

AG  
T

*Algebraic & Geometric  
Topology*

Volume 25 (2025)

**A group-theoretic framework for low-dimensional topology  
Or: how not to study low-dimensional topology?**

SARAH BLACKWELL

ROBION KIRBY

MICHAEL KLUG

VINCENT LONGO

BENJAMIN RUPPIK

# A group-theoretic framework for low-dimensional topology

## Or: how not to study low-dimensional topology?

SARAH BLACKWELL  
 ROBION KIRBY  
 MICHAEL KLUG  
 VINCENT LONGO  
 BENJAMIN RUPPIK

A correspondence, by way of Heegaard splittings, between closed, oriented 3-manifolds and pairs of surjections from a surface group to a free group has been studied by Stallings, Jaco and Hempel. This correspondence, by way of trisections, was recently extended by Abrams, Gay and Kirby to the case of smooth, closed, connected, oriented 4-manifolds. We unify these perspectives and generalize this correspondence to the case of links in closed, oriented 3-manifolds and links of knotted surfaces in smooth, closed, connected, oriented 4-manifolds. The algebraic manifestations of these four subfields of low-dimensional topology (3-manifolds, 4-manifolds, knot theory and knotted surface theory) are all strikingly similar, and this correspondence perhaps elucidates some unique character of low-dimensional topology.

57K40; 20F05, 57M05

|   |      |
|---|------|
| 1. Introduction                                     | 4667 |
| 2. Trivial tangles in handlebodies from algebra     | 4672 |
| 3. Closed 3-manifolds and bridge split links        | 4691 |
| 4. Closed 4-manifolds and bridge trisected surfaces | 4703 |
| 5. Examples and consequences                        | 4711 |
| References  | 4716 |

## 1 Introduction

All manifolds and submanifolds discussed in this paper are smooth and, with the exception of surfaces in 4-manifolds, oriented. We use decompositions of manifolds in dimensions three and four, possibly together with links, to give a group-theoretic framework for studying these spaces. We begin by reviewing the simplest case of closed 3-dimensional manifolds, where this work has already been carried out by Stallings and Jaco [40; 20].

A *Heegaard decomposition*, or *Heegaard splitting*, of a closed 3-manifold  $M^3$  is a pair of handlebodies  $H_1$  and  $H_2$  embedded inside of  $M$  with boundaries a common genus  $g$  surface  $\Sigma_g$  such that  $M = H_1 \cup_{\Sigma_g} H_2$ . Every such 3-manifold admits a Heegaard decomposition (for example by triangulating  $M$  and taking a regular neighborhood of the 1-skeleton). By choosing a basepoint on  $\Sigma_g$ , we then obtain the following pushout diagram between fundamental groups, where the maps are induced by inclusion:

$$\begin{array}{ccc} \pi_1(\Sigma_g, *) & \longrightarrow & \pi_1(H_1, *) \\ \downarrow & & \downarrow \\ \pi_1(H_2, *) & \longrightarrow & \pi_1(M, *) \end{array}$$

Note that  $\pi_1(H_1, *)$  and  $\pi_1(H_2, *)$  are both free groups of rank  $g$ , and the maps are surjections. Jaco proved that, given a surjective homomorphism  $\phi: \pi_1(\Sigma_g, *) \twoheadrightarrow F_g$ , there is a unique handlebody  $H(\phi)$  with  $\partial H(\phi) = \Sigma_g$  such that the map induced on  $\pi_1$  by inclusion of  $\Sigma_g$  as the boundary agrees with  $\phi$  (see Jaco [19], Lemma 2.3, and also Leininger and Reid [29, Lemma 2.2] for a simpler proof in this case). From this, it follows that a pair of surjective homomorphisms  $(\phi_1, \phi_2)$  with  $\phi_i: \pi_1(\Sigma_g, *) \twoheadrightarrow F_g$  determines a 3-manifold  $H(\phi_1) \cup_{\Sigma_g} H(\phi_2)$ , and that every closed 3-manifold arises in this way. Jaco referred to these pairs of maps as *splitting homomorphisms*.

One concrete application of this is the following group-theoretic recasting of the 3-dimensional Poincaré conjecture. Writing  $\pi_1(\Sigma_g, *) = \langle a_1, b_1, \dots, a_g, b_g \mid [a_1, b_1] \cdots [a_g, b_g] = 1 \rangle$ , there is a surjective homomorphism

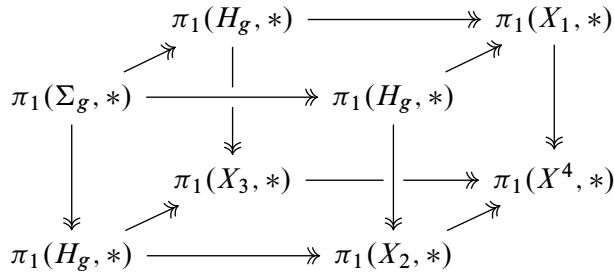
$$\pi_1(\Sigma_g, *) \twoheadrightarrow \langle x_1, \dots, x_g \rangle \times \langle y_1, \dots, y_g \rangle, \quad a_i \mapsto (x_i, 1), \quad b_j \mapsto (1, y_j).$$

The Poincaré conjecture is equivalent to the statement that this is the unique surjective homomorphism of these groups modulo precomposing with automorphisms and postcomposing with products of automorphisms (see Hempel [16]). Thus, by Perelman's work [35], this result follows, and we are left in the state where the only known proof of this perhaps innocent-looking group-theoretic result involves a careful analysis of Ricci flow.

In addition to the observation that every closed 3-manifold admits a Heegaard decomposition, there is a corresponding uniqueness theorem called the *Reidemeister–Singer theorem*, which states that any two Heegaard decompositions of a fixed 3-manifold differ by a sequence of simple inverse geometric operations called stabilization and destabilization [36; 39]. Jaco proposed a way of incorporating the Reidemeister–Singer theorem into the construction of 3-manifolds from appropriate pairs  $(\phi_1, \phi_2)$  to obtain a bijective correspondence [20].

More recently, a 4-dimensional analogue of Heegaard splittings, called *trisections*, together with a corresponding uniqueness theorem has been introduced by Gay and Kirby [8]. A trisection of a closed 4-manifold  $X^4$  is a decomposition  $X = X_1 \cup X_2 \cup X_3$  into 4-dimensional 1-handlebodies  $X_i$ , which pairwise intersect in genus  $g$  handlebodies  $H_g$ , and with triple intersection a genus  $g$  surface  $\Sigma_g$ . Every

smooth, closed, connected, oriented 4-manifold admits a trisection, which is unique up to a stabilization operation [8]. (See Section 4A for a further review of trisections.) The inclusion maps between the various components of a trisection of a 4-manifold induce maps between their fundamental groups, which produces the following commutative diagram, where every face is a pushout and every homomorphism is surjective, and in which the basepoint is chosen to lie on  $\Sigma_g$ :



Abrams, Gay and Kirby [1] noticed that the analogue of being able to recover a 3-manifold from a pair of surjective homomorphisms  $(\phi_1, \phi_2)$  holds in dimension four via trisections. Namely, given three surjective homomorphisms  $(\phi_1, \phi_2, \phi_3)$  with  $\phi_i: \pi_1(\Sigma_g, *) \twoheadrightarrow F_g$  such that the pairwise pushout of any pair  $\phi_i$  and  $\phi_j$  is a free group  $F_k$ , since  $\#^k(S^1 \times S^2)$  is the unique closed, orientable 3-manifold with fundamental group  $F_k$  (by Perelman’s work [35]) we obtain a closed 4-manifold by realizing three handlebodies  $H(\phi_i)$ , gluing them along their common boundary  $\Sigma_g$ , and filling in their pairwise unions, which are diffeomorphic to  $\#^k(S^1 \times S^2)$ , with three 4-dimensional 1-handlebodies (uniquely by Laudenbach and Poénaru [28]). They called this triple of maps (which then determine the entire cube pictured above) a *group trisection*, where the object being trisected is the group resulting from pushing out the three maps into a cube (in this case  $\pi_1(X^4, *)$ ).

Additionally, Abrams, Gay and Kirby [1] use the uniqueness theorem for trisections to obtain results analogous to those previously mentioned in dimension three. Namely, they obtain a group-theoretic statement that is equivalent to the smooth 4-dimensional Poincaré conjecture and, by modding out the set of such triples  $(\phi_1, \phi_2, \phi_3)$ , they obtain a bijection between a group-theoretically defined set and the set of all smooth, closed, connected, oriented 4-manifolds.

