gamma))$ is both upward and fixable for a given α, then there is some $p' \in [p]_\alpha$, and upward triangular regions T_p and $T_{p'}$ such that the entirety of α is contained in their edges; it follows that in this case every triangular region of $(\Sigma, \varphi, \Gamma)$ is upward. As such, we refer to a region with fixable vertices as *upward* if some vertex is in an upward triangular region.

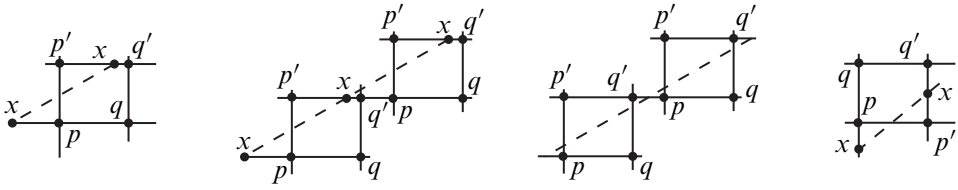


Figure 23: Figures for Lemma 3.25

Our argument depends on being able to consider a given region of $\mathcal{R}(\mathcal{P}^\alpha)$ as the "image" under τ_α of a region in (possibly an augmentation of) $\mathcal{R}(\mathcal{P})$. Now, it is clear intuitively that a region should not have an image if it is collapsed, or if its vertices are not all fixable under τ_α. We will further impose a minimality condition, to obtain:

Definition 3.26 Let $(\Sigma, \varphi, \Gamma)$ be an augmented open book, and $\alpha \in \mathrm{SCC}(\Sigma)$. Let $B \in \mathcal{R}(\mathcal{P})$ be non-collapsed (for α), and have fixable $\circ$–points. Then B is a *preimage for i_α* if

(1) B is not upward, and no downward triangular region on either $\bullet$–point is contained in B, or

(2) B is upward, and the equivalence class of neither $\circ$–point contains any point in the interior of an edge of B.

Definition 3.27 Given $B \in \mathcal{R}(\mathcal{P})$, we refer to any region of $\mathcal{R}(\mathcal{P}^\alpha)$ whose vertices are each in the respective image sets of the vertices of B as an *image* of B.

As some justification for this terminology, we have the following three lemmas relating pre-images and images, from which the proof of Theorem 1.1 will follow easily.

Lemma 3.28 *Let $B \in \mathcal{R}(\mathcal{P})$ be a preimage for i_α. Then there is an image B^α of B, such that i_α maps the $\circ$–points of B to those of B^α.*

Proof To simplify the picture, let $\widetilde{B}$ be a lift of B to the universal cover $\widetilde{\Sigma}$, and $\{\widetilde{\alpha}_i\}$ the connected components of $\rho^{-1}(\alpha)$ (where $\rho\colon \widetilde{\Sigma} \to \Sigma$ is the covering map). The lifted map $\widetilde{D}_\alpha$ has disjoint support, each connected component of which contains some $\widetilde{\alpha}_i$; let $D_{\widetilde{\alpha}_i}$ denote the restriction of $\widetilde{D}_\alpha$ to this component of the support. Now, a given $D_{\widetilde{\alpha}_i}$ affects $\widetilde{B}$ non-trivially only if $\widetilde{\alpha}_i$ has non-empty intersection with at least one of the edges $\widetilde{e}_1$ and $\widetilde{e}_2$. These $\widetilde{\alpha}_i$ can then be distinguished as those which intersect exactly one of the edges $\widetilde{e}_1$ and $\widetilde{e}_2$, which we refer to as a *corner* intersection, and those which intersect both edges, which we refer to as a *vertical* intersection.

Now, if $\widetilde{\alpha}_i$ has a vertical intersection, $D_{\widetilde{\alpha}_i}$ will slice each lift of B, and reattach to form new regions whose vertices map to the correct vertices under ρ (see Figure 24(a)). We are then done as long as no vertex is contained in a bigon in the resulting configuration. This however follows easily from our assumptions: suppose otherwise, so $\widetilde{\alpha}_i \cap \widetilde{B}$ is contained in an edge of an upward triangular region, say on $\widetilde{y}_1$. Now, as y_1 is fixable, there is $y_1' \in [y_1]_\alpha$, distinct from y_1. But then condition (2) of Definition 3.26 implies that each edge of B with endpoint y_1 is contained in the span of $[y_1]$ (see Figure 24(b)). Thus B is collapsed, a contradiction.

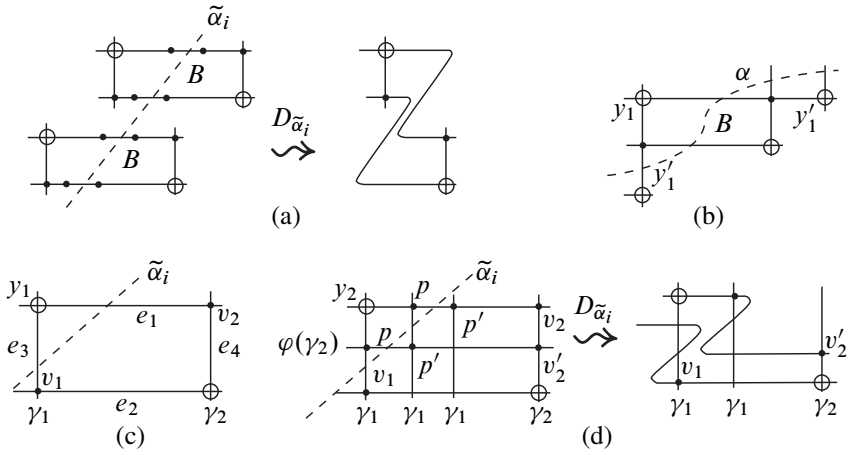


Figure 24: (a): A vertical intersection and the effect of the corresponding $D_{\widetilde{\alpha}_i}$ on the lifted region. (b) If a vertical segment of $\alpha \cap B$ is contained in an upward triangle, then B cannot be a pre-image. (c) A corner intersection, and (d) the effect of $D_{\widetilde{\alpha}_i}$.

Consider then an (upward) corner intersection; such $\widetilde{\alpha}_i$ is upward on a $\circ$–point, say $\widetilde{y}_1$ (see Figure 24(c)). As y_1 is fixable, and B not collapsed, there are then points $p, p' \in \gamma_1 \cap e_1$, $p <_{\gamma_1} p'$ consecutive in $[p]_\alpha$, such that $\widetilde{\alpha} \cap [\widetilde{p}, \widetilde{p'}]_{\widetilde{e}_1}$ is a single point. As $\widetilde{\alpha}_i$ exits $\widetilde{B}$ through $\widetilde{e}_3$, the point p also lies along e_3 (ie the edge of B contained in γ_1) (see Figure 24(d)). Now, traveling along $\widetilde{e}_1$ from $\widetilde{y}_1$, turning right onto $\widetilde{\alpha}_i$ and traveling exactly one length of α, one thus arrives at a second lift of $\varphi(\gamma_2)$, at a point in the interior of $\widetilde{B}$. Letting $\widetilde{v_2'}$ denote the intersection of this lift with $\widetilde{\gamma}_2$, by Lemma 3.25 we have $v_2' := \rho(\widetilde{v_2'}) \in [v_2]_{\alpha'}$. In particular, $i_\alpha(v_2') \in \mathcal{P}_2^\alpha$. The map $D_{\widetilde{\alpha}_i}$ cuts each component of $\rho^{-1}(\varphi(\gamma_2))$ at its intersection with $\widetilde{\alpha}_i$, and reattaches by inserting copies of (a fundamental domain of) α. So, after the cutting, $\widetilde{\varphi(\gamma_2)}$ is attached to the lift of $\varphi(\gamma_2)$ through $\widetilde{v_2'}$ via a path in $\widetilde{B}$.

Thus, after performing $\widetilde{D}_\alpha$, we have a region $\widetilde{B}^\alpha$ such that all bigons on the vertices are exterior to the region, and contain no other vertices. We may therefore isotope

the $\widetilde{D}_\alpha \rho^{-1}(\varphi(\gamma_i))$ so as to eliminate bigons, and preserve $\widetilde{B}^\alpha$. As all isotopic bigon-free configurations are isotopic through bigon-free configurations, the result is thus equivalent to $\rho^{-1}\tau_\alpha(\varphi(\gamma_i))$, and so $\rho(\widetilde{B}^\alpha) = B^\alpha \in \mathcal{R}(\mathcal{P}^\alpha)$ is a region as desired. □

More generally, we would like to be able to find an image for *any* non-collapsed region with fixable $\circ$–points. We break the argument into two lemmas, for the cases of upward and non-upward regions.

Lemma 3.29 *Let $\mathcal{P} = \mathcal{P}(\Sigma, \varphi, \{\gamma_1, \gamma_2\})$ be a right position. Suppose $B \in \mathcal{R}(\mathcal{P})$ is a non-collapsed upward region with fixable $\circ$–points $y_i \in \gamma_i$, $\alpha \in \mathrm{SCC}(\Sigma)$, and $\hat{y}_i \in i_{[\alpha]}(y_i)$. Then $\hat{y}_1$ and $\hat{y}_2$ are the $\circ$–points of an image of B.*

Proof Let $\mathcal{P}'$ denote the maximal right position which both contains $\mathcal{P}$, and is such that $(\mathcal{P}')^\alpha = \mathcal{P}^\alpha$. We will show that $\mathcal{R}(\mathcal{P}')$ contains a preimage for i_α which maps to the desired region.

By assumption, at least one of the $\circ$–points, say y_1, is upward. We then may fix a nice α such that $y_1^\alpha = \hat{y}_1$ and $[y_1]_\alpha$ contains more than one element.

Let S' denote the set of regions of $\mathcal{R}(\mathcal{P}')$ whose orientations match that of B, and whose $\circ$–points are *simultaneously* fixable to the given $\hat{y}_1$ and $\hat{y}_2$, ie such that for some $\alpha' \in [\alpha]$, $i_{\alpha'}$ sends the $\circ$–points to $\hat{y}_1$ and $\hat{y}_2$. We will show firstly that S' contains a non-collapsed element.

Suppose then that $B \notin S'$, ie for $i = 1, 2$ the sets $\{\alpha \in [\alpha] \mid y_i^\alpha = \hat{y}_i\}$ are disjoint. Thus by Lemma 3.21 $[y_1]_\alpha$ has exactly two elements: y_1, and a point we label y_1'. Let T_{y_1} denote the upward triangular region on y_1. Then as y_2 is also fixable, there is an upward triangle T_{y_2} and a shift isotopy $h_t^{T_{y_2}}$ such that $y_2^{\alpha'} = \hat{y}_2$, where $\alpha' = h_1^{T_{y_2}}\alpha$ (see Figure 25(a)). As $B \notin S'$, this isotopy must cross y_1, so y_1 lies in the interior of T_{y_2}. Again by Lemma 3.21, the class $[y_2]_{\alpha'}$ also contains exactly 2 points: y_2, and a point we label y_2'. The points y_1' and y_2' thus lie in the exterior of T_{y_2}, so $y_i'^{\alpha'} = y_i'^{\alpha} = \hat{y}_i$ for either i. Now, consider a neighborhood of y_1 chosen small enough that its intersection with $\gamma_1 \cup \varphi(\gamma_2)$ is a pair of arcs intersecting at y_1, and let T_ϵ denote the intersection of this neighborhood with T_{y_1}. We have two cases:

(1) $T_\epsilon \not\subset B$ (see Figure 25(b)). Pushing to the universal cover, let $\widetilde{\varphi(\gamma_2)}'$ denote the lift of $\varphi(\gamma_2)$ through $\widetilde{y}_1'$; this arc enters $\widetilde{T}_{y_2}$ at its intersection with $\widetilde{\alpha}$, so continues to intersect $\widetilde{\gamma}_2$ at a point we denote $\widetilde{v}_2{}'$. Note then that using Lemma 3.25, $w' := \rho(\widetilde{v}_2{}')$ is in $\mathcal{P}_2'$. Letting B_1 denote the region obtained by extending B along the γ_i from v_2 and y_1 to v_2' and y_1', and isotoping α to α', we have $B_1 \in S'$ (see Figure 25(c)). As no edge is contained in the span of the equivalence class of its $\circ$–point, B_1 is not collapsed, as desired.

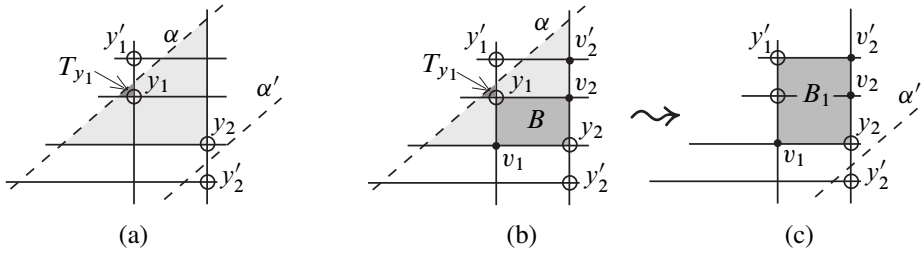


Figure 25: Figures for Lemma 3.29. The region T_{y_2} is lightly shaded in each.