Not only can every 3-manifold be split into a union of two handlebodies, but additionally, given a link  $L \subset M$ , we have a Heegaard splitting  $M = H_1 \cup_{\Sigma_g} H_2$  such that the tangles  $T_1 = L \cap H_1$  and  $T_2 = L \cap H_2$  are trivial (that is, consist of arcs that can all be simultaneously isotoped in  $H_i$  into  $\Sigma_g$ ). This is called a *bridge splitting* of  $L \subset M$ . Note that the complement of  $L$  in each handlebody is again a handlebody and hence has free fundamental group. In the case of  $M = S^3$  with the Heegaard splitting into balls, this is the classical setting of *bridge position* of links (see [38]).

One dimension up, a similar story emerges. A *knotted surface* is a closed (potentially nonorientable or disconnected) surface smoothly embedded in a 4-manifold. Meier and Zupan showed that a knotted surface in a trisected 4-manifold can always be isotoped to be in *bridge position*, meaning that it

$$\begin{array}{ccc}
\{\phi_1, \phi_2: \pi_1(\Sigma_g, *) \twoheadrightarrow F_g\}/\sim & \xrightarrow[\text{Heegaard splittings}]{\cong} & \{3\text{-manifolds}\}/\text{diff.} \\
\{\phi_1, \phi_2: \pi_1(\Sigma_g - \{2b \text{ pts}\}, *) \twoheadrightarrow F_{g+b}\}/\sim & \xrightarrow[\text{bridge splittings}]{\cong} & \{\text{links in 3-manifolds}\}/\text{diff.} \\
\{\phi_1, \phi_2, \phi_3: \pi_1(\Sigma_g, *) \twoheadrightarrow F_g\}/\sim & \xrightarrow[\text{trisections}]{\cong} & \{4\text{-manifolds}\}/\text{diff.} \\
\{\phi_1, \phi_2, \phi_3: \pi_1(\Sigma_g - \{2b \text{ pts}\}, *) \twoheadrightarrow F_{g+b}\}/\sim & \xrightarrow[\text{bridge trisections}]{\cong} & \{\text{surfaces in 4-manifolds}\}/\text{diff.}
\end{array}$$

Figure 1: A summary of the main results of this paper; the equivalences via bridge splittings and bridge trisections are novel, while the other equivalences are a recasting and unification of previous work. All manifolds are smooth and, with the exception of surfaces in 4-manifolds, oriented, and the diffeomorphisms are orientation-preserving. All maps  $\phi_i$  shown are surjective homomorphisms, and satisfy two algebraic conditions (see [Definition 2.1](#)). The homomorphisms  $\phi_i$  in the third and fourth rows need to satisfy the additional condition that they push out pairwise to free groups of an appropriate rank (see [Sections 4A](#) and [4B](#)).

intersects the trisected 4-manifold in such a way that the surface inherits its own trisection, called a *bridge trisection* [[31](#); [32](#)]. This is unique up to a stabilization operation [[31](#); [18](#)]. (See [Section 4B](#) for a further review of bridge trisections.) Given the existence and uniqueness of such a decomposition in this setting, it is natural to wonder whether knotted surfaces in 4-manifolds can also be given such a group-theoretic framework. Achieving this goal was the initial motivation for this work.

The main results of this paper are bijective correspondences from group-theoretic sets to the set of 3-manifolds together with a link and 4-manifolds together with a knotted surface, and are summarized in [Figure 1](#). Just as the cases of 3-manifolds and 4-manifolds are facilitated by Heegaard splittings and trisections, respectively, our results for links in 3-manifolds and surfaces in 4-manifolds use bridge splittings and bridge trisections, respectively.

In order to get off the ground constructing these spaces from appropriate group homomorphisms, we need to know how to recover a trivial tangle  $T(\phi)$  with boundary points  $\{p_1, \dots, p_{2b}\}$  in a handlebody  $H(\phi)$  with boundary  $\Sigma_g$  from a suitable homomorphism

$$\phi: \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \twoheadrightarrow F_{g+b},$$

where  $F_{g+b}$  is playing the role of the fundamental group of the complement of the trivial tangle. (For the precise algebraic conditions on  $\phi$  needed for this construction, we defer to [Definition 2.1](#).) [Section 2](#) is dedicated to this task. Our first main result (see [Theorem 2.10](#)), and the result underlying all of the constructions of spaces from group homomorphisms, is a method for constructing  $H(\phi)$  and  $T(\phi)$  algorithmically. This method is inspired by the procedure for computing the corresponding homomorphism given the topological data of the surface together with curves indicating the handlebody and the trivial tangle (see [Lemma 2.4](#)). When naively trying to construct diagrams for  $H(\phi)$  and  $T(\phi)$ , we run into the

possibility of constructing diagrams with too many curves. We fix this using bands to connect curves together, where the combinatorics of how the bands connect curves is guided by a process called Stallings folding [41], whose behavior is guaranteed to serve our purposes by the conditions placed on  $\phi$  (see the proof of [Theorem 2.10](#)).

With this construction in hand, the constructions of the maps in [Figure 1](#) are straightforward and surjectivity follows from existence of the various geometric decompositions. In [Sections 3](#) and [4](#) we discuss in detail the various geometric descriptions, the map in [Figure 1](#), and the various algebraically defined relations that need to be collectively modded out by on the set of homomorphisms in order to obtain a bijection. This latter part involves setting up appropriate relations on the set of homomorphisms that mimic the geometric moves needed in the corresponding uniqueness theorem.

For example, in the case of closed 3-manifolds, by way of Heegaard splittings, all such 3-manifolds can be described by a *Heegaard diagram*, and two Heegaard diagrams result in the same 3-manifold if and only if they are related by a sequence of handleslides, diffeomorphisms of the surface applied to the diagram, and stabilizations (this is a diagrammatic restating of the Reidemeister–Singer theorem; see [Theorem 3.1](#)). In this case, we need to mod out our set of homomorphisms by an equivalence relation generated by three relations  $\sim_h$ ,  $\sim_m$  and  $\sim_s$  ( $h$  for handleslide,  $m$  for mapping class and  $s$  for stabilization) that algebraically mimic the corresponding diagrammatic moves. In [Section 3A](#), we carry out this process for closed 3-manifolds and in [Sections 3B](#), [4A](#) and [4B](#) we do the analogous procedure for links in 3-manifolds, closed 4-manifolds and surfaces in 4-manifolds, respectively. In the “relative” cases of [Sections 3B](#) and [4B](#), there are additional relations needed, corresponding to the different types of stabilizations available in these settings.

It is unclear if this formalism will prove useful in deriving topological results (see our subtitle). However, in [Section 5](#) we give some additional examples and pose some questions regarding potential applications. One curious consequence of our work is that although smoothly knotted surfaces in the 4-sphere cannot be distinguished by fundamental groups (or even their complements), they *can* be distinguished by group trisections (see [Corollary 4.9](#)).

## Acknowledgments

We were inspired by Stallings’ paper *How not to prove the Poincaré conjecture* [40] in more ways than one.

This project started at the 2020 virtual *Summer trisections workshop*, which was financially supported by the NSF. We are grateful for this wonderful opportunity to exchange ideas, and would like to thank all of the participants for their input. We want to give special credit to Michelle Chu, David Gay, Gabriel Islambouli, Jason Joseph, Chris Leininger, Jeffrey Meier, Puttipong Pongtanapaisan, Arunima Ray and Alexander Zupan. We also would like to sincerely thank the referees for their careful reading of our paper and their insightful comments. Blackwell, Klug and Ruppik would additionally like to thank the Max Planck Institute for Mathematics in Bonn for its hospitality and financial support.

## 2 Trivial tangles in handlebodies from algebra

Let  $\Sigma_g$  denote the genus  $g$  oriented surface with basepoint  $*$  and marked points  $p_1, \dots, p_{2b}$  as in [Figure 2](#). We abuse notation and let  $p_1, \dots, p_{2b}$  also denote the fundamental group elements as pictured. Using the notation  $[a, b] = aba^{-1}b^{-1}$  for the commutator of  $a$  and  $b$ , we have

$$\pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) = \langle p_1, \dots, p_{2b}, a_1, b_1, \dots, a_g, b_g \mid p_1 \cdots p_{2b} = [a_1, b_1] \cdots [a_g, b_g] \rangle,$$

where the  $a_i, b_i$  are generators of  $\pi_1(\Sigma_g, *)$ . If  $b = 0$ , by convention this is the fundamental group of a closed genus  $g$  surface,

$$\pi_1(\Sigma_g, *) = \langle a_1, b_1, \dots, a_g, b_g \mid [a_1, b_1] \cdots [a_g, b_g] = 1 \rangle,$$

while for  $g = 0$  this is the group of a  $2b$ -times punctured sphere,

$$\pi_1(S^2 - \{p_1, \dots, p_{2b}\}, *) = \langle p_1, \dots, p_{2b} \mid p_1 \cdots p_{2b} = 1 \rangle.$$

Observe that for  $b \geq 1$  this is a free group on  $2g + 2b - 1$  generators, although it will be useful for us to instead remember the relation  $p_1 \cdots p_{2b} = [a_1, b_1] \cdots [a_g, b_g]$ , which we call the *surface relation*.