(2) $T_\epsilon \subset B$ (see Figure 26(a)). As T_{y_1} does not contain y_2, it follows that $\alpha \cap B \neq \varnothing$. Note that if y_1' lies along neither e_1 nor e_3, then each of these is contained in the span of $[y_1]_\alpha$ (see Figure 26(b)). But then as α may be assumed nice, each of e_2 and e_4 is contained in the span of $[y_2]_\alpha$, so B is collapsed, a contradiction. Thus y_1' lies along e_1 or e_3 (see Figure 26(c) and (d)). Then, again using Lemma 3.25, y_1' is the $\circ$–point of a subregion B_1 in S' whose edge along $\varphi(\gamma_2)$ is not contained in the span of $[y_1']_\alpha$, so B_1 is not collapsed (of course if y_1' lies in $e_1 \cap e_3$ we have 2 such subregions, either of which serves our purpose). Furthermore, as y_1' does not lie in T_{y_2}, we may shift α over T_{y_2} to realize B_1 in S'.

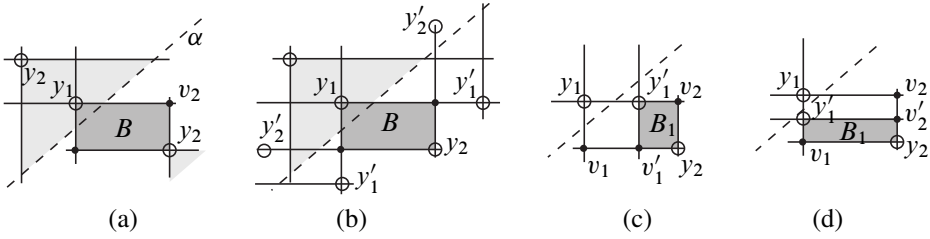


Figure 26: Figures for Lemma 3.29, case (2)

Finally, let $B_2 \in S'$ be minimal among non-collapsed regions, in the sense that it contains no other. Observe then that no element of $[y_i]_\alpha$, for either i, lies in the interior of an edge of B_2. Indeed, suppose otherwise, so that eg $y_1' \in [y_1]_\alpha$ is such that y_1' lies along e_1 (the other cases are essentially identical). Then y_1' is the endpoint of a segment of $\gamma_1 \cap B$ which cuts across B_2, and whose other endpoint lies in $\mathcal{P}'$, thus determining a subregion. Such a subregion will again be non-collapsed, contradicting minimality.

The region B_2 is thus a preimage whose $\circ$–points map to the given $\hat{y}_1$ and $\hat{y}_2$. By Lemma 3.28 these are then the $\circ$–points of an image of B as desired. $\square$

We require a similar statement for the non-upward case:

Lemma 3.30 *Let $\mathcal{P} = \mathcal{P}(\Sigma, \varphi, \{\gamma_1, \gamma_2\})$ be a right position. Suppose $B \in \mathcal{R}(\mathcal{P})$ is a non-upward region with $\circ$–points y_1 and y_2, $\alpha \in \mathrm{SCC}(\Sigma)$, and let $\hat{y}_i \in i_{[\alpha]}(y_i)$ for $i = 1, 2$. Then $\hat{y}_1$ and $\hat{y}_2$ are the $\circ$–points of an image of B.*

Proof Ideally, one would like to find $\alpha' \in [\alpha]$ such that B is a preimage for $i_{\alpha'}$ which maps to the desired region. Supposing for the moment that this is possible, we will break the isotopy $\alpha \rightsquigarrow \alpha'$ into a sequence of shifts, as follows: if $y_i^\alpha \neq \hat{y}_i$ for either i, there is a unique triangular region T_i such that $y_1^{h_1^{T_i}(\alpha)} = \hat{y}_i$ (see Figure 27 (a)). Assuming these shifts may be performed simultaneously, we obtain a curve α'' (see Figure 27(b)) which gives the correct $\circ$–point images. Now, for each $\bullet$–point, let T_3 (for v_1) and T_4 (for v_2) denote the maximal triangular region contained in B and with vertex the given $\bullet$–point. Isotoping over each of these, we then obtain α' with only vertical intersections with B, so B is a preimage with image $B^{\alpha'}$ satisfying the desired properties (see Figure 27(c)).

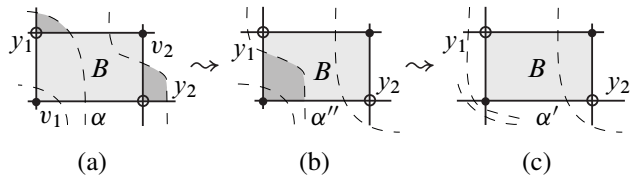


Figure 27: Shift-isotoping α so that B is the prescribed pre-image

Of course these shifts are not in general simultaneously realizable; we may however get around this by passing to the universal cover. To see this, let $\widetilde{B}$ be a lift of B to the universal cover, and $\widetilde{T}_i$, for each i, the lift of T_i (defined in the previous paragraph) with a vertex in common with $\widetilde{B}$. Of course we may always assume α has been isotoped over T_1. But then, as $\widetilde{T}_2$ has no interior intersection with $\widetilde{B}$, we may isotope $\rho^{-1}\alpha$ over $\widetilde{T}_2$ without crossing $\widetilde{y}_1$. Similarly, any downward region contained in $\widetilde{B}$ contains neither $\widetilde{y}_i$, while any pair of such regions is either disjoint, or such that one contains the other. In particular, the composition $h_t^{\widetilde{T}} := h_t^{\widetilde{T}_4} h_t^{\widetilde{T}_3} h_t^{\widetilde{T}_2}$ isotopes $\rho^{-1}\alpha$ to a representative $(\rho^{-1}\alpha)'$ with the desired properties.

We then define $D_{h_t^{\widetilde{T}}(\rho^{-1}\alpha)}\colon \widetilde{\Sigma} \to \widetilde{\Sigma}$ as the composition of the $D_{h_t^{\widetilde{T}}\widetilde{\alpha}_j} := h_t^{\widetilde{T}} \widetilde{D}_{\alpha_j} (h_t^{\widetilde{T}})^{-1}$ for each connected component $\widetilde{\alpha}_j$ of $\rho^{-1}\alpha$. The map $D_{h_1^{\widetilde{T}}(\rho^{-1}\alpha)}$ is thus isotopic to the lift $\widetilde{D}_\alpha$, and fixes each $\widetilde{y}_i$ to a lift of y_i^α (recall that intersection points of $\gamma_i \cap \varphi(\gamma_j)$ are only considered up to isotopy of $\varphi(\gamma_j)$ which involves no bigons). As all intersections with $\widetilde{B}$ are vertical, $\widetilde{B}$ is a preimage, and in particular the $\widetilde{y}_i^\alpha$ bound a region of

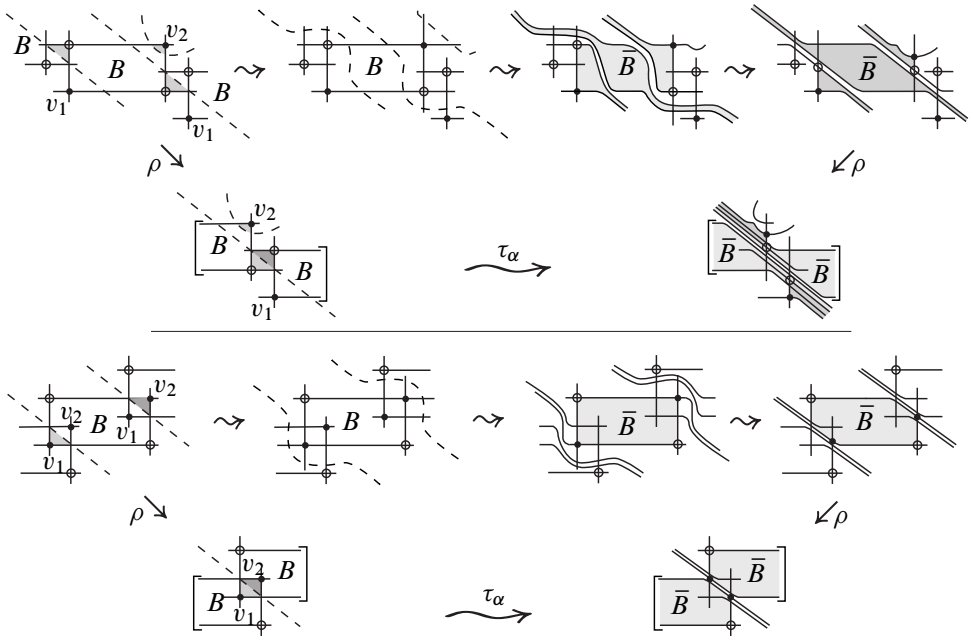


Figure 28: Examples of the steps outlined in Lemma 3.30. In each of the two figures, the lower left illustration is of some collection of triangular regions in Σ whose shifts cannot be done simultaneously, while each upper row is in the universal cover, and the squiggly arrows refer to, from left to right, the isotopy $(\rho^{-1}\alpha) \rightsquigarrow (\rho^{-1}\alpha)'$, the map $D_{h_t^{\tilde{T}}(\rho^{-1}\alpha)}$ applied to the $\rho^{-1}(\varphi(\gamma_j))$, and finally the isotopy of these images back to the $\rho^{-1}(\tau_\alpha(\varphi(\gamma_j)))$.

$\mathcal{R}(\rho^{-1}(\mathcal{P}^\alpha))$, which is then mapped by ρ to a region of $\mathcal{R}(\mathcal{P}^\alpha)$ as desired. As an illustration, Figure 28 goes through the steps for 2 particular configurations. □

We arrive finally at the proof of Theorem 1.1.

Proof of Theorem 1.1 Let γ_1 and γ_2 be disjoint properly embedded arcs in Σ, and $\varphi \in \mathrm{MCG}(\Sigma)$ admit a positive factorization $\tau_{\alpha_n} \cdots \tau_{\alpha_1}$. We will show by induction on n that if $\mathcal{P}^j$ denotes the right position for the augmented open book $(\Sigma, \tau_{\alpha_j} \cdots \tau_{\alpha_1}, \{\gamma_1, \gamma_2\})$ consisting of those points of $\gamma_k \cap \varphi(\gamma_k)$, for $k = 1, 2$, which are fixable under each successive τ_{α_i}, then $\mathcal{P}^j$ is consistent for each j.

For the base case $n = 0$, so $\varphi = \mathrm{Id}$, we observe that, isotoping each $\varphi(\gamma_k)$ to be to the right of γ_k at each endpoint and have no positive interior intersections (see Figure 29), $\mathcal{P}^0$ is given by the trivial position $\partial\Gamma$. As this has no regions, it is trivially consistent.

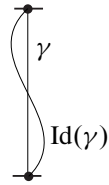


Figure 29

Suppose then that $\mathcal{P}^{j-1}$ is consistent, and let $B^\alpha \in \mathcal{R}(\mathcal{P}^j)$. Now, if B^α is created by α, then again by Lemma 3.24 it is completed, so we assume otherwise. Letting v_1^α and v_2^α denote the $\bullet$–points of B^α, by definition there are v_1 and v_2 in $\mathcal{P}^{j-1}$ such that $v_1^\alpha \in i_{[\alpha]}(v_1)$ and $v_2^\alpha \in i_{[\alpha]}(v_2)$. Let $\bar{y}_1, \bar{y}_1^\alpha, \bar{y}_2$ and $\bar{y}_2^\alpha$ denote the reflections of v_2, v_2^α, v_1 and v_1^α, respectively (see Figure 30(a)). We then have $\bar{y}_i \in i_{[\alpha]}(\bar{y}^\alpha)$, for $i = 1, 2$ (see Remark 3.22). Then, as $\bar{y}_1^\alpha$ and $\bar{y}_2^\alpha$ are the $\circ$–points of the reflection $\bar{B}^\alpha$ of B^α, Lemmas 3.29 (if $\bar{B}^\alpha$ is upward with respect to α) and 3.30 (otherwise) ensure the $\bar{y}_i$ are $\circ$–points for a region $\bar{B}$. The reflection of $\bar{B}$, which we denote B, is thus an element of $\mathcal{R}(\mathcal{P}^{j-1})$, and such that B^α is an image of B (see Figure 30(b)).

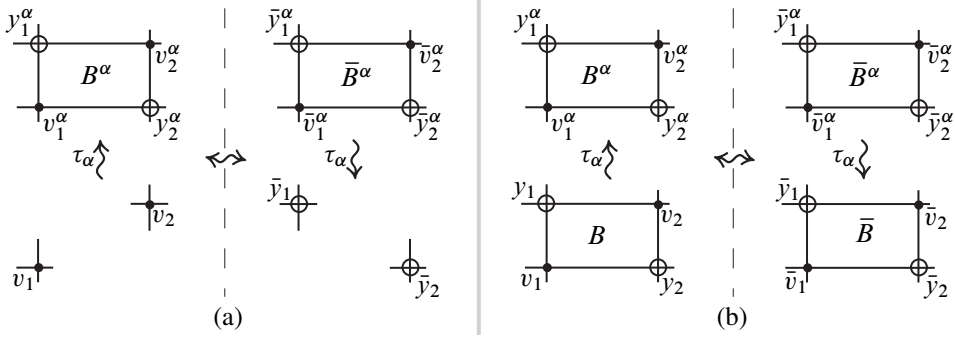


Figure 30

By Lemmas 3.29 and 3.30, any non-collapsed region with the opposite orientation of B and $\circ$–points in $[y_i]_\alpha$, for $i = 1, 2$, will have an image which completes B^α. Consider then the set S of regions of $\mathcal{R}(\mathcal{P}^{j-1})$ with $\circ$–points in $[y_i]_\alpha$, $i = 1, 2$, and the opposite orientation of B. We will show that S contains a non-collapsed element. Note firstly that by consistency, B itself has a completion B' which lies in S, so S is non-empty.

Now, we may order S by considering $\circ$–points along γ_1 in accordance with its standard orientation (ie pointing from y_1 to v_1); let B'_m denote a minimal element. Suppose then that B'_m is collapsed. Then by Lemma 3.24 there is a region B_{m+1} of opposite

orientation and the same $\bullet$–point set as B'_m, and whose $\circ$–points y_i^{m+1} lie in $[y_i]_\alpha$ (see Figure 31). In particular, B_{m+1} lies in $\mathcal{R}(\mathcal{P}^{j-1})$, and so is completed by some B'_{m+1} which lies in S, but is such that its $\circ$–point along γ_1 is further along γ_1 than that of B'_m, contradicting minimality. Thus B'_m is non-collapsed, and so has an image which completes B^α. □

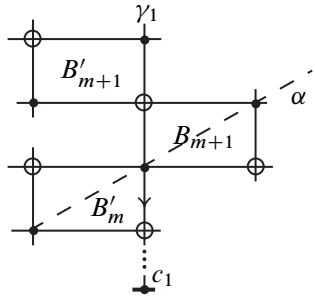


Figure 31

4 Restrictions on $pe(\varphi)$

In this section, we switch focus back to the entirety of a pair of distinct properly embedded arcs γ_1, γ_2 in a surface Σ, and the images of these arcs under a right-veering diffeomorphism φ. The motivating observation here is that, as the endpoints of each arc are by definition included in any right position, the property of the initial horizontal segments of the images determining an initially parallel region is independent of right position, and so is a property of the pair $\varphi(\gamma_1), \varphi(\gamma_2)$. In particular, for *positive* φ, if $\varphi(\gamma_1)$ and $\varphi(\gamma_2)$ are initially parallel, they must admit a consistent right position in which the initially parallel region is completed (see Example 3.7).

Definition 4.1 For a surface Σ and positive mapping class φ, we define the *positive extension* $pe(\varphi) \subset \mathrm{SCC}(\Sigma)$ as the set of all $\alpha \in \mathrm{SCC}(\Sigma)$ such that τ_α appears in some positive factorization of φ.

We are interested in understanding what necessary conditions on $\alpha \in pe(\varphi)$ can be derived from the information that $\varphi(\gamma_1)$ and $\varphi(\gamma_2)$ are initially parallel.

4.1 Restrictions from initial horizontal segments

Throughout the section, we will be considering rectangular regions R in Σ, and classifying various curves and arcs by their intersection with such regions. We need:

Definition 4.2 Let R be an embedded rectangular region with distinguished oriented edge e_1, which we call the *base* of R. We label the remaining edges e_2, e_3, e_4 in order, using the orientation on e_1 (see Figure 32). We classify properly embedded (unoriented) arcs on R by the indices of the edges corresponding to their boundary points, so that an arc with endpoints on e_i and e_j is of *type* $[i, j]$ (on R). We are only concerned with arcs whose endpoints are not on a single edge. An arc on R is then

- *horizontal* if of type $[2, 4]$,
- *vertical* if of type $[1, 3]$,
- *upward* if of type $[1, 2]$ or $[3, 4]$,
- *downward* if of type $[1, 4]$ or $[2, 3]$,
- *non-diagonal* if horizontal or vertical.

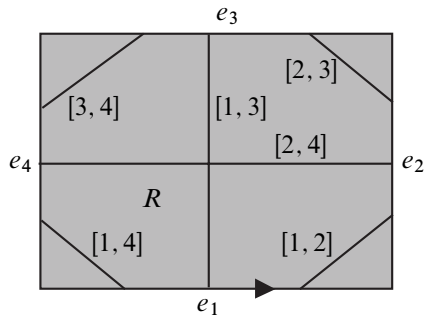


Figure 32: Representatives of each possible type of arc on R

Definition 4.3 Let R in Σ be a rectangular region with distinguished base as in Definition 4.2. We say that $\alpha \in \mathrm{SCC}(\Sigma)$ is *not downward on* R if each arc $\alpha \cap R$ is not downward on R.

Remark 4.4 Following Definition 4.2, we use a given pair of properly embedded arcs to define a rectangular region D of Σ as follows. Suppose γ_1 and γ_2 are disjoint properly embedded arcs, where $\partial\gamma_i = \{c_i, c_i'\}$, and $\bar{\gamma}_1$ is a given properly embedded arc with endpoints c_1' and c_2. Let $\bar{\gamma}_2$ be a parallel copy of $\bar{\gamma}_1$ with endpoints isotoped along γ_1 and γ_2 to c_1 and c_2'. We then define D to be the rectangular region bounded by $\gamma_1, \gamma_2, \bar{\gamma}_1, \bar{\gamma}_2$ (the endpoints are labeled so that c_1 and c_2 are diagonally opposite), and base $\bar{\gamma}_2$, oriented away from c_1 (see Figure 33). For the remainder of this section, D will always refer to this construction. Of course, for a given pair of arcs, D is unique only up to the choice of $\bar{\gamma}_1$.

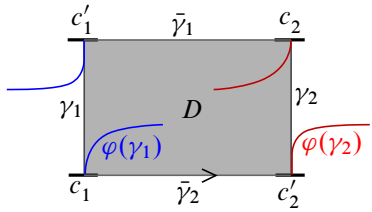


Figure 33: Disc construction

Definition 4.5 We say the pair $\varphi(\gamma_1), \varphi(\gamma_2)$ is *initially parallel (on D)* if there exists D such that $\varphi(\gamma_1), \varphi(\gamma_2), \gamma_1$ and γ_2 bound a rectangular region $B \hookrightarrow D$ on which c_1 and c_2 are vertices. An initially parallel (on D) pair of images is *flat* (on D) if $\varphi(\gamma_i) \cap \bar{\gamma}_j = \varnothing$ for $i, j \in \{1, 2\}$.

As expected, then, the pair of arcs $\varphi(\gamma_1), \varphi(\gamma_2)$ is initially parallel exactly when the horizontal segments h_{c_1} and h_{c_2} originating from the boundary of each pair of right positions determine an initially parallel region. Using this, we immediately obtain:

Lemma 4.6 *Let φ be positive, D be as in Remark 4.4 and the pair $\varphi(\gamma_1), \varphi(\gamma_2)$ be initially parallel on D. Then $\alpha \in pe(\varphi)$ only if α is not downward on D.*

Proof Suppose otherwise. Then $\alpha \cap D$ has an arc of type $[1, 4]$ or $[2, 3]$ on D. Now, if $\alpha \in pe(\varphi)$, there is some positive factorization of φ in which τ_α is the initial Dehn twist. However, if $\alpha \cap D$ has an arc of type $[1, 4]$ (resp. $[2, 3]$), then the initial segment of $\tau_\alpha(\gamma_1)$ (resp. $\tau_\alpha(\gamma_2)$) is strictly to the right of $\varphi(\gamma_1)$ (resp. $\varphi(\gamma_2)$), a contradiction. □

4.1.1 Motivating examples We begin with a pair of examples, meant to motivate the various definitions to come.

Example 4.7 Consider a positive $\varphi \in \mathrm{MCG}(\Sigma)$, and a pair of properly embedded arcs γ_1 and γ_2 such that the images $\varphi(\gamma_1)$ and $\varphi(\gamma_2)$ are flat in some disc D (see Figure 34(a)). The segments h_{c_1} and h_{c_2} are of course initially parallel (along some region B), so there is some completing region (it is easiest to keep in mind the simplest case in which the $\bullet$–points of this completing region are the opposite endpoints of γ_1 and γ_2, as in the figure). Let α be a simple closed curve such that $\alpha \cap D$ has exactly two diagonal arcs, each upward, and connected such that an orientation of α gives each of these arcs the same orientation. The arcs $\tau_\alpha^{-1}(\varphi(\gamma_i))$ are thus again initially parallel, with $\circ$–points which map to those of B under i_α. One may then verify that there is a unique completing region, which is collapsed by α. Lemma 3.24 then guarantees that the $\bullet$–points of this completing region define initially parallel segments $h_{v'}$ and $h_{w'}$.

This new initially parallel region is then completed by (an image under i_α^{-1} of) the original completing region for B. Furthermore, the $\circ$–points of the regions are each fixable. The regions and points involved are illustrated in Figure 34(b).

Consider then a slightly more complicated configuration, in which α has 4 diagonal arcs, arranged as in Figure 34(c). Again, one may verify that any consistent right positions for $((\tau_\alpha^{-1} \circ \varphi), \gamma_i)$ (see Figure 34(d)) must contain one of the two positions indicated in Figure 34(e) and (f). Again, each of these determines a sequence of (collapsed) regions and points which ends with the original completions. The goal of the remainder of this section is to show that for *any* simple closed curve α, any pair of consistent right positions for $((\tau_\alpha^{-1} \circ \varphi), \gamma_i)$ must contain a similar sequence, which then determines necessary conditions on $\alpha \in pe(\varphi)$.

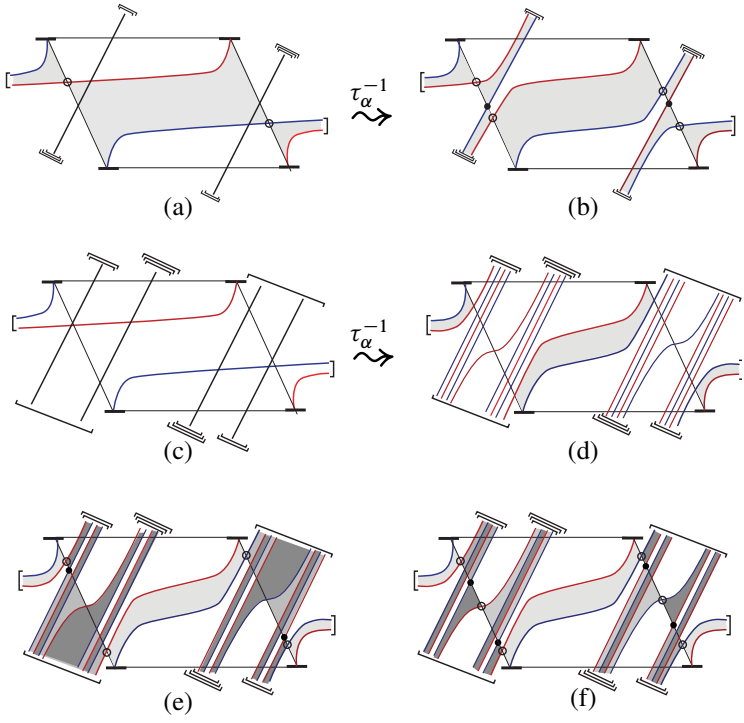


Figure 34

4.1.2 Notation and conventions Given a pair of arc images $\varphi(\gamma_1)$ and $\varphi(\gamma_2)$ flat on some D, and a non-downward curve α, let $B^\alpha \hookrightarrow D$ denote the initially parallel region with $\bullet$–points c_1 and c_2, and $y_i^\alpha \in \gamma_i$, $i = 1, 2$, the $\circ$–points. Consider the set of upward triangles of $(\alpha, \gamma_1, \varphi(\gamma_2))$ with vertex y_1^α (see Figure 35(a)); each of course has a distinct vertex in $\gamma_1 \cap \alpha$, which we label $x_1, x_3, \dots, x_{m_1}$ so that indices

decrease along γ_i in the direction of the orientation of γ_1. Similarly we have upward triangles with vertex y_2^α, and vertices $x_2, x_4, \dots, x_{m_2}$ along γ_2 (see Figure 35(b)). Let $a_1, a_2, \dots a_m$, where $a_1 = 1$, be the indices of the x_i with order given by the order in which they are encountered traveling along α to the right from x_1. We associate to α the list $\pi_\alpha = (a_1, a_2, \dots a_m)$, which for reasons that will become apparent in the following subsection we define only up to cyclic permutation. We decorate an entry with a bar, $\bar{a}_j$, if for odd (resp. even) j, the orientation of the intersection point x_j agrees (resp. disagrees) with that of x_1.

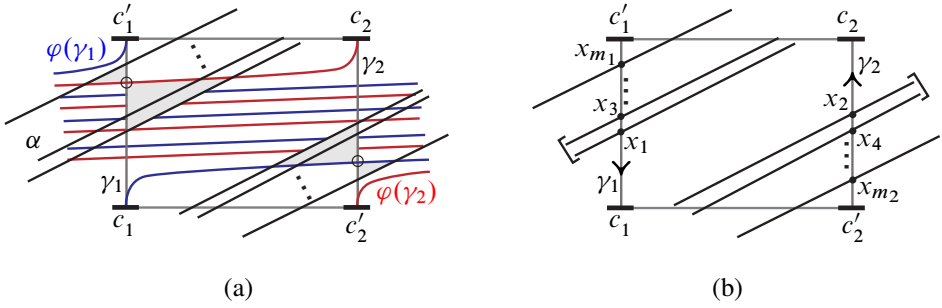


Figure 35: (a) Triangular regions on the $\circ$–points of B^α. (b) Labeling of arcs and intersections. The paths $\eta_{2,3}$ and $\eta_{4,1}$ are components of a symmetric multipath.