We take [Figure 2](#) to be our standard model for the genus  $g$  oriented surface  $\Sigma_g$  with basepoint  $*$  and marked points  $p_1, \dots, p_{2b}$ . Proponents of the right-hand rule may be disappointed in our model, as each  $a_i$  and  $b_i$  pair violate this convention, but our choice of surface relation mandates the labels and orientations of the  $a_i$  and  $b_i$  curves. We choose to write the relation in this form for notational convenience on the algebraic side.

A genus  $g$  *handlebody*  $H$  is a compact orientable 3-manifold whose boundary is a genus  $g$  closed surface, with the property that  $H$  can be cut along 2-dimensional disks so that the resulting space is a set of 3-dimensional balls. A  $b$ -component *trivial tangle*  $T$  in a handlebody  $H$  is a collection of  $b$  properly embedded arcs in  $H$  such that all of the arcs can be simultaneously isotoped into the boundary of  $H$ . Given two handlebodies containing trivial tangles  $(H_1, T_1)$  and  $(H_2, T_2)$  with the property that  $\partial H_1 = \partial H_2$  and  $\partial T_1 = \partial T_2$ , we say that  $(H_1, T_1)$  and  $(H_2, T_2)$  are *equivalent* if there exists a diffeomorphism  $H_1 \rightarrow H_2$  mapping  $T_1$  to  $T_2$  that is the identity on  $\partial H_1$ . In the special case where  $T_1$  and  $T_2$  are empty, we then say that the handlebodies  $H_1$  and  $H_2$  are equivalent.

We will be concerned with the set of equivalence classes of handlebodies and trivial tangles  $(H, T)$  such that  $\partial H = \Sigma_g$  and  $\partial T = \{p_1, \dots, p_{2b}\}$ . Any equivalence class of such a handlebody and trivial tangle can be described by a *diagram*  $\mathcal{D} = (C, S)$  on the surface  $\Sigma_g$ , made up of a collection of  $g$  disjoint homologically linearly independent simple closed curves  $C = \{C_1, \dots, C_g\}$  (referred to as a *cut system*) together with  $b$  pairwise disjoint embedded arcs  $S = \{S_1, \dots, S_b\}$  whose endpoints are  $\{p_1, \dots, p_{2b}\}$  (referred to as a *shadow diagram*). In [\[32\]](#), this collection of cut system curves together with the shadow arcs is referred to as a “curve-and-arc system”.

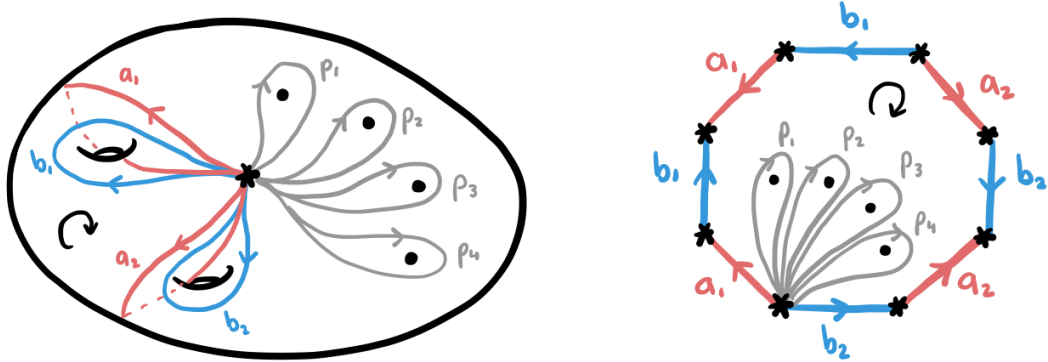


Figure 2: The genus  $g$  oriented surface  $\Sigma_g$  with basepoint  $*$  and marked points  $p_1, \dots, p_{2b}$ , represented two ways. Gluing the edges of the polygon on the right gives the surface on the left. The fundamental group  $\pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *)$  is generated by  $a_i, b_i$  and (abusing notation)  $p_i$ .

The handlebody  $H$  can be constructed from the cut system by taking  $\Sigma_g \times [0, 1]$ , attaching  $g$  3-dimensional 2-handles to  $\Sigma_g \times \{1\}$  along all of the curves  $C_1, \dots, C_g$ , and attaching a 3-ball to the resulting 2-sphere boundary component. Given a shadow diagram in addition to the cut system, a trivial tangle in the resulting handlebody can be constructed by taking the arcs of the tangle to be the union of  $\{p_1, \dots, p_{2b}\} \times [0, \frac{1}{2}]$  together with the arcs  $S_1, \dots, S_b$  in the shadow diagram considered as arcs in  $\Sigma_g \times \{\frac{1}{2}\}$ . Conversely, given a handlebody  $H$  and a trivial tangle  $T$  we can obtain a diagram  $\mathcal{D} = (C, S)$  for  $(H, T)$  by choosing a set of  $g$  disjoint embedded disks in  $H$  that cut  $H$  into a 3-ball and letting the cut system  $C$  be the boundary of these disks, and taking an isotopy relative to the boundary of  $T$  into  $\Sigma_g$  and letting the shadow diagram  $S$  denote the end result of this isotopy.

We now give a name to these disks, as well as a few other disks referred to in some of the following proofs. Given a handlebody  $H$  and a trivial tangle  $T$ , we refer to disjoint properly embedded disks bounded in  $H$  by its cut system curves as *cut disks*, and disjoint properly embedded disks that are the endpoint union of a shadow arc and the associated tangle component as *bridge disks*. One choice for these bridge disks is the track  $S_i \times [0, \frac{1}{2}]$ . A *bubble disk* in  $H - T$  is a properly embedded disk which encloses a bridge disk of the tangle strand; see Figure 3.

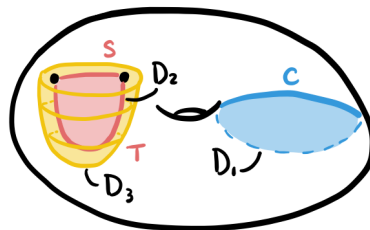


Figure 3: A cut disk  $D_1$  bounded by a cut system curve  $C$ , a bridge disk  $D_2$  bounded by a shadow arc  $S$  and tangle strand  $T$ , and a bubble disk  $D_3$ .

## 2A Bounding homomorphisms

The following definition is motivated by [Lemma 2.4](#), and examples are given in [Examples 2.5](#) and [2.6](#).

**Definition 2.1** (bounding homomorphism) A *bounding homomorphism* is an epimorphism from a (possibly) punctured surface group to a free group

$$\phi: \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \twoheadrightarrow \langle t_1, \dots, t_b, h_1, \dots, h_g \rangle$$

with the following properties:

- (1) The image of the subgroup generated by  $a_1, b_1, \dots, a_g, b_g$  surjects onto the quotient obtained by setting the  $t_i = 1$ .
- (2) Each  $p_i$  maps to a conjugate of one of the  $t_j$ , where each of  $t_j$  and its inverse  $t_j^{-1}$  appears exactly once as the central letter. More precisely, there exists a bijection

$$f: \{p_1, \dots, p_{2b}\} \rightarrow \{t_1, t_1^{-1}, \dots, t_b, t_b^{-1}\}$$

and there are group elements  $g_i \in \langle t_1, \dots, t_b, h_1, \dots, h_g \rangle$  with  $\phi(p_i) = g_i f(p_i) g_i^{-1}$ .

There are two special cases worth mentioning. If  $b = 0$ , this is an epimorphism from a closed surface group to a free group. Topologically, this will correspond to having no tangle strands. If  $g = 0$ , this corresponds to a trivial tangle in the 3-ball. The necessity of properties (1) and (2) will be seen in the proof of [Theorem 2.10](#), but, roughly speaking, property (1) will allow us to distinguish between the handlebody and tangle, and property (2) is a natural condition coming from the proof of [Lemma 2.4](#).