Given a pair $x_i \in \alpha \cap \gamma_2$ and $x_j \in \alpha \cap \gamma_1$ (so i is even, j odd), the pair splits α into two disjoint arcs. If the orientations of x_i and x_j disagree, we label these arcs $\eta_{i,j}$ and $\eta_{j,i}$ (where for even i, $\eta_{i,j}$ refers to the arc connecting x_i and x_j and initially *exterior* to D at each endpoint, so for odd j, a neighborhood of each endpoint of $\eta_{j,i}$ is contained in D). We refer to each such arc as a *path* of the pair (α, D). We say paths η and η' are *parallel* if, after sliding endpoints along the γ_i so as to coincide, they are isotopic as arcs (relative to their endpoints). Then a *multipath* of (α, D) is a collection of more than one pairwise parallel path.

Of particular interest will be the following special case:

Definition 4.8 A multipath as described in the previous paragraph is *symmetric* if its components are $\eta_{a_1,b_1}, \eta_{a_2,b_2}, \dots, \eta_{a_r,b_r}$, where each pair $\{a_i, b_{r+1-i}\}$ is $\{j, j-1\}$ for some even j. So, for example, any symmetric *path* is $\eta_{j,j-1}$ or $\eta_{j-1,j}$ for some even j. Figure 35(b) contains an example of a symmetric multipath with components $\eta_{2,3}$ and $\eta_{4,1}$.

4.2 Necessary conditions for $\alpha \in pe(\varphi)$

This subsection is devoted to introducing the terminology of, and proving, Theorem 1.2.

Definition 4.9 Let $\alpha \in \mathrm{SCC}(\Sigma)$ be not downward on a region D. We call α *balanced* (with respect to D) if each x_i is an endpoint in either a symmetric path or a component of a symmetric multipath. We say a balanced α is *nested* with respect to D if its associated π_α can be reduced to the empty word by successive removals of pairs (i, j) of consecutive entries, where $\eta_{i,j}$ is a symmetric path or component of a symmetric multipath.

Note that in the absence of symmetric multipaths of more than one component, these properties are encoded in π_α. For example, $(12\bar{3}4)$ is not balanced, (135462) is balanced but not nested, and $(1\bar{4}56\bar{3}2)$ is nested. On the other hand, $(12\bar{4}\bar{5}\bar{7}\bar{8}\bar{6}\bar{3})$ is not balanced in the absence of symmetric multipaths, but nested if $\eta_{4,5}$ and $\eta_{6,3}$ are components of a symmetric multipath. Also note that each condition requires the number of intersections $|\alpha \cap \gamma_1|$ to equal $|\alpha \cap \gamma_2|$, ie $m_1 + 1 = m_2$.

We want to characterize the nested condition in terms of paths of α. We need:

Definition 4.10 Let $\eta_{a,b}, \eta_{c,d} \subset \alpha$ be paths, with a, b, c and d distinct. We refer to the pair as *nested* if either one contains the other, or they are disjoint. Otherwise, the pair is *non-nested*.

Observe then that if for some balanced α, each pair of components of symmetric multipaths is nested, then there is some "innermost" multipath whose paths contain no other. In particular the indices of the endpoints of these paths are consecutive pairs in π_α. We have:

Lemma 4.11 *Let balanced α be such that each pair of components of symmetric multipaths is nested. Then α is nested.* □

To keep track of intersection points $\gamma_i \cap \tau_\alpha^{-1}(\varphi(\gamma_j))$, for $i, j \in \{1, 2\}$, we begin with the observation that each $\circ$–point of the initially parallel region with $\bullet$–points c_1 and c_2 is fixable under τ_α^{-1}. As in the previous subsection we label these $y_1^\alpha \in \gamma_1 \cap \varphi(\gamma_2)$ and $y_2^\alpha \in \gamma_2 \cap \varphi(\gamma_1)$. There are then exactly $(m_1 + 3)/2$ points to which y_1^α may be fixed by i_α^{-1}; we label these $y_1, y_3, \dots, y_{m_1+1}$ so that the indices increase along γ_1 from c_1. Similarly, there are $(m_2 + 2)/2$ possible fixed preimages of y_2, which we label $y_2, y_4, \dots, y_{m_2+2}$ along γ_2 from c_2.

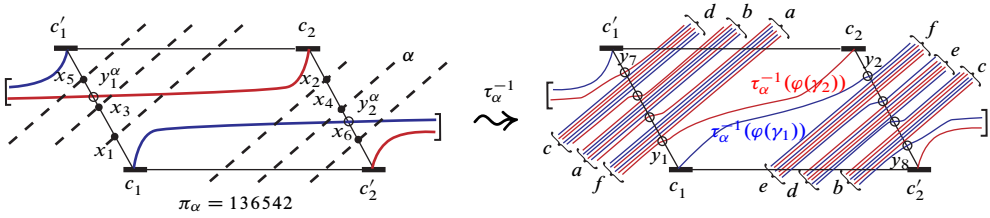


Figure 36: Above left, the various points of $\alpha \cap \gamma_i$ and $\alpha \cap \varphi(\gamma_i)$. Above right, the inverse image $\tau_\alpha^{-1}(\varphi(\gamma_i))$, for α with $\pi_\alpha = 136542$ (strands terminating in brackets with like letters are meant to be identified).

4.2.1 The right position $\mathcal{P}^*$ Our immediate goal is to show that, if the $\tau_\alpha^{-1}(\varphi(\gamma_i))$, $i = 1, 2$, admit a consistent right position, then any such right position has a particular subposition whose regions extend non-trivially along the entire length of each arc, as described in Example 4.7. Theorem 1.2 will then follow from various properties of this right position.

Lemma 4.12 *Let $\varphi(\gamma_1)$ and $\varphi(\gamma_2)$ be flat with respect to some D, $\alpha \in pe(\varphi)$, and $\mathcal{P}$ be a consistent right position of $(\Sigma, \tau_\alpha^{-1} \circ \varphi, \{\gamma_i\})$. Then there is a right position $\mathcal{P}^* = (\{v_j\}_{j=1}^n \cup \{w_j\}_{j=1}^n) \subset \mathcal{P}$, and $\{e_k\}_{k=1}^{2n} \subset \mathbb{N}$ such that:*

(1) *$v_1 = c_1$ and $v_n = c_1'$, and $w_1 = c_2$ and $w_n = c_2'$, and the indices $1, \dots, n$ increase along each of γ_i and $\tau_\alpha^{-1}(\varphi(\gamma_i))$ from the endpoint c_i.*

(2) *For each $j < n$, segments $h_{v_j, v_{j+1}}$ and $h_{w_j, w_{j+1}}$ are initially parallel along a region B_j, and completed along a region B_j' with $\bullet$–points v_{j+1} and w_{j+1}. Furthermore, the set of $\circ$–points for these regions are fixable, and the $\circ$–points of the completing pair B_j, B_j' are $y_{e_{2j-1}}$ and $y_{e_{2j}}$ (in the set $\{y_k\}$ of fixable points described above).*

See Figure 37 for a schematic illustration.

Proof By assumption, h_{c_1} and h_{c_2} are initially parallel along a region B_1, with fixable $\circ$–points y_1 and y_2, which for consistency must be completed by some region B_1' with $\bullet$–points which we label v_2 and w_2. We need then to show that the horizontal segments $h_{v_2, c_1'}$ and $h_{w_2, c_2'}$ are themselves initially parallel along a region with $\circ$–points again in the fixable point set $\{y_k\}$. The result will then follow, as we may repeat the construction until we hit one of the opposite endpoints.

Now, each of the sets $\{y_1, y_3, \dots, y_{m_1+2}\}$ and $\{y_2, y_4, \dots, y_{m_2+2}\}$ is of course just the α–equivalence class (Definition 3.19) of each of its elements. In particular, each

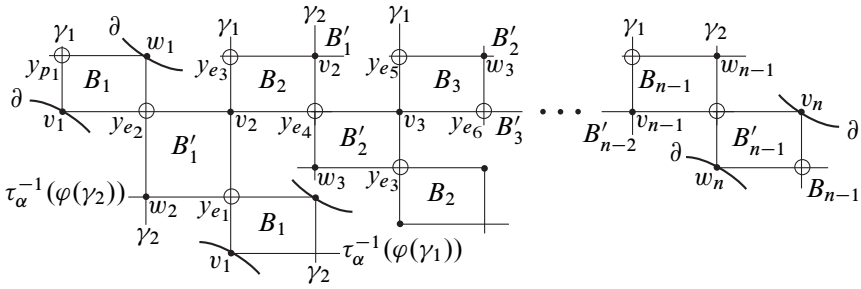


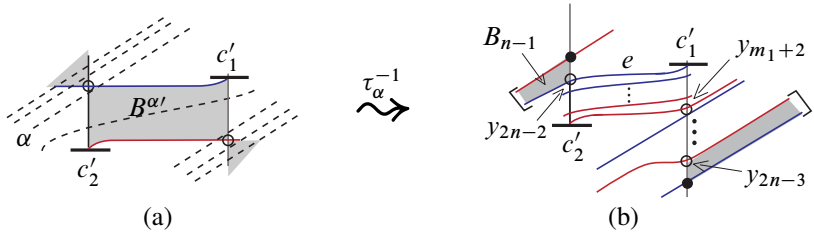
Figure 37: A schematic overview of the position $\mathcal{P}^*$ constructed in Lemma 4.12. The picture is of the universal cover of Σ, along a lift of $\tau_\alpha^{-1}(\varphi(\gamma_1))$. As usual, all lifts of an object are given the same label as the original object.

edge of B_1' is contained in the span of the α–equivalence class of its $\circ$–point; as α may be assumed to intersect B_1' non-trivially, B_1' is collapsed by α, and so by Lemma 3.24 there is a complementary $(B_1')^c$ with $\circ$–points in $[y_i]_\alpha$, as desired.

Continuing in this fashion, for each j, $B_j = (B_{j-1}')^c$, we obtain a sequence of points $\{v_j\}$ and $\{w_j\}$ satisfying condition (2), which continues until we hit an endpoint, ie (assuming $m_1 \geq m_2 - 1$), $w_n = c_2'$. Condition (1) will then be satisfied if $m_1 = m_2 - 1$, so that the corresponding v_n is the other endpoint c_1'.

Suppose otherwise; we will show that B_{n-1} has no completion, a contradiction. As elsewhere in the paper, we will use the notation $[a,b]_\sigma$ to denote the segment of an arc σ between points a and b. Let $B^{\alpha\prime}$ denote the completion of B^α, and consider firstly the case that $\alpha \cap B^{\alpha\prime}$ has no vertical arcs (ie arcs connecting $\varphi(\gamma_1)$ and $\varphi(\gamma_2)$, see Figure 38(a)). Then, by construction, y_{2n-2} is such that $e := [y_{2n-2}, c_1']_{\tau_\alpha^{-1}(\varphi(\gamma_1))}$ has no interior intersection with $[y_1, c_1']_{\gamma_1}$ (see Figure 38(b)). As such, any completion of B_{n-1} has e as an edge, and c_1' as a $\bullet$–point. Thus $[y_{2n-3}, c_1']_{\gamma_1}$ is another edge. However, it is clear that, for instance, the intersection number of this edge with $\tau_\alpha^{-1}(\varphi(\gamma_1))$ is greater than that of any candidate edge along γ_2, so we have no possible region. The remaining case, that $\alpha \cap B^{\alpha\prime}$ has vertical arcs, is almost identical: we now take e to be $[y_{2n-2}, c_2']_{\gamma_2}$, which now has no interior intersection with either $\tau_\alpha^{-1}(\varphi(\gamma_i))$, so is an edge of any completion. Thus $[y_{2n-3}, c_2']_{\tau_\alpha^{-1}(\varphi(\gamma_2))}$ is another edge, and again considering intersections of $\tau_\alpha^{-1}(\varphi(\gamma_1))$ with the edges along the γ_i we conclude that no region exists. $\square$

We have the following immediate corollary, giving us "half" of Theorem 1.2:



(a) (b)

Figure 38

Corollary 4.13 *Let $\varphi(\gamma_1)$ and $\varphi(\gamma_2)$ be flat with respect to some D. Then $\alpha \in pe(\varphi)$ only if α is balanced with respect to D.* □

4.2.2 Nestedness The proof of the remainder of Theorem 1.2, that *nestedness* is a further necessary condition for consistency, in fact also follows easily from the previous lemma, but keeping track of everything involved requires quite a bit of notation. We begin by using Lemma 4.12 to characterize nestedness in terms of regions and •–points.

As in the previous section, we denote by $\mathcal{R}(\mathcal{P})$ the set of initially parallel/completed regions in a right position $\mathcal{P}$. Recall from Lemma 3.24 that a collapsed region of length $l > 1$ and width $w = 1$ (resp. length $l = 1$ and width $w > 1$) contains $l - 1$ (resp. $w - 1$) collapsed subregions each of $l = w = 1$.

Definition 4.14 Given a pair of complementary paths $\eta_{a,b}$ and $\eta_{b,a}$, where a is *even*, let $B_{a,b}$ and $B_{b,a}$ be the collapsed pair of rectangular regions bounded by $[y_a, y_{a+2}]_{\gamma_2}$, $[y_a, y_{a+2}]_{\tau_\alpha^{-1}(\varphi(\gamma_2))}$, $[y_b, y_{b+2}]_{\gamma_1}$, and $[y_b, y_{b+2}]_{\tau_\alpha^{-1}(\varphi(\gamma_1))}$, where $B_{a,b}$ has vertices y_a and y_b, and $B_{b,a}$ has vertices y_{a+2} and y_{b+2} (see Figure 39(a)). Thus $B_{a,b}$ and $B_{b,a}$ are collapsed regions satisfying length = width = 1. We refer to each such region as *basic*. We refer to a pair of basic regions $B_{a,b}$ and $B_{c,d}$ as *nested* or *non-nested* if the associated paths $\eta_{a,b}$ and $\eta_{c,d}$ are nested or non-nested. From the definitions, we see that if $B_{a,b}$ and $B_{c,d}$ are nested, they are either disjoint, or intersect such that no vertex is in the intersection, ie in "strips" as in Figure 39(b). Conversely, if the regions are non-nested, they intersect in a "corner overlap" as in Figure 39(c). Finally, we call a region $B \in \mathcal{R}(\mathcal{P}^*)$ *nested* if each basic subregion of B is nested with each basic subregion of all other regions of $\mathcal{R}(\mathcal{P}^*)$.

We can thus characterize nestedness in terms of regions as follows:

Observation 4.15 A balanced curve α is nested if for any two regions $A, B \in \mathcal{R}(\mathcal{P}^*)$, any pair consisting of a basic region of A and a basic region of B is nested. Equivalently, if each $B \in \mathcal{R}(\mathcal{P}^*)$ is nested.

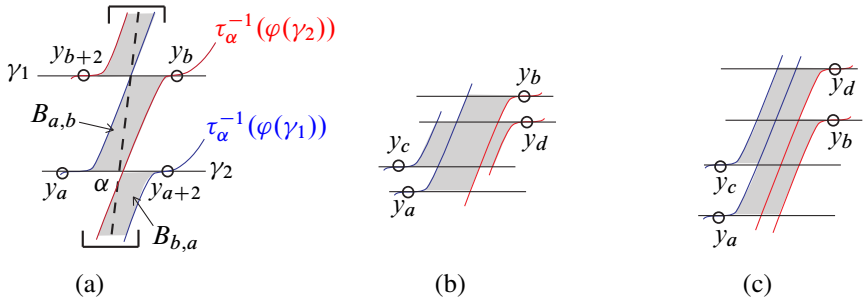


Figure 39: (a) The region $B_{a,b}$ corresponding to the path $\eta_{a,b}$, and complementary $B_{b,a}$ corresponding to $\eta_{b,a}$. The curve α is the dashed line. (b) Nested, and (c) non-nested regions.

We need:

Definition 4.16 Given right positions $\mathcal{P}_1$ and $\mathcal{P}_2$, we call a region $B \in \mathcal{R}(\mathcal{P}_1)$ *empty* (in $\mathcal{P}_2$) if neither the interior of B nor the interior of the edges which form ∂B contain any points of $\mathcal{P}_2$. Conversely, any point y which is a $\circ$–point of some region in $\mathcal{R}(\mathcal{P}_1)$ is *isolated* (in $\mathcal{P}_2$) if, for each $B \in \mathcal{R}(\mathcal{P}_2)$, y is in neither the interior of B nor the interior of the edges which form ∂B.

Then:

Lemma 4.17 *Let $\mathcal{P}^*$ be the right position constructed in Lemma 4.12. Then if $B \in \mathcal{R}(\mathcal{P}^*)$ is empty (in $\mathcal{P}^*$) with isolated $\circ$–points, then B is nested.*

Proof Suppose otherwise, so B is empty with isolated $\circ$–points, and there is an $A \in \mathcal{R}(\mathcal{P}^*)$ such that some basic region $B_{a,b}$ of B and some basic region $B_{c,d}$ of A are non-nested. Now, as in the proof of Lemma 3.24, any collapsed region has either width or length equal to 1; the cases are essentially identical, so we assume $l(B) = 1$. Now, if $l(A) \neq 1$, note that $B_{a,b}$ and $B_{d,c} = (B_{c,d})^c$ are also non-nested (see Figure 40(a)), so we may replace A with its complementary region A^c, which then satisfies $l(A^c) = 1$. It is then trivial to verify that if two collapsed regions each of length one (or each of width one) have non-nested basic regions, then each of the pair either contains a $\bullet$–point of, or has a $\circ$–point contained in, the other (see Figure 40(b)). We therefore arrive at a contradiction. $\square$

Lemma 4.18 *Let $\mathcal{P}^*$ be as in Lemma 4.12. Then if $B \in \mathcal{R}(\mathcal{P}^*)$ is empty (in $(\mathcal{P}^*)$) with isolated $\circ$–points, then the same holds for B^c.*

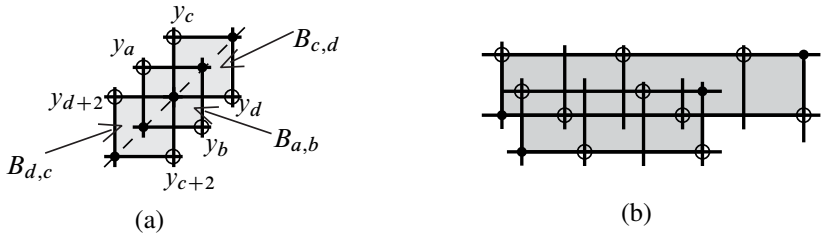


Figure 40: (a) If a basic region is non-nested, so is its complementary region. (b) Two collapsed, non-nested regions of length 1. Each is either non-empty, or has non-isolated $\circ$–points.

Proof Suppose that B^c is not empty, so contains some $\bullet$–point v of some region $A \in \mathcal{R}(\mathcal{P}^*)$. Now, the complement in B of the basic regions of B is the same as the complement in B^c of the basic regions of B^c, so, as B is empty, v lies in a basic region of B^c (see Figure 41 (a)). But then the basic subregion of A with vertex v is non-nested with a basic subregion of B, contradicting the previous lemma.

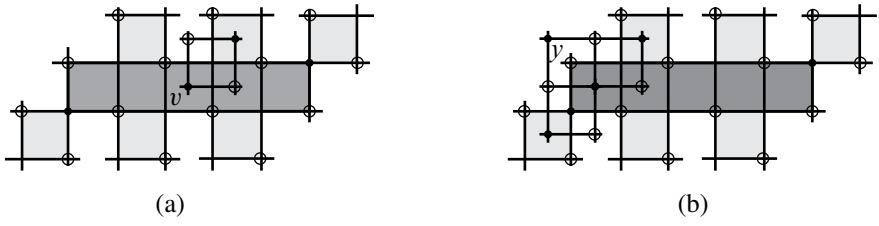


Figure 41: Figures for Lemma 4.18. The region B^c is darkly shaded, the basic regions of B lightly.

Similarly, if a $\circ$–point y of B^c is not isolated, it is contained in some A (see Figure 41(b)). Then, as by construction no basic region contains the $\circ$–point of any other, y must lie in the complement of the basic regions of A, ie in $A \cap A^c$. But then A and B again contain non-nested basic regions, a contradiction. □

Now, to connect $\mathcal{P}^*$ with arbitrary consistent $\mathcal{P}$, we require:

Lemma 4.19 *Let $\gamma_i, \varphi, \mathcal{P}, \mathcal{P}^*$, $i = 1, 2$, be as in Lemma 4.12. Suppose region $B_j \in \mathcal{R}(\mathcal{P}^*)$ is empty in $\mathcal{P}$, and the $\circ$–points $y_{e_{2j-1}}$ and $y_{e_{2j}}$ are isolated in $\mathcal{P}$. Then the completing region B'_j is also empty in $\mathcal{P}$.*

Proof Suppose that B_j is empty, while B'_j is not, so there is some $v \in \mathcal{P}$ such that $v \in B'_j$. Then v is the $\bullet$–point of an initially parallel region $B \subset B'_j$ whose remaining

$\bullet$–point is a $\bullet$–point of B'_j (see Figure 42, where we have illustrated the case $i = 1$). As B is a region of $\mathcal{R}(\mathcal{P})$, it has a completion B'. Then, letting v' be the $\bullet$–point of B' along γ_1, we see that, as the boundary of B' is assumed to contain neither of the pair $y_{e_{2j-1}}, y_{e_{2j}}$, then v' is in the interior of two distinct edges of B'_j. We thus have a contradiction, as each region of $\mathcal{R}(\mathcal{P}^*)$ is embedded by construction. $\square$

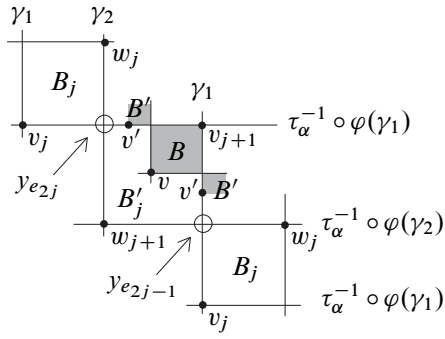


Figure 42: Figure for Lemma 4.19.

Observe then that, referring back to the sequence of regions created in Lemma 4.12, if B_1 is empty, Lemma 4.19 implies that B'_1 is empty, and so by Observation 4.15 it follows that the collection of entries of π_α up to the largest index of a basic region of B'_1 (ie $\{1, 2, \dots, e_3 - 1\}$) appear in π_α as a nested (not necessarily contiguous) subword. Our strategy is then to proceed by modifying the surface Σ via a sequence of punctures, which will have the effect of removing all regions $\mathcal{R}(\mathcal{P}) \setminus \mathcal{R}(\mathcal{P}^*)$, and isolating each of the fixed points y_{e_j}. This will allow us to generalize Lemma 4.17 to apply to each region of $\mathcal{R}(\mathcal{P}^*)$, and then show that the "overlapping regions" which characterize the non-nested case are an obstruction to consistency.

We begin by defining our sequence of modifications. Given $\mathcal{P}^*$ as in Lemma 4.12, let Σ^j be the surface given by deleting a pair of points from a neighborhood of each of the $\circ$–points $\{y_{e_k} \mid e_k \leq 2j\}$ as in Figure 43. To be precise, we choose a neighborhood about each of these points such that the restriction of $\{\gamma_1, \gamma_2, \tau_\alpha^{-1}(\varphi(\gamma_1)), \tau_\alpha^{-1}(\varphi(\gamma_2))\}$ to this neighborhood is a pair of arcs with a single intersection point, so these arcs divide the neighborhood into 4 "quadrants", two of which are filled by the pair of regions for which the given $\circ$–point is a vertex. We then delete a neighborhood of a point in the interior of each of the remaining two quadrants (see Figure 43). Note that this has the effect of removing from $\mathcal{R}(\mathcal{P})$ each region which contains either of these points, so that each $\circ$–point is isolated in the new region set.

Observe that our arcs γ_i as well as their images are preserved by this construction, so may be viewed as lying on each of the Σ^j, and similarly the mapping class $\tau_\alpha^{-1} \circ \varphi$ can be viewed as an element of each Γ_{Σ^j}.

We then inductively define a sequence of right-positions $\mathcal{P}^j$ for the augmented open book $(\Sigma, \tau_\alpha^{-1} \circ \varphi, \{\gamma_i\})$ for each surface Σ^j, where $\mathcal{P}^0 = \mathcal{P}$, and $\Sigma^0 = \Sigma$, and for $j > 0$, each $\mathcal{P}^j$ is obtained from $\mathcal{P}^{j-1}$ by removing any point which lies either in the interior of B_j or in the interior of its boundary.

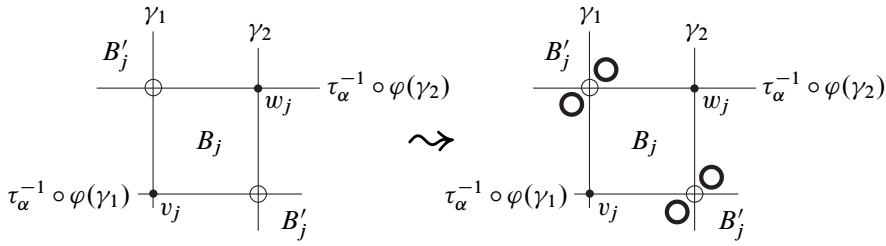


Figure 43: Modification $\Sigma^{j-1} \rightsquigarrow \Sigma^j$. The thick circles indicate boundary components introduced by the modification.

Finally:

Lemma 4.20 *The above modification process is such that, for each Σ^j, $\mathcal{P}^j$ is consistent, and furthermore $\mathcal{R}(\mathcal{P}^*) \subset \mathcal{R}(\mathcal{P}^j)$.*

Proof The proof is by induction on j. The base step, that on Σ the original positions $\mathcal{P}_i$ satisfy the conditions, is trivial. Suppose then that the conditions hold for Σ^{j-1} and $\mathcal{P}^{j-1}$. Observe that, as the deletion process ensures that for each $k < j$, B_k is empty, and each pair $y_{e_{2k-3}}, y_{e_{2k-2}}$ isolated, it follows from Lemma 4.19 that each region $B_1, B'_1, B_2, \dots, B_{j-1}, B'_{j-1}$ is empty in $(\mathcal{P}^{j-1})$. Consider then Σ^j and $\mathcal{P}^j$.