**Definition 2.2** (topological realization) A *topological realization* of a bounding homomorphism  $\phi: \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \twoheadrightarrow \langle t_1, \dots, t_b, h_1, \dots, h_g \rangle$  is a trivial tangle  $T$  in a handlebody  $H$  with  $\partial H = \Sigma_g$  and  $\partial T = \{p_1, \dots, p_{2b}\}$  such that there is an isomorphism  $\psi: \pi_1(H - T, *) \xrightarrow{\cong} \langle t_1, \dots, t_b, h_1, \dots, h_g \rangle$  that makes the diagram

$$\begin{array}{ccc} & & \pi_1(H - T, *) \\ & \nearrow \iota_* & \cong \downarrow \psi \\ \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) & & \langle t_1, \dots, t_b, h_1, \dots, h_g \rangle \\ & \searrow \phi & \end{array}$$

commute, where the map  $\iota_*: \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \twoheadrightarrow \pi_1(H - T, *)$  is induced by inclusion.

**Lemma 2.3** (uniqueness of realization) Let  $(H_1, T_1)$  and  $(H_2, T_2)$  be two trivial tangles in handlebodies with  $\partial H_1 = \partial H_2 = \Sigma_g$  and  $\partial T_1 = \partial T_2 = \{p_1, \dots, p_{2b}\}$  such that there is an isomorphism  $\rho$  between the fundamental groups of the tangle complements which makes the diagram

$$\begin{array}{ccc} & & \pi_1(H_1 - T_1, *) \\ & \nearrow \iota_{1*} & \cong \downarrow \rho \\ \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) & & \pi_1(H_2 - T_2, *) \\ & \searrow \iota_{2*} & \end{array}$$

commute, where the maps  $\iota_{i*} : \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \rightarrow \pi_1(H_i - T_i, *)$  for  $i = 1, 2$  are induced by inclusion. Then  $(H_1, T_1)$  and  $(H_2, T_2)$  are equivalent.

**Proof** We will construct a diffeomorphism  $H_1 \rightarrow H_2$  mapping  $T_1$  to  $T_2$  that extends the identity on the boundary  $\partial H_1$ , by defining it first on 2-cells in the complement of  $T_1$  in  $H_1$ , and then extending over 3-balls.

Fix a cut system for the handlebody  $H_1$  and shadow arcs for the tangle  $T_1$ , and choose whiskers connecting each curve and arc to the basepoint. Let  $\lambda$  be a based cut system curve which thus bounds a cut disk in  $H_1$ . From commutativity of the diagram, we know that  $\lambda$  is homotopically trivial in the tangle complement  $H_2 - T_2$ , and so by Dehn's lemma it bounds an embedded disk in  $H_2 - T_2$ . Extend the identity on the boundary to the cut disk bounded by  $\lambda$  in  $H_1 - T_1$  by mapping it to the disk obtained by Dehn's lemma in  $H_2 - T_2$ . In the same manner, extend the map to a complete system of cut disks for the handlebody  $H_1$ .

Let  $\eta$  be the based boundary of a closed tubular neighborhood of one of the shadow arcs of  $T_1$ . This curve bounds a bubble disk in  $H_1 - T_1$ ; recall Figure 3. Again from commutativity of the diagram,  $\eta$  is null-homotopic in  $H_2 - T_2$ , and, by another application of Dehn's lemma, bounds a disk. Use these disks to extend the map over all of the bubble disks in  $H_1 - T_1$ .

To finish the construction, use the Alexander trick to extend the map over the 3-cells in  $H_1$ . Observe that each of the bubble disks cuts off a single tangle strand on one of its sides. Combined with the observation that there is a unique trivial 1-strand tangle in the 3-ball, this shows that the diffeomorphism  $H_1 \rightarrow H_2$  we constructed can be arranged to map  $T_1$  to  $T_2$ . □

Now we will set up some notation in preparation for the following lemma. Let  $\mathcal{D} = (C, S)$  be a diagram for a handlebody and trivial tangle  $(H, T)$  with  $\partial H = \Sigma_g$  and  $\partial T = \{p_1, \dots, p_{2b}\}$ , together with an ordering of the curves in the cut system  $C_1, \dots, C_g$ , an ordering of the arcs in the shadow diagram  $S_1, \dots, S_b$ , and a choice of an orientation for each of the  $C_i$  and  $S_j$ . Observe that cutting the surface  $\Sigma_g$  along the cut system curves and shadow arcs creates a connected, planar surface, and thus we will be able to choose dual loops  $h_i$  and  $t_j$  as follows.

For each curve  $C_i$ , pick a closed loop  $h_i$  based at  $*$  that does not intersect the arcs  $S_k$  for any  $k$  or the curves  $C_k$  for  $k \neq i$ , and that intersects the curve  $C_i$  in a single point. Orient  $h_i$  so that, at the point of intersection of  $h_i$  and  $C_i$ , the orientation of  $h_i$  followed by the orientation of  $C_i$  agrees with the ambient orientation of  $\Sigma_g$  (which we have assumed to be clockwise; see Figure 2).

Similarly, for each arc  $S_j$ , choose a loop  $t_j$  based at  $*$  that intersects  $S_j$  in exactly one point and does not intersect any of the other arcs or curves. Orient  $t_j$  so that, at the point of intersection of  $t_j$  and  $S_j$ , the orientation of  $t_j$  followed by the orientation of  $S_j$  agrees with the ambient orientation of  $\Sigma_g$ . Therefore, we now have  $h_i, t_j \in \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *)$ . See Figure 4.

**Lemma 2.4** (diagrams to maps) *Using the notation defined above, we have the following:*

- (1) *The loops representing  $h_i$  and  $t_j$  are well defined in  $\pi_1(H - T, *)$ , independent of choices.*

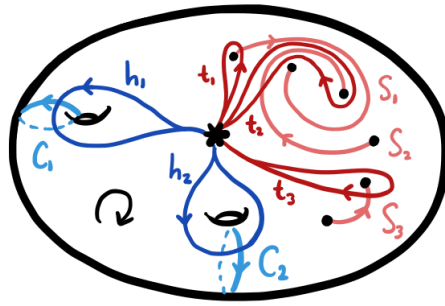


Figure 4: Finding loops  $h_i$  dual to the cut system curves  $C_i$ , and loops  $t_j$  dual to the shadow arcs  $S_j$ .

(2) The map

$$\psi_{\mathbb{Q}}: \pi_1(H - T, *) \rightarrow \langle t_1, \dots, t_b, h_1, \dots, h_g \rangle, \quad h_i \mapsto h_i, \quad t_j \mapsto t_j,$$

is an isomorphism.

(3) The composition of the map induced by inclusion and  $\psi_{\mathbb{Q}}$ ,

$$\phi: \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \xrightarrow{L^*} \pi_1(H - T, *) \xrightarrow{\psi_{\mathbb{Q}}} \langle t_1, \dots, t_b, h_1, \dots, h_g \rangle,$$

on an element  $\lambda \in \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *)$  is computed as follows. Represent  $\lambda$  by a based, closed, immersed curve on  $\Sigma_g$  that is transverse to the cut system curves  $C_i$  and the shadow arcs  $S_i$ . We call this curve  $\lambda$  again. The image of  $\lambda$  under this composition of maps is given by traversing  $\lambda$  and building a word in the elements  $t_1, \dots, t_b, h_1, \dots, h_g$  and their inverses as follows. Start with the empty word. For each intersection between  $C_i$  and  $\lambda$  we concatenate  $h_i^{\pm 1}$  on the right, and for each intersection between  $S_j$  and  $\lambda$  we concatenate  $t_j^{\pm 1}$  on the right, where the sign is determined by the sign of the intersection of the oriented curves as in Figure 5.

(4) The map  $\phi$  in (3) is a bounding homomorphism, and, with the choice of isomorphism  $\psi_{\mathbb{Q}}$ , the trivial tangle  $(H, T)$  is a topological realization of  $\phi$ .

**Proof** (1) Observe that the choices of these elements  $h_i$  and  $t_j$  are not unique when considered as elements in  $\pi_1(\Sigma_g, *)$ ; see Figure 6. However we now prove that they are unique in the group  $\pi_1(H - T, *)$ . By choosing a collection of disjoint cut disks and bridge disks, and cutting  $H - T$  along these disks, we obtain a 3-ball as in Figure 7. From this it follows that the choices of  $h_i$  and  $t_j$  are unique as elements of  $\pi_1(H - T, *)$ , because there is a unique homotopy class of curves connecting points in a 3-ball.



Figure 5: Sign convention for intersection points, where the pink arrow is a cut system curve  $C_i$  or shadow arc  $S_i$ , and the gray arrow is an element in the fundamental group of the punctured surface, represented by a based, closed immersed curve  $\lambda$  on the surface.