We first show that $\mathcal{P}^j$ is consistent. There is an obvious inclusion of $\mathcal{R}(\mathcal{P}^j)$ in $\mathcal{R}(\mathcal{P}^{j-1})$, so we must show that if $B \in \mathcal{R}(\mathcal{P}^{j-1})$ does not survive the deletion process (ie if B has an $\bullet$–point in B_j or contains one of $y_{e_{2j-1}}, y_{e_{2j}}$), then the same is true for any completing region B'. Figure 44 indicates the setup. Observe that B cannot contain B_j; this is of course obvious for $j = 1$, while for $j > 1$ follows easily, as each such region either has an $\bullet$–point in B'_{j-1} or contains one of the points $\{y_{e_k} \mid k \leq 2j-2\}$. Suppose then that B is a region with $\bullet$–point v in B_j (see Figure 44(a)). Then, as B'_{j-1} is empty (for $j > 1$), and the pair $y_{e_{2j-3}}, y_{e_{2j-2}}$ isolated, B has a $\circ$–point in B_j. But then any completing region B' either contains one of the points $\{y_{e_{2j-1}}, y_{e_{2j}}\}$, or has an $\bullet$–point in B_j. In either case B' does not survive. If, on the other hand, B has no $\bullet$–point in B_j, then B must contain one of the points $\{y_{e_{2j-1}}, y_{e_{2j}}\}$ (see Figure 44(b)), and again have a $\circ$–point in B_j, so any completing B' again does not survive.

It is left then to show that no region of $\mathcal{R}(\mathcal{P}^*)$ is removed through the deletion process. For the initial deletion, this follows by construction, as B_1 is of course empty and has isolated $\circ$–points in $(\mathcal{P}^*)$, while the general case follows easily from Lemma 4.18. Indeed, in the pair $(\Sigma^{j-1}, \mathcal{P}^{j-1})$ the region B'_{j-1} is empty and the points $\{y_{e_{2j-3}}, y_{e_{2j-2}}\}$ are isolated, so by Lemma 4.18 the same holds for $(B'_{j-1})^c = B_j$. □

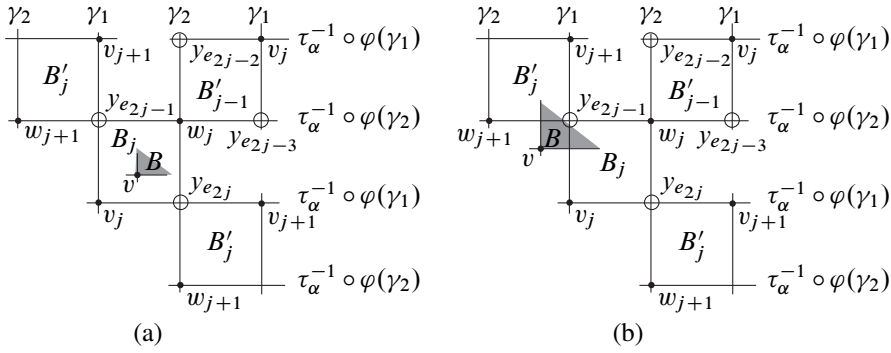


Figure 44: The setup for the argument of Lemma 4.20. As usual, the figure illustrates the situation in the universal cover of the surface; lifts of a common object are labeled identically.

We are now able to complete Theorem 1.2.

Proof of Theorem 1.2 Let $\alpha \in pe(\varphi)$. By Corollary 4.13, α is balanced, so let $\mathcal{P}$ be a consistent right position for $(\Sigma, \tau_\alpha^{-1} \circ \varphi, \{\gamma_1, \gamma_2\})$, and let $\mathcal{P}^* \subset \mathcal{P}$ be as in Lemma 4.12. But by Lemma 4.20, each region $B \in \mathcal{R}(\mathcal{P}^*)$ is empty in $\mathcal{P}^*$, with isolated $\circ$–points, and so by Lemma 4.17 each B is nested. Thus, using Observation 4.15, α is nested. □

5 Non-positive open books of Stein-fillable contact 3–folds

In this section we prove Theorem 1.3 by constructing explicit examples of open book decompositions which support Stein-fillable contact structures yet whose monodromies have no positive factorizations. We first introduce a construction, based on a modification of the lantern relation, which allows us to introduce essential left-twisting into a stabilization-equivalence class of open book decompositions. As, by Giroux, elements of such an equivalence class support a common contact manifold, proving the essentiality of this left-twisting (accomplished here using the methods of the previous sections) is sufficient to produce examples of non-positive open books which support Stein-fillable contact structures.

5.1 Immersed lanterns

For our construction, we recall the definition of a stabilization of an open book decomposition.

Definition 5.1 Given an open book decomposition (Σ, φ), let σ be a properly embedded arc in Σ. Then if Σ' denotes the result of adding a 1–handle to Σ with attaching sphere $\partial(\sigma)$, and s the (unique up to isotopy) simple closed curve in Σ' obtained by extending σ over the core of this handle, then the open book decomposition $(\Sigma', \tau_s \circ \varphi)$ is a *stabilization* of (Σ, φ). The inverse operation is referred to as *destabilization*.

We start out then with a surface $\Sigma_{0,4}$ of genus zero with 4 boundary components, and a mapping class with factorization $\tau_{\epsilon_1}\tau_{\epsilon_2}$, where ϵ_1 and ϵ_2 intersect exactly twice, with opposite sign (so $\Sigma_{0,4}$ is just a regular neighborhood of the pair, see Figure 45 (a)). It is easy to see that this defines an open book decomposition of $S^1 \times S^2$ with the standard (and unique) Stein-fillable contact structure. We use the well-known lantern relation to give a mapping class equivalence between the words indicated in Figure 45(a) and (b), thus introducing a left twist about the curve α into the positive monodromy. To make room for what is to come, we must enlarge the surface by adding a 1–handle (see Figure 45(c)) to the surface, so that the associated open book decomposition is now of $\#^2(S^1 \times S^2)$, with its (also unique) Stein-fillable contact structure. We then stabilize, use this stabilization curve to "braid" two of the lantern curves into a new configuration in which the lantern relation does not apply, and then destabilize to a book in which, as the lantern is unavailable, the left twist about α can no longer be canceled. The steps are indicated clearly in the remainder of Figure 45. As none of the steps (c) through (f) affect the supported contact structure, we have obtained an open book decomposition, supporting a Stein-fillable contact structure, which has no obvious positive factorization. We refer to this construction as an *immersed* lantern relation. To motivate this terminology, observe that the surface and curves of Figure 45 (f) are obtained from those of Figure 45(b) by a self plumbing of the surface. Figure 46 summarizes the construction.

The remainder of this section uses the results of Section 4 to show:

Theorem 5.2 *Let $\varphi_{e_1,e_2} = \tau_{\delta_1}\tau_{\delta_2}\tau_{\beta_1}^{e_1}\tau_{\beta_2}^{e_2}\tau_{\alpha}^{-1}$, where all curves, and the surface Σ in which they exist, are as in Figure 45 (f). Then for each pair of positive integers e_1 and e_2, the open book decomposition $(\Sigma, \varphi_{e_1,e_2})$ supports a Stein-fillable contact structure, yet φ_{e_1,e_2} admits no factorization into positive Dehn twists.*

As a simple example, for the case $e_2 = 1$ it is simple to verify (eg by tracing backwards through Figure 45) that we obtain an open book decomposition of $(S^1 \times S^2) \# L(e_1 - 1, e_1 - 2)$ for each $e_1 \geq 2$ (see Figure 47).

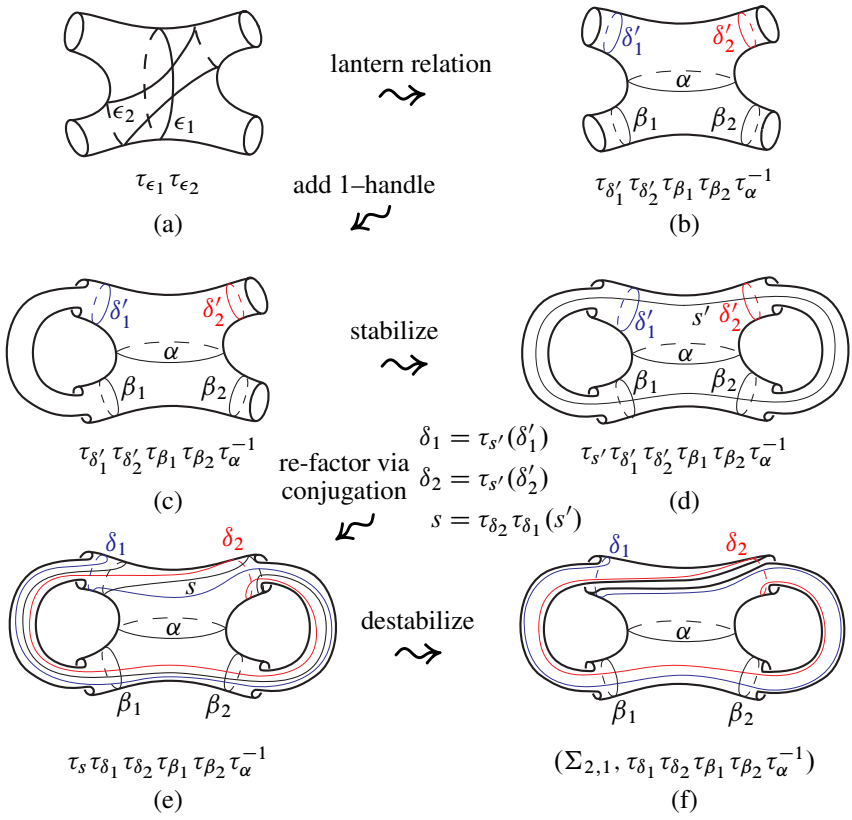
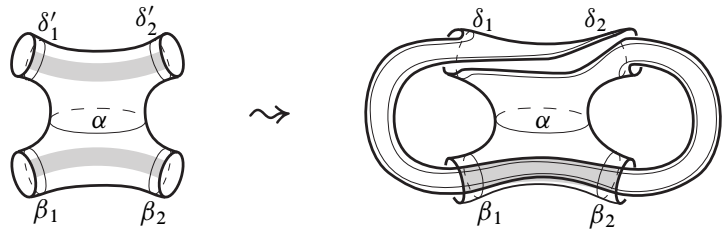


Figure 45: $(\#^2(S^1\times S^2), \xi_{st})$ is supported by $(\Sigma_{2,1}, \tau_{\delta_1}\tau_{\delta_2}\tau_{\beta_1}\tau_{\beta_2}\tau_\alpha^{-1})$

5.2 The positive extension of φ_{e_1,e_2}

Recall from Definition 4.1 that the *positive extension*, $pe(\varphi)$, of a positive mapping class φ is simply the set of all simple closed curves α such that τ_α appears in a positive factorization of φ. Our method of demonstrating essentiality of the left twisting introduced by the immersed lantern is as follows. Let $\varphi = \varphi_{e_1,e_2}$ be as in Theorem 5.2, for arbitrary positive integers e_1 and e_2, and define $\varphi' := \tau_\alpha \circ \varphi = \tau_{\delta_1}\tau_{\delta_2}\tau_{\beta_1}^{e_1}\tau_{\beta_2}^{e_2}$. Note that α has no intersection with any of the other curves, so τ_α commutes with each twist. Suppose that φ has positive factorization $\tau_{\alpha_n}\cdots\tau_{\alpha_1}$. Then we may factorize $\varphi' = \tau_\alpha\tau_{\alpha_n}\cdots\tau_{\alpha_1}$. Our goal is then to derive a contradiction by showing that $\alpha \notin pe(\varphi')$. The argument, to which the remainder of this subsection is devoted, comprises two steps. Firstly, we show that α has trivial intersection with each curve in $pe(\varphi')$, and secondly use this to conclude that $\alpha \notin pe(\varphi')$.



$\alpha \in pe(\tau_{\delta_1'}\tau_{\delta_2'}\tau_{\beta_1}\tau_{\beta_2})$ $\qquad$ $\alpha \notin pe(\tau_{\delta_1}\tau_{\delta_2}\tau_{\beta_1}\tau_{\beta_2})$

Figure 46: Immersing a lantern. To the left, curves of the lantern relation. In particular, a positive twist about each boundary parallel curve is sufficient to cancel a negative twist about the curve α. To the right, the curves in an "immersed" configuration for which the cancellation no longer holds. Topologically, one obtains the immersed configuration by a self-plumbing of the surface, ie an identification of the shaded rectangles via a 90° twist as in the figure. As open book decompositions, one gets from one picture to the other by the steps of Figure 45. The effect on the contact manifold is a connect sum with $S^1 \times S^2$, with the standard (Stein-fillable) contact structure.

$(\Sigma_{0,4}, \tau_{\epsilon_1}\tau_{\epsilon_2}\tau_{\beta_1}^{e_1-1})$ $\qquad$ $(\Sigma_{1,1}, \tau_{\beta_1}^{e_1-1})$

Figure 47: One may easily verify that the open book to the left may be destabilized twice to give the one to the right, which obviously supports $((S^1 \times S^2) \# L(e_4-1, e_4-2), \xi_{st})$.

For the first step, we wish to apply the results of Section 4 to the pair γ_1, γ_2 shown in Figure 48 to conclude that each curve in $pe(\varphi')$ may be isotoped so as not to intersect α. Note that $\varphi(\gamma_i)$ is isotopic to $\tau_{\delta_1}\tau_{\delta_2}(\gamma_i)$ for each i; in particular, this data does not depend on e_1 and e_2. Now, $\Sigma \setminus \{\gamma_1, \gamma_2\}$ is a pair of pants bounded by β_1, β_2, and α, and so these three curves are the only elements of $\mathrm{SCC}(\Sigma)$ which have no intersection with $\{\gamma_1, \gamma_2\}$. The pair $\varphi(\gamma_1), \varphi(\gamma_2)$ is flat (Definition 4.5), and so by Theorem 1.2, each curve $\epsilon \in pe(\varphi') \setminus \{\beta_1, \beta_2, \alpha\}$ is nested.

So as to simplify our view of the situation, we begin by cutting Σ along $\bar{\gamma}_i, i = 1, 2$ (see Figure 48). The resulting surface is a pair of pants Σ' with boundary components

$\partial_i, i = 1, 2, 3$, labeled such that α is parallel to ∂_1. As ϵ is balanced, we have $\#|\epsilon \cap \bar{\gamma}_1| = \#|\epsilon \cap \bar{\gamma}_2|$.

Examples are given in Figure 49.

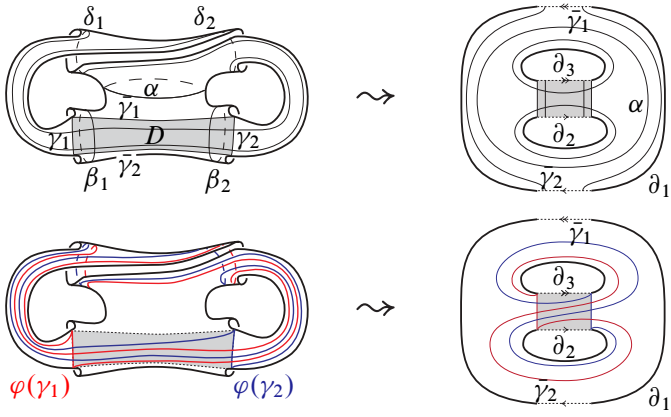


Figure 48: To the left, various curves and arcs in the surface Σ. To the right, the result Σ' of cutting Σ along the $\bar{\gamma}_i$.

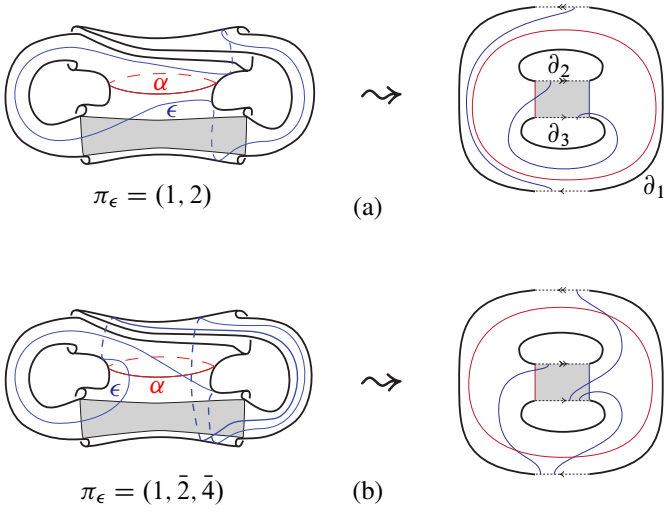


Figure 49: (a) A nested curve ϵ in Σ, and the result of cutting Σ as described above. (b) A curve ϵ which is neither nested nor balanced.

We summarize various observations concerning arcs $\epsilon \cap \Sigma'$, for $\epsilon \in \mathrm{SCC}(\Sigma)$, in the following lemma and corollary.

Lemma 5.3 *Let Σ, Σ', α and D be as above. Then if ϵ is a nested curve on D,*

(1) *No component of $\epsilon \cap \Sigma'$ has both endpoints in ∂_2 or both endpoints in ∂_3.*

(2) *Any component of $\epsilon \cap \Sigma'$ with both endpoints in ∂_1 is boundary-parallel.*

Proof For (1), suppose there is some component γ of $\epsilon \cap \Sigma'$ with each endpoint on ∂_2. Now if γ is boundary-parallel, then it corresponds to either a bigon bounded by ϵ and one of the $\bar{\gamma}_i$, contradicting intersection minimality, or a downward arc $\epsilon \cap D$, contradicting Lemma 4.6. The same argument shows that there is no boundary parallel arc with both endpoints on ∂_3. If however γ is *not* boundary parallel, it separates Σ' into two annulus components, one of which contains a proper subset of $\epsilon \cap \partial_2$ on one boundary component, while the other boundary component is the original ∂_3. As nestedness implies that $\#|\epsilon \cap \partial_2| = \#|\epsilon \cap \partial_3|$, this second annulus must then contain some boundary parallel arc connecting points of ∂_3, a contradiction.

For (2), any arc with each endpoint on ∂_1 which is not boundary-parallel would separate Σ' into two annuli so that one annulus had the points $\epsilon \cap \partial_2$ on one boundary component, and some subset of the points $\epsilon \cap \partial_1$ on the other, while the other annulus had the points $\epsilon \cap \partial_3$ on one boundary component, again some subset of the points $\epsilon \cap \partial_1$ on the other. But, letting $m = \#|\epsilon \cap \partial_2| = \#|\epsilon \cap \partial_3| = \frac{1}{2}(\#|\epsilon \cap \partial_1|)$, there are a total of $2m-2$ remaining points of $\epsilon \cap \partial_1$. This would then necessitate some arc parallel to ∂_2 or ∂_3, contradicting (1). □

Corollary 5.4 *Given Σ, δ_1, and δ_2 as above (see Figure 48), let $\psi \in \mathrm{Dehn}^+(\Sigma)$ be such that $\psi(\gamma_i)$ is isotopic to $\tau_{\delta_1}\tau_{\delta_2}(\gamma_i)$, for $i = 1, 2$. Then for each $\epsilon \in pe(\psi)$ which is not isotopic to any δ_i or γ_i, any horizontal component of $\epsilon \cap D$ is contained in a component of $\epsilon \cap \Sigma'$ which intersects neither $\psi(\gamma_i)$ and has endpoints on ∂_2 and ∂_3. Similarly, any vertical component of $\epsilon \cap D$ is contained in a component of $\epsilon \cap \Sigma'$ which intersects neither γ_i and has endpoints on ∂_2 and ∂_3.*

Proof Let σ be a connected component of $\epsilon \cap \Sigma'$ which contains a horizontal segment of $\epsilon \cap D$. Now, such a σ is not parallel to ∂_1, so Lemma 5.3 tells us that the endpoints of σ are on distinct boundary components. Assume then that σ has an endpoint on ∂_2 (the case of an endpoint on ∂_3 is identical). We consider firstly the case that the remaining endpoint is on ∂_3. As δ_1 and δ_2 are isotopic to ∂_2 and ∂_3 respectively, σ is isotopic to $\tau_{\delta_1}^{t_1}\tau_{\delta_2}^{t_2}(\gamma)$, for some integers t_1 and t_2, and properly embedded arc γ parallel to γ_2. Now, if either t_i were negative, we would have a downward arc of $\epsilon \cap D$ (see Figure 50(a)). On the other hand, if either t_i were greater than 1, we would have (for $i = 2$) $\tau_\epsilon(\gamma_2) > \psi(\gamma_2)$ (from c_2'), see Figure 50(b). We thus have $t_i \in \{0, 1\}$ for

each i. It is immediate to check that the case $t_1 = t_2 = 1$ is the unique case containing a horizontal arc, and further that it intersects neither $\varphi(\gamma_i)$, as desired.

Suppose then that the remaining endpoint is on ∂_1, so that σ is isotopic to $\tau_{\delta_1}^{t_1}\tau_\alpha^{t_3}(\gamma')$, where γ' is an arc from ∂_1 to ∂_3 satisfying $\gamma' \cap \gamma_1 = \varnothing$ and $\#|\gamma' \cap \gamma_2| = 1$ (see Figure 50(c)). Then, for horizontality, we must have $t_1 > 0$, in which case again $\tau_\epsilon(\gamma_2) > \psi(\gamma_2)$ (from c_2'), giving a contradiction.

The vertical case is of course trivial, as a vertical arc of $\epsilon \cap D$ is by definition itself an arc of $\epsilon \cap \Sigma'$ which intersects neither γ_i and has endpoints on ∂_2 and ∂_3. □

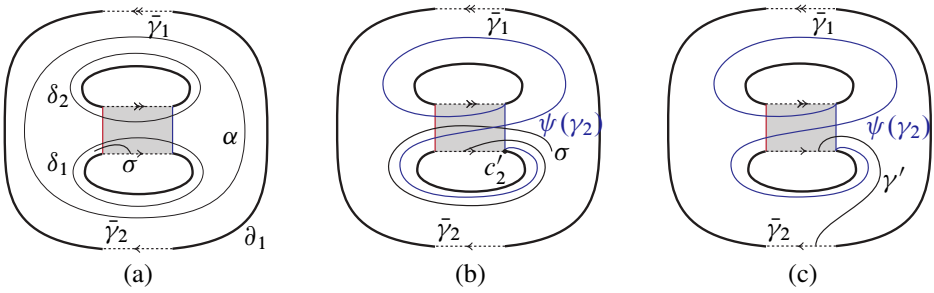


Figure 50: Arcs of $\epsilon \cap \Sigma'$

Throughout the rest of this section, we will be interested in positive factorizations of a given monodromy only up to conjugation. As such, we write $\omega \overset{b}{\sim} \omega'$ to indicate that $\omega, \omega' \in \mathrm{Fac}^+(\psi)$ are related through conjugation, and $[\omega]_b$ for the associated equivalence class. Our main lemma is the following.

Lemma 5.5 *Let ψ be as in the previous corollary, and ω a positive factorization of ψ. Then there is $\omega' \overset{b}{\sim} \omega$ such that no curve in ω' has any diagonal arc on D.*

Proof Observe firstly that, through conjugation, we may assume any given factorization of ψ is $\tau_\alpha^e \omega_h \omega_d \omega_v$, for some $e \geq 0$, where the restriction to D of each curve involved in ω_h is a collection of horizontal arcs, those in ω_v are vertical, and each element of ω_d has some diagonal arc. We will further assume our given factorization is chosen from its conjugation-equivalence class such that the length of ω_d is minimized, and non-zero, and derive a contradiction. Let $\omega_d = \tau_{\epsilon_n}\tau_{\epsilon_{n-1}}\cdots\tau_{\epsilon_1}$.

Throughout the remainder of the proof, we will refer to a curve $\epsilon \in \mathrm{SCC}(\Sigma)$ as *vertical (horizontal)* if its restriction to D contains a vertical (horizontal) arc. We begin with the case that no ϵ_i is horizontal. Recalling the ordering on arcs with common endpoint described at the beginning of this paper in Example 2.1, observe that we may order

$\mathrm{SCC}(\Sigma)$ by considering the images of a given arc under twists about the curves. For our purposes, for $\epsilon, \epsilon' \in \mathrm{SCC}(\Sigma)$, we set $\epsilon \geq \epsilon' \iff \tau_\epsilon(\gamma_1) \geq \tau_{\epsilon'}(\gamma_1)$ from c_1.

We let S denote the set of elements of $[\omega_d]_b$ which maximize the index $m := \max\{i \mid \epsilon_i \geq \epsilon_j \,\forall\, j \leq i\}$, and then $S' \subset S$ those elements of S for which the intersection number $|\epsilon_m \cap \gamma_1|$ is minimized. Note that, as each curve is nested, any curve whose intersection with D contains a diagonal arc intersects each of γ_1 and γ_2; in particular $|\epsilon_m \cap \gamma_1| > 0$. We may then assume $\omega_d \in S'$.

Now, let x denote the initial point of $\tau_{\epsilon_m}(\gamma_1) \cap \bar{\gamma}_1$ from c_1. As ϵ_m is not horizontal, $[c_1, x]_{\tau_{\epsilon_m}(\gamma_1)}$ is contained in D. As $\psi(\gamma_1)$ is flat, there is some $k > m$ such that x is a vertex of a downward triangular region of the triple $(\epsilon_k, \tau_{\epsilon_m}(\gamma_1), \bar{\gamma}_1)$ (see Figure 51(a)). After conjugation, we may assume $k = m+1$. Now, using Lemma 4.6, the segment of $\epsilon_k \cap D$ which extends from this triangular region cannot be downward on c_2, thus is vertical. We then have a bigon of $D_{\epsilon_k}(\tau_{\epsilon_m}(\gamma_1))$ and $\bar{\gamma}_2$; as $\tau_{\epsilon_k}\tau_{\epsilon_m}(\gamma_1)$ cannot be further to the right from c_1 than $\psi(\gamma_1)$, we must have a canceling bigon (see Figure 51(b)), and point $x' \in \tau_{\epsilon_m}(\gamma_1) \cap \bar{\gamma}_2$ corresponding to a vertex of this bigon (see Figure 51(c)).