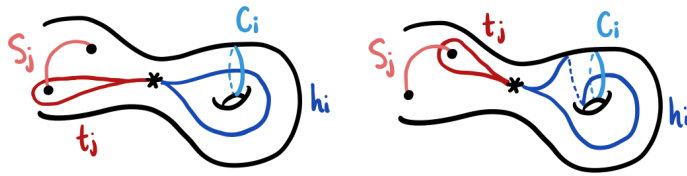


Figure 6: Inequivalent choices of  $h_i$  and  $t_j$  in the surface minus the points, which become isotopic in the handlebody minus the tangle determined by the  $C_i$  and  $S_j$ .

(2) Now define the map

$$\langle t_1, \dots, t_b, h_1, \dots, h_g \rangle \rightarrow \pi_1(H - T, *)$$

by sending  $t_j \mapsto t_j$  and  $h_i \mapsto h_i$ . We now argue that this map is an isomorphism. First observe that  $\pi_1(H - T, *)$  is a free group of rank  $g + b$ , because  $H - T$  deformation retracts onto a spine obtained in the following way. As above, cutting along (a choice of) cut disks and bridge disks results in a 3-ball, and thus the  $h_i$  and  $t_j$  make up a spine for the tangle complement  $H - T$ . This also means that the homomorphism above is surjective, and, since free groups are Hopfian, it must be an isomorphism.

(3) Using the spine from the proof of part (2), we apply the cut system–spine duality from [22] to see that the image of an element  $\lambda \in \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *)$  under these maps can be computed by recording intersections of  $\lambda$  with the cut disks and bridge disks, since traversing an edge in the spine corresponds to hitting its dual disk. For curves  $\lambda$  that live on the surface  $\Sigma_g$ , these intersections occur on the boundaries of the disks. See Figure 7.

(4) To check the first condition for a bounding homomorphism, we glue 2-handles to the meridians of the tangle strands to kill the normal closure of the generators  $t_i$ , and observe that the  $h_i$  make up a spine for the resulting handlebody.

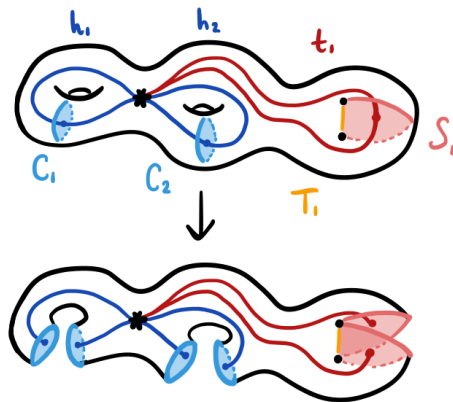


Figure 7: Cutting along cut disks (bounded by cut system curves  $C_i$ ) and bridge disks (bounded by shadow arcs  $S_i$  and tangles  $T_i$ ) gives a 3-ball. The figure also shows how to build a spine of the tangle complement in the handlebody.

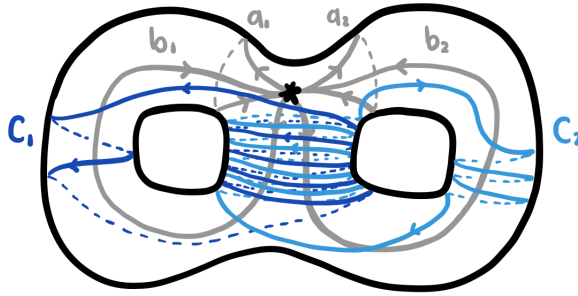


Figure 8: One of the handlebodies in a genus 2 Heegaard splitting of the Poincaré homology sphere, where the other handlebody is the solid genus 2 handlebody filling the interior. The generators  $a_1, b_1, a_2, b_2$  of the fundamental group of the surface are shown in gray, and the two curves  $C_1, C_2$  of the cut system are shown in light and dark blue.

For the second condition, we represent the generators  $p_i \in \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *)$  in the following way. Choose a simple system of rays (or whiskers)  $r_i$  connecting  $*$  to each point in  $\{p_1, \dots, p_{2b}\}$ . Then  $p_i$  corresponds to a loop that runs out along  $r_i$ , goes around the puncture  $p_i$  in a negative direction, as dictated by Figure 2, and returns along  $r_i$ . On punctured surfaces this is also known as a Hurwitz arc system [25, Section 2.3]. The sequence in which the whisker of the loop around  $p_i$  intersects the cut system and tangle shadows will read off a word  $g_i$  in the free group. Then running around the puncture reads off one of the generators  $t_i$ , followed by returning to  $*$  along  $r_i$  contributing  $g_i^{-1}$ . □

**Example 2.5** (handlebody with empty tangle) Here we give an example for the case  $b = 0$ . In this case the conditions on a bounding homomorphism  $\phi: \pi_1(\Sigma_g, *) \twoheadrightarrow \langle h_1, \dots, h_g \rangle$  ensure that it is an epimorphism from a surface group to a free group. Consider the genus 2 handlebody shown in Figure 8, which is one of the handlebodies in a genus 2 Heegaard splitting of the Poincaré homology sphere. The bounding homomorphism for this handlebody can be read off by recording the sequence of intersections of the generators  $a_i$  and  $b_i$  with the curves  $C_j$ :

$$\pi_1(\Sigma_2, *) = \langle a_1, b_1, a_2, b_2 \mid [a_1, b_1][a_2, b_2] = 1 \rangle \twoheadrightarrow \langle h_1, h_2 \rangle$$

is given by

$$a_1 \mapsto h_1^{-1}, \quad b_1 \mapsto (h_1 h_2)^5 h_1^{-2}, \quad a_2 \mapsto (h_1 h_2)^5 h_2^3, \quad b_2 \mapsto h_2.$$

**Example 2.6** (running example) The following map is a bounding homomorphism which is realized by a trivial 2-bridge tangle in a solid genus 1 handlebody. It appears as the green tangle in the bridge trisection of  $\mathbb{R}P^2$  in  $\mathbb{C}P^2$  from [32, Figure 2; 23, Figure 3]. See Figure 9. We will use this bounding homomorphism as a running example in the proof of Theorem 2.10:

$$\pi_1(\Sigma_1 - \{p_1, \dots, p_4\}, *) = \langle p_1, p_2, p_3, p_4, a_1, b_1 \mid p_1 p_2 p_3 p_4 = [a_1, b_1] \rangle \twoheadrightarrow \langle t_1, t_2, h_1 \rangle$$

is given by

$$p_1 \mapsto t_2 h_1 t_1 h_1^{-1} t_2^{-1}, \quad p_2 \mapsto t_2, \quad p_3 \mapsto h_1 t_1^{-1} h_1^{-1}, \quad p_4 \mapsto h_1 t_2^{-1} h_1^{-1}, \quad a_1 \mapsto t_2 h_1, \quad b_1 \mapsto h_1.$$

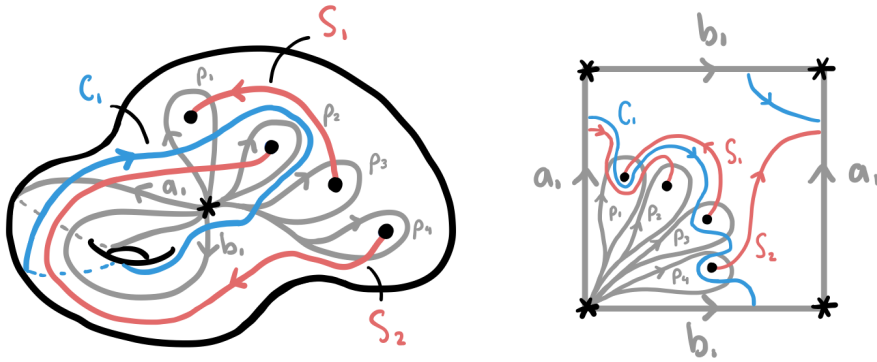


Figure 9: A trivial 2-bridge tangle in a solid genus 1 handlebody, which appears as the green tangle in the bridge trisection of  $\mathbb{R}P^2$  in  $\mathbb{C}P^2$  from [32, Figure 2; 23, Figure 3].

### 2B Stallings folding

We now discuss a technique, due to Stallings, called folding [41], which in our context will be used to give a topological realization of any bounding map. There are several applications of folding in the study of finitely generated free groups (eg for the membership problem or determining the index and normality of a subgroup; see [5]). However, for our purposes we only need one application, namely that folding gives a convenient algorithmic method to determine if a set of elements  $w_1, \dots, w_k \in F_n$  generate  $F_n$ , where  $F_n$  is the free group generated by the elements  $x_1, \dots, x_n$ .

We now describe this algorithm. We begin by forming a directed graph  $\Gamma$  with edges labeled by elements of  $\{x_1, \dots, x_n\}$ , where  $\Gamma$  is topologically a wedge of  $k$  circles and each of the circles is subdivided and labeled according to the words  $w_1, \dots, w_k$  as in Figure 10. We can change  $\Gamma$  by a move called a fold to obtain a new such graph.

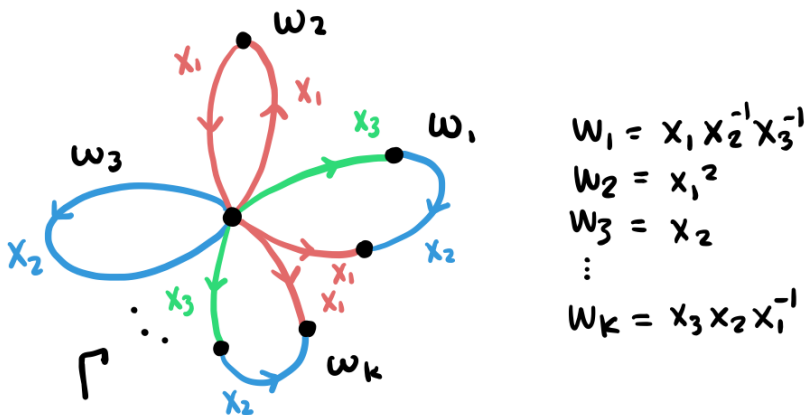


Figure 10: A directed graph  $\Gamma$  with edges labeled by elements of  $\{x_1, \dots, x_n\}$ , where  $\Gamma$  is topologically a wedge of  $k$  circles and each of the circles is subdivided and labeled according to the words  $w_1, \dots, w_k$ .

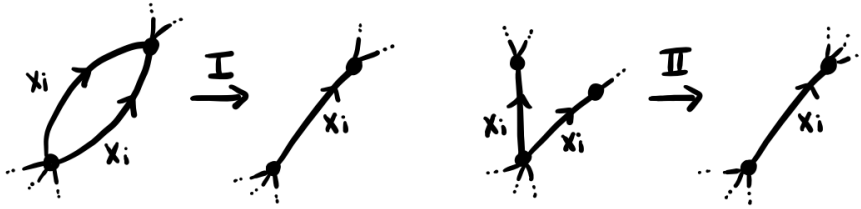


Figure 11: Types of folds. The labels and orientations of the edges being folded must match.

**Definition 2.7** (fold) A (Stallings) fold is a move on a labeled, directed graph which takes two edges and replaces them with one single edge, with the following restrictions. The original edges must

- (1) have the same label,
- (2) share a vertex, and
- (3) be oriented either both in or both out of the shared vertex.

The label and orientation of the new edge are induced by those of the original edges. If the original edges share a second vertex, we call this a *type I fold*, and if not, a *type II fold*. In the case of the type II fold, the unshared vertices are identified together after replacing the original edges with the new edge. See Figure 11.

**Lemma 2.8** (Stallings) The elements  $w_1, \dots, w_k \in F_n$  generate  $F_n$  if and only if there exists a sequence of folds beginning with the graph  $\Gamma$  given by  $w_1, \dots, w_k$  as in Figure 10, and any such sequence of folds terminates in the graph  $R_n$  in Figure 12.

The proof of Lemma 2.8 follows as a special case of [5, Theorem 4.7]. Before we begin the proof of our main technical theorem, we mention one last lemma that will be used to check that a given set of closed curves constitutes a cut system.

**Lemma 2.9** Let  $w_1, \dots, w_k$  be elements in a free group  $F_n$  with free generating set  $x_1, \dots, x_n$  such that  $w_1, \dots, w_k$  generate  $F_n$ . Let  $\text{exp}_{x_i}: F_n \rightarrow \mathbb{Z}$  denote the exponent sum homomorphism, that is, the

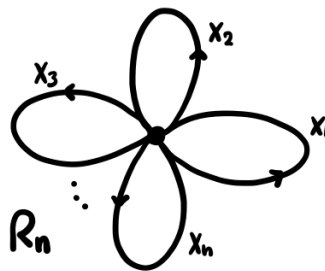


Figure 12: The directed graph  $R_n$  is topologically a wedge of  $n$  circles, with each circle uniquely labeled by an element  $x_i$ .

signed count of occurrences of the letter  $x_i$ . Then the  $n$  vectors

$$v_i = (\exp_{x_i}(w_1), \exp_{x_i}(w_2), \dots, \exp_{x_i}(w_k)) \in \mathbb{Z}^k$$

are linearly independent in  $\mathbb{Z}^k$ .

**Proof** Consider the  $n \times k$  matrix where the rows are given by the vectors  $v_i$ . The  $j^{\text{th}}$  column is made up of all of the exponent sums of the word  $w_j$ , and thus computes the image of  $w_j$  under the abelianization map  $\text{ab}: F_n \rightarrow \mathbb{Z}^n$ , where the images of the generators  $x_i$  form the basis of the codomain. Since the words  $w_1, \dots, w_k$  generate  $F_n$ , the columns of the matrix generate  $\mathbb{Z}^n$  and thus its column rank is  $n$ . The claim now follows from the equality of row and column rank and the observation that  $n \leq k$ .  $\square$

**Theorem 2.10** (existence of realization) *For every bounding homomorphism*

$$\phi: \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \twoheadrightarrow \langle t_1, \dots, t_b, h_1, \dots, h_g \rangle,$$

*there exists a topological realization (where the realizing tangle  $(H, T)$  is unique by Lemma 2.3) and, further, we give a (polynomial-time) algorithm to construct a diagram for the topological realization.*

**Proof** The plan is to describe a topological realization of  $\phi$  by an explicit diagram that we will produce in such a way that Lemma 2.4 ensures that it is indeed a topological realization of  $\phi$ . In the **first stage** we produce a preliminary diagram  $\mathfrak{D}$  that, if it were to consist of only a cut system and a shadow diagram, would be a realization of  $\phi$ . This preliminary diagram will not necessarily be unique, and could possibly include “extra” components. After this, in the **second stage** we will apply Stallings folding from Lemma 2.8 to guide band sums which eliminate any extra components of the preliminary diagram. At this point, we will have obtained a unique diagram, regardless of the choices made while producing the preliminary diagram. Finally, in the **third stage** we argue why the necessary bands always exist.

**First stage (preliminary diagram)** For the first stage, it is necessary to initially make the distinction between words in the generators and inverses of generators in a free group, and elements of the free group. For now, when we write  $\phi(p_i)$ ,  $\phi(a_j)$ ,  $\phi(b_k)$ , we mean the unique freely reduced words representing these elements. To produce  $\mathfrak{D}$ , look at two (not necessarily freely reduced) words, the first given by concatenating the freely reduced  $\phi(p_i)$ , namely

$$w_1 = \phi(p_1)\phi(p_2) \cdots \phi(p_{2b}),$$

and the second given by concatenating the freely reduced  $\phi(a_i)$ ,  $\phi(b_j)$ , namely

$$w_2 = \phi(a_1)\phi(b_1) \cdots \phi(a_g)\phi(b_g).$$

Note that since  $\phi$  is a homomorphism these words are equal as elements of  $\langle t_1, \dots, t_b, h_1, \dots, h_g \rangle$ .

The first step in drawing the preliminary diagram  $\mathfrak{D}$  is to represent  $\Sigma_g$  by a polygon with edges  $a_1, b_1, \dots, a_g, b_g$  and punctures as in Figure 13, which will begin our main running example for the proof. (If  $g = 0$ , view the sphere as the plane with a point at infinity, place all of the punctures in a

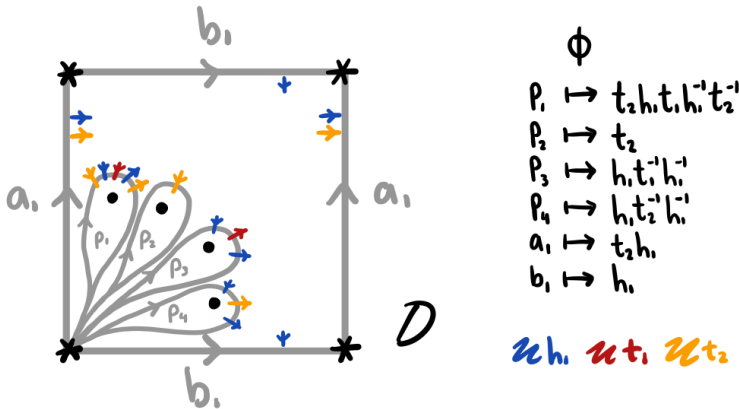


Figure 13: Our main running example for this proof comes from the bounding homomorphism in Example 2.6. Represent  $\Sigma_1 = T^2$  by a polygon with punctures. Mark each of the circles representing  $a_1$  and  $b_1$  with “oriented dashes” labeled (using color) by the respective elements in  $\phi(a_1)$  and  $\phi(b_1)$ . Additionally, mark the loops around each punctured point  $p_i$  with “oriented dashes” labeled by the respective elements in  $\phi(p_1), \dots, \phi(p_4)$ . Note that for the oriented dashes we use colors corresponding to  $h_1, t_1$  and  $t_2$  to see this correspondence easily, but keep in mind that throughout this running example we are really recovering the curves  $C_1, S_1$  and  $S_2$ , which is why the colors in our end result, Figure 22, differ slightly from those in Example 2.6.

line, and place the basepoint at infinity. See [3, Section 4.2.3] for an example of this.) Mark each of the circles representing  $a_1, b_1, \dots, a_g, b_g$  on  $\Sigma_g$  with “oriented dashes” labeled by the respective elements in  $\phi(a_1), \phi(b_1), \dots, \phi(a_g), \phi(b_g)$ . Additionally, mark the loops around each punctured point  $p_i$  on  $\Sigma_g$  with “oriented dashes” labeled by the respective elements in  $\phi(p_1), \dots, \phi(p_{2b})$ .