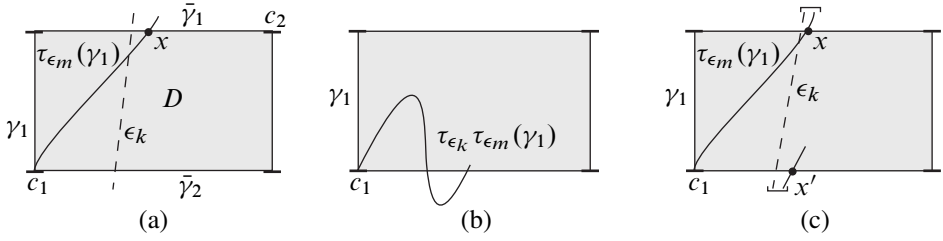


Figure 51

Let A denote a closed support neighborhood of the twist diffeomorphism D_{ϵ_m}. We label the points $\gamma_1 \cap \partial A$ as $z_0, z_1, z_2, \ldots$ increasing along γ_1 from c_1, so for instance for even i, $[z_i, z_{i+1}]_{\gamma_1}$ is contained in A (see Figure 52(a)). Each z_i is thus in $D_{\epsilon_m}(\gamma_1) \cap \gamma_1$; as $D_{\epsilon_m}(\gamma_1)$ and γ_1 bound no bigons, we may assume $\tau_{\epsilon_m}(\gamma_1)$ passes through each z_i, so for even i, $[z_i, z_{i+1}]_{\tau_{\epsilon_m}(\gamma_1)}$ is isotopic to $\tau_{\epsilon_m}([z_i, z_{i+1}]_{\gamma_1})$. Let r denote $\max\{i \mid z_i \in [x, x']_{\tau_{\epsilon_m}(\gamma_1)}\}$, so $\epsilon_k \cap D$ has $r+1$ connected components (see Figure 52(b)).

We note firstly that $r \geq 1$. Indeed, supposing otherwise, letting y denote the unique point $\epsilon_m \cap [z_0, z_1]_{\gamma_1}$ we see that y is not upward in $(\epsilon_k, \gamma_1, \epsilon_m)$ (see Figure 53(a)), and so $\tau_{\epsilon_k}(\epsilon_m) > \epsilon_m > \epsilon_k$ (see Figure 53(b)). We may then conjugate $\tau_{\epsilon_k}\tau_{\epsilon_m} \rightsquigarrow \tau_{\tau_{\epsilon_k}(\epsilon_m)}\tau_{\epsilon_k}$, obtaining a new element of $[\omega_d]_b$ which contradicts our maximality assumption on the index m.

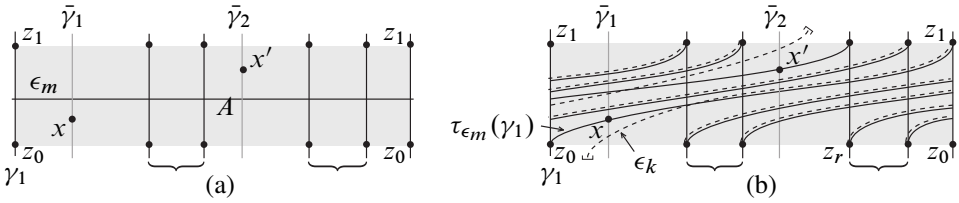


Figure 52: The neighborhood A of ϵ_m, and the various arcs it intersects. Dark vertical lines are segments of γ_1.

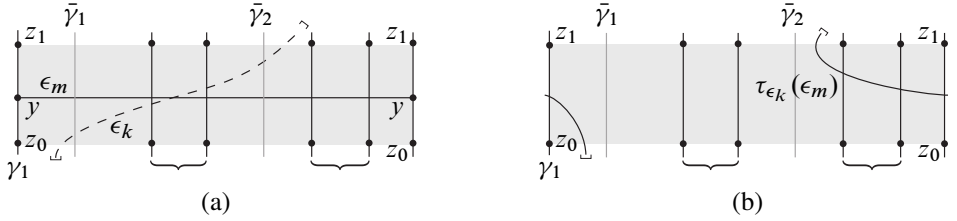


Figure 53

In general then, we see that ϵ_m intersects *each* component of $A \cap \gamma_1$ (see Figure 54(a)), while $\tau_{\epsilon_m}^{-1}(\epsilon_k)$ intersects each such component at most once, and does not intersect the component $[z_r, z_{r+1}]_{\gamma_1}$ (see Figure 54(b)). As all intersections $\epsilon_k \cap \gamma_1$ were by construction contained in A, we conclude that $\#|\tau_{\epsilon_m}^{-1}(\epsilon_k) \cap \gamma_1| < \#|\epsilon_m \cap \gamma_1| < \#|\epsilon_k \cap \gamma_1|$. But then through conjugation $\tau_{\epsilon_k}\tau_{\epsilon_m} \rightsquigarrow \tau_{\epsilon_m}\tau_{\tau_{\epsilon_m}^{-1}(\epsilon_k)}$, we have $\tau_{\epsilon_m}^{-1}(\epsilon_k) > \epsilon_m$, so that our new word is in S, and thus a contradiction of our minimality assumption on intersection with γ_1.

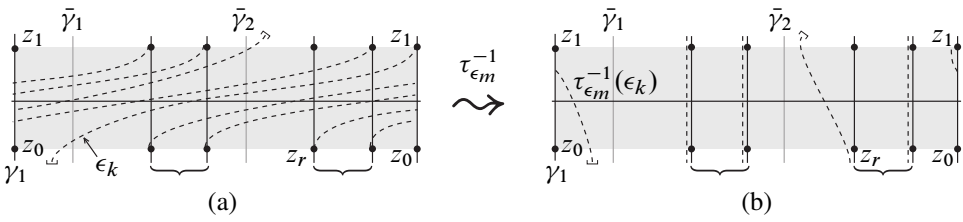


Figure 54

The remaining case, that some ϵ_i is horizontal, follows from a similar argument. In particular, we now order our curves with respect to γ_2 and its endpoint c_2', so $\epsilon \geq \epsilon' \iff \tau_\epsilon(\gamma_2) \geq \tau_{\epsilon'}(\gamma_2)$ from c_2', and again assume ω_d is chosen from $[\omega_d]_b$ such that the index $m := \max\{i \mid \epsilon_i \geq \epsilon_j \,\forall\, j \leq i\}$ is maximized. Now x will refer to the initial point of $\tau_{\epsilon_m}(\gamma_2) \cap \bar\gamma_1$, and there is some $k > m$ such that x is a vertex of a downward triangular region of the triple $(\epsilon_k, \epsilon_m, \bar\gamma_2)$. As $\epsilon_k \cap D$ contains no vertical component (else $\tau_{\epsilon_k}\epsilon_m$ is downward), the component extending from this triangle is

upward on c_1' (see Figure 55(a)). Note that as Corollary 5.4 ensures that any horizontal curve is strictly greater than any non-horizontal curve, ϵ_m is horizontal, and the point of $\epsilon_m \cap \gamma_2$ closest to c_2' is contained in an arc component σ of $\epsilon_m \cap \Sigma'$ which connects ∂_2 and ∂_3, intersecting neither of the $\psi(\gamma_i)$.

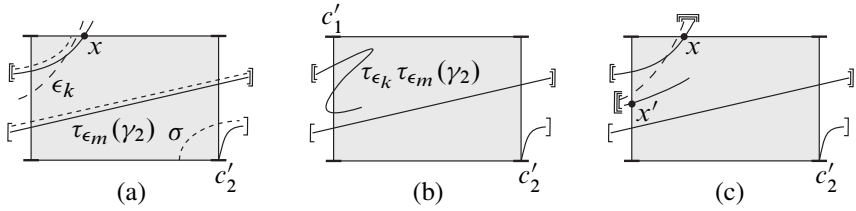


Figure 55: The neighborhood A of ϵ_m, and the various arcs it intersects. Dark vertical lines are now segments of γ_2.

In particular, the segment $[c_2', x]_{\tau_{\epsilon_m}(\gamma_2)}$ is isotopic to $[c_2', x]_{\psi(\gamma_2)}$ up to its 2nd intersection with γ_1. Thus it again follows that we have a canceling pair of bigons in $D_{\epsilon_k}(\tau_{\epsilon_m}(\gamma_2))$, else $\tau_{\epsilon_k}\tau_{\epsilon_m}(\gamma_2)$ is further to the right from c_2' than $\psi(\gamma_2)$ (see Figure 55(b)). We now let $x' \in \tau_{\epsilon_m}(\gamma_2) \cap \gamma_1$ be the point corresponding to a vertex of this bigon (see Figure 55(c)).

We again consider the support neighborhood A of τ_{ϵ_m} and label the points in $\gamma_2 \cap \partial A$ by $z_0, z_1, z_2, \dots$, increasing along γ_2 from c_2' (see Figure 56(a)). Exactly as in the vertical case, our maximality condition on m implies that $z_1 \in [x, x']_{\tau_{\epsilon_m}(\gamma_2)}$, but then $\#|\tau_{\epsilon_m}^{-1}(\epsilon_k) \cap \gamma_2| < \#|\epsilon_k \cap \gamma_2|$ (see Figure 56(b)), contradicting our minimality condition. □

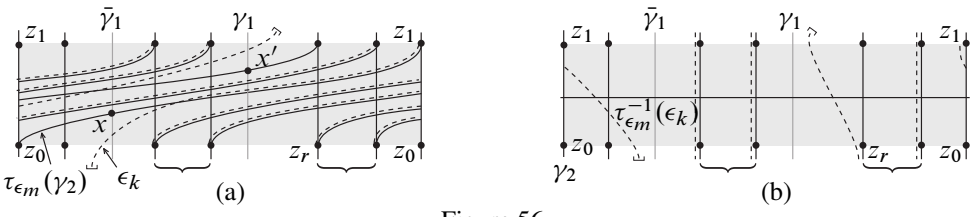


Figure 56

We bring all of this together with:

Proof of Theorem 5.2 Suppose that φ_{e_1,e_2} is positive, so $\tau_\alpha^{-1}\tau_{\delta_1}\tau_{\delta_2}\tau_{\beta_1}^{e_1}\tau_{\beta_2}^{e_2}$ has some positive factorization $\tau_{\alpha_n}\cdots\tau_{\alpha_1}$. Then letting $\varphi' = \tau_\alpha \circ \varphi_{e_1,e_2}$, we have $\varphi' = \tau_{\delta_1}\tau_{\delta_2}\tau_{\beta_1}^{e_1}\tau_{\beta_2}^{e_2} = \tau_\alpha\tau_{\alpha_n}\cdots\tau_{\alpha_1}$; ie $\alpha \in pe(\varphi')$. At the same time, Lemma 5.5 ensures that $[\tau_\alpha\tau_{\alpha_n}\cdots\tau_{\alpha_1}]_b$ contains an element $\tau_{\epsilon_n}\cdots\tau_{\epsilon_1}$ with the property that for all i, ϵ_i is

either disjoint from, or isotopic to, α. As this property is preserved under conjugation, it holds for $\tau_\alpha \tau_{\alpha_n} \cdots \tau_{\alpha_1}$ as well. It follows that if γ is a properly embedded arc which intersects α exactly once, then there are points $x, x' \in \varphi'(\gamma) \cap \gamma$ such that $[x, x']_{\varphi'(\gamma)} \cup [x, x']_\gamma$ is a simple closed curve isotopic to α. The choice of γ shown in Figure 57 then gives a contradiction. We have drawn the image for the case $e_1 = e_2 = 1$; the general case follows immediately. □

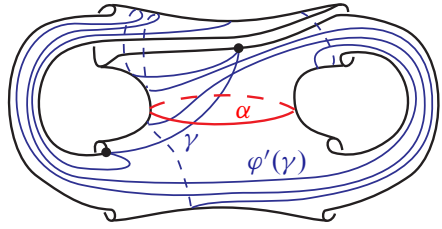


Figure 57: The arc γ, and its image $\varphi'(\gamma)$.

References

[1] **S Akbulut**, **B Ozbagci**, *Lefschetz fibrations on compact Stein surfaces*, Geom. Topol. 5 (2001) 319–334 MR1825664

[2] **K L Baker**, **J B Etnyre**, **J Van Horn-Morris**, *Cabling, contact structures and mapping class monoids*, J. Differential Geom. 90 (2012) 1–80 MR2891477

[3] **B Farb**, **D Margalit**, *A primer on mapping class groups*, Princeton Mathematical Series 49, Princeton Univ. Press (2012) MR2850125

[4] **E Giroux**, *Géométrie de contact: de la dimension trois vers les dimensions supérieures*, from: "Proceedings of the International Congress of Mathematicians, II", (T Li, editor), Higher Ed. Press, Beijing (2002) 405–414 MR1957051

[5] **N Goodman**, *Overtwisted open books from sobering arcs*, Algebr. Geom. Topol. 5 (2005) 1173–1195 MR2171807

[6] **K Honda**, **W H Kazez**, **G Matić**, *Right-veering diffeomorphisms of compact surfaces with boundary*, Invent. Math. 169 (2007) 427–449 MR2318562

[7] **Y Lekili**, *Planar open books with four binding components*, Algebr. Geom. Topol. 11 (2011) 909–928 MR2782547

[8] **A Loi**, **R Piergallini**, *Compact Stein surfaces with boundary as branched covers of B^4*, Invent. Math. 143 (2001) 325–348 MR1835390

Laboratoire de Mathématiques Jean Leray, Université de Nantes
2, rue de la Houssinière, 44322 Nantes, France

wand@math.berkeley.edu

Proposed: Danny Calegari
Seconded: Dmitri Burago, Peter S. Ozsváth

Received: 9 October 2009
Revised: 3 October 2012