The second step is to freely reduce both of the words  $w_1$  and  $w_2$ , and, as cancellations occur in the free reductions, draw arcs between the corresponding dashes as in Figure 14. The arcs retain the respective

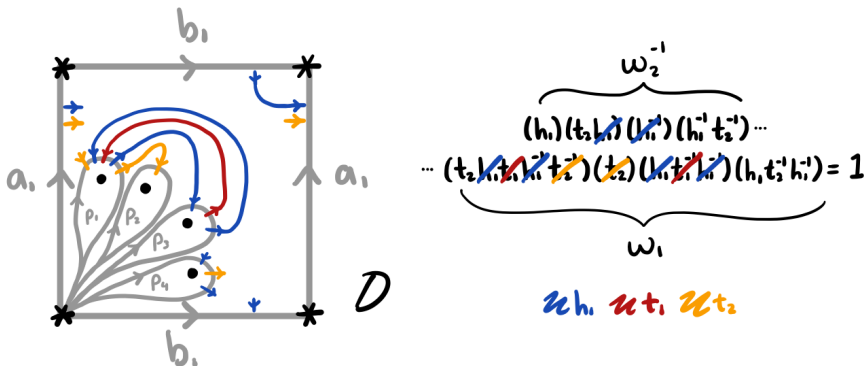


Figure 14: Running example. Drawing arcs between the dashes as corresponding cancellations occur in the free-reductions of the words  $w_1$  and  $w_2$ . There are multiple ways to freely reduce these words; this running example shows one possible choice. In fact, there exists a different choice of cancellations here which would avoid the need to proceed to the second stage.

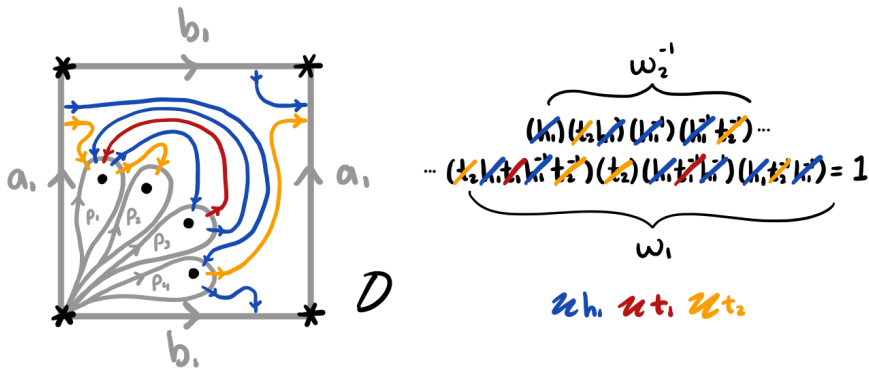


Figure 15: Running example. Continuing to draw arcs between dashes until all cancellations have occurred.

labelings and have orientations induced by the dashes. Let  $w'_1$  and  $w'_2$  denote the resulting freely reduced words, which are equal as freely reduced words since they are equal as elements of the free group. Therefore  $(w'_2)^{-1}w'_1$  freely reduces to yield the trivial word.

The final step is to continue to carry out this free reduction down to the trivial word, drawing arcs with each cancellation as in the second step. See Figure 15. This reduction will not necessarily be unique and could produce different diagrams, but this indeterminacy will be fixed in the second stage. The loops around the punctures each intersect an odd number of dashes, which are now each part of an arc. Connect the middle dash to the punctured point, and connect the rest of the dashes in pairs that go around the opposite side of the puncture, as in Figure 16.

We consider the preliminary diagram  $\mathcal{D}$  to be the resulting collection of disjoint, oriented arcs and closed curves on  $\Sigma_g$ , each labeled by a generator of  $\langle t_1, \dots, t_b, h_1, \dots, h_g \rangle$ . In general  $\mathcal{D}$  at this stage will consist of too many closed loops and will not give a realization of  $\phi$ . Note that by condition (1) in the definition of a bounding homomorphism, there is at least one closed curve with each label  $h_j$ . Similarly, by condition (2), there is exactly one arc with each label  $t_i$ . However, in general  $\mathcal{D}$  will consist of additional closed curves with labels  $h_j$  and  $t_j$ .

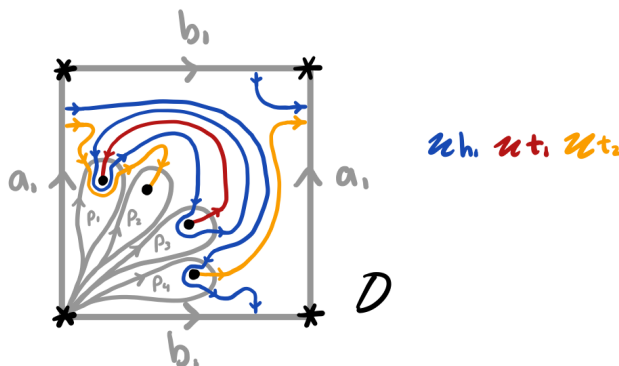


Figure 16: Running example. Connecting the middle dash to the puncture and the other dashes in pairs.

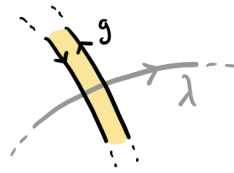


Figure 17: Following the curve  $a$  running across a band, we read off the canceling pair of intersections  $gg^{-1}$ .

Suppose that at this stage there are no such additional curves in  $\mathcal{D}$ ; namely, for each  $h_i$  there is exactly one closed curve with that label and for each  $t_j$  there are no closed curves with that label. We now show that the curves labeled by the elements  $h_i$  form a cut system, namely that they are homologically linearly independent.

Let  $C_1, \dots, C_g$  denote the closed curves in this case, with corresponding labels  $h_1, \dots, h_g$ , respectively. The curves  $a_1, b_1, \dots, a_g, b_g$  in Figure 2 give a basis for  $H_1(\Sigma_g; \mathbb{Z})$ . Using that the intersection product on  $H_1(\Sigma_g; \mathbb{Z})$  is given by  $a_i \cdot b_j = \delta_{ij}$ , we find that, by construction, in  $H_1(\Sigma_g; \mathbb{Z})$  we have

$$C_i = \exp_{h_i}(\phi(b_1))a_1 + \exp_{h_i}(\phi(a_1))b_1 + \dots + \exp_{h_i}(\phi(b_g))a_g + \exp_{h_i}(\phi(a_g))b_g,$$

where

$$\exp_{h_i}: \langle t_1, \dots, t_b, h_1, \dots, h_g \rangle \rightarrow \mathbb{Z}$$

maps an element to the exponent sum of  $h_i$ . By property (1) of being a bounding homomorphism, together with Lemma 2.9, it follows that the elements of  $C_i$  are linearly independent in  $H_1(\Sigma_g; \mathbb{Z})$  and therefore in this case the curves  $C_1, \dots, C_g$  form a cut system.

From this it follows that, in this case,  $\mathcal{D}$  is a diagram for a handlebody and trivial tangle. Furthermore, in this case, we know by Lemma 2.4 that  $\mathcal{D}$  is a diagram for a topological realization of  $\phi$ .

**Second stage (Stallings folding)** In the second stage, we will modify the preliminary diagram  $\mathcal{D}$  in steps by performing orientation-preserving band sums between components in order to eliminate the extra closed curves, as in Figure 22. These bands may pass through the  $a_i, b_j$  and  $p_k$  representing the elements in  $\pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *)$ , but may not pass through the curves and arcs we drew in the previous stage. Assuming for a moment that we successfully banded together all of the curves and arcs, so that there is exactly one curve/arc with each label, the above argument still applies and the resulting diagram will be a diagram for a cut system and trivial tangle (namely, it will consist of a cut system together with a shadow diagram) since each band introduces a pair of canceling intersections as in Figure 17. Furthermore, because the added intersections cancel, applying Lemma 2.4 shows that the resulting diagram is in fact a diagram for a topological realization of  $\phi$ .

The algorithm for finding these bands is as follows. Note that, here and beyond, we use the word *color* to describe the “labels” as mentioned in Definition 2.7, and give the word *label* a new, specific meaning. (In particular, we use *color* colloquially to mean “labeled by a color”, as opposed to the graph-theoretic

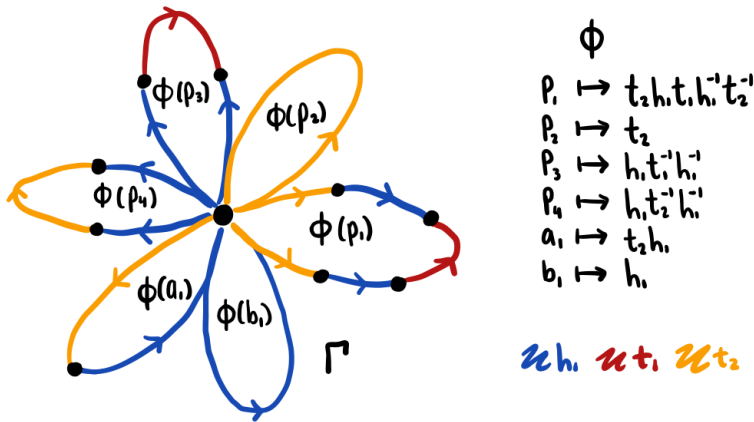


Figure 18: Running example. The graph  $\Gamma$  is topologically a wedge of circles, where each circle is colored by the words  $\phi(p_1), \dots, \phi(p_4), \phi(a_1), \phi(b_1)$ . Here colors are used to denote this coloring.

notion.) Let  $\Gamma$  be the graph that is topologically a wedge of circles such that each circle is colored by the words

$$\phi(p_1), \dots, \phi(p_{2b}), \phi(a_1), \phi(b_1), \dots, \phi(a_g), \phi(b_g)$$

as in Figure 18. Add an additional label to each edge of  $\Gamma$ , namely label each edge of  $\Gamma$  with the corresponding closed curve or arc in the preliminary diagram  $\mathcal{D}$  as in Figure 19. We will modify the graph  $\Gamma$  by folding and this will dictate how to modify the diagram  $\mathcal{D}$  by band sums. At each stage, we will refer to the new graph again by  $\Gamma$  and the new diagram again by  $\mathcal{D}$ .

Since the elements

$$\phi(p_1), \dots, \phi(p_{2b}), \phi(a_1), \phi(b_1), \dots, \phi(a_g), \phi(b_g)$$

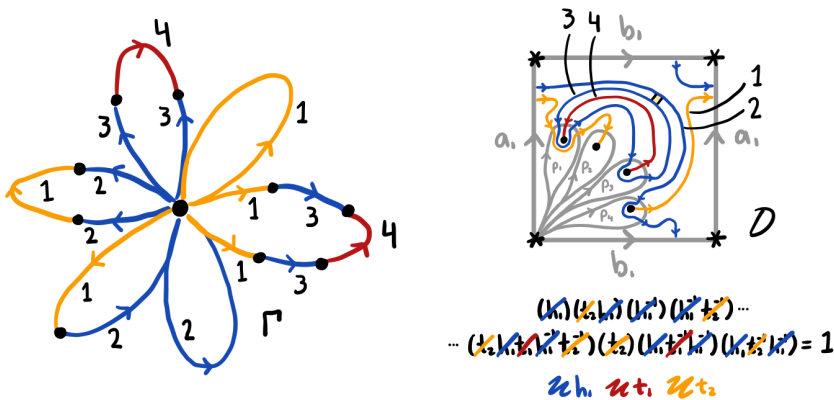


Figure 19: Running example. An additional label is added to each edge of  $\Gamma$  which represents the corresponding curve/arc in  $\mathcal{D}$ . Here numbers are used to denote these labels.

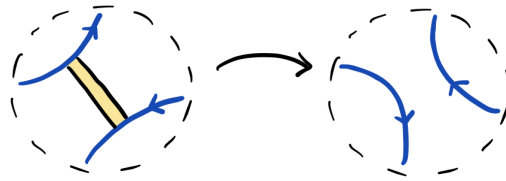


Figure 20: An orientation-preserving band sum between two curves/arcs in a neighborhood within the diagram  $\mathcal{D}$ . The curves/arcs are the same color but have different labels; that is, they are different curves/arcs in  $\mathcal{D}$  that are colored by the same word.

generate the free group  $\langle t_1, \dots, t_b, h_1, \dots, h_g \rangle$ , by Lemma 2.8 there exists a sequence of foldings of the graph  $\Gamma$  to the graph  $R_n$  in Figure 12, where  $n = b + g$  is the rank of the free group. Choose a sequence of such foldings. Recall that foldings must occur between edges with the same coloring, but not necessarily the same labels (using the language of the previous paragraph). Whenever two edges of  $\Gamma$  with the same label are folded, the diagram  $\mathcal{D}$  is left unchanged and the edge of  $\Gamma$  resulting from the fold is given the same label as the edges it came from. Whenever two edges of  $\Gamma$  with different labels are folded, an orientation-preserving band between the two corresponding curves/arcs is chosen, disjoint from all of the other curves/arcs, and the diagram  $\mathcal{D}$  is modified by performing a band sum along this band as in Figure 20. The edge of  $\Gamma$  resulting from the fold is given a new labeling that identifies the

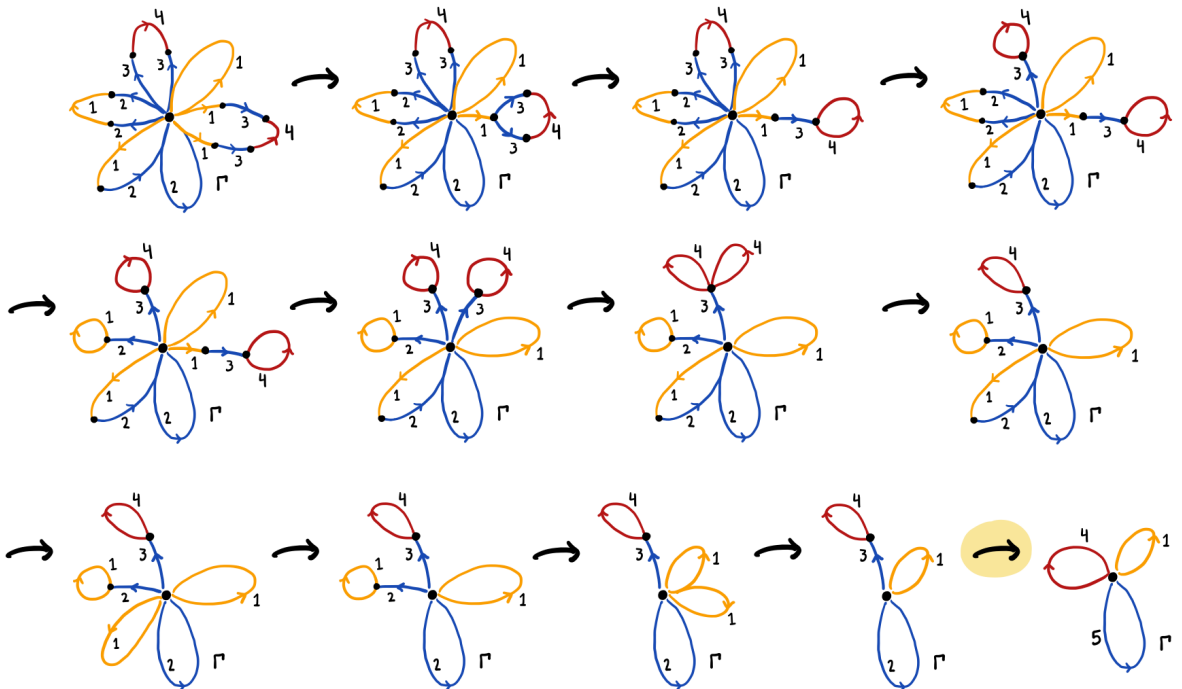


Figure 21: Running example. One possible sequence of Stallings folds. The diagram  $\mathcal{D}$  remains unchanged throughout all of the folds except the last fold (highlighted), after which a band sum is performed as in Figure 22.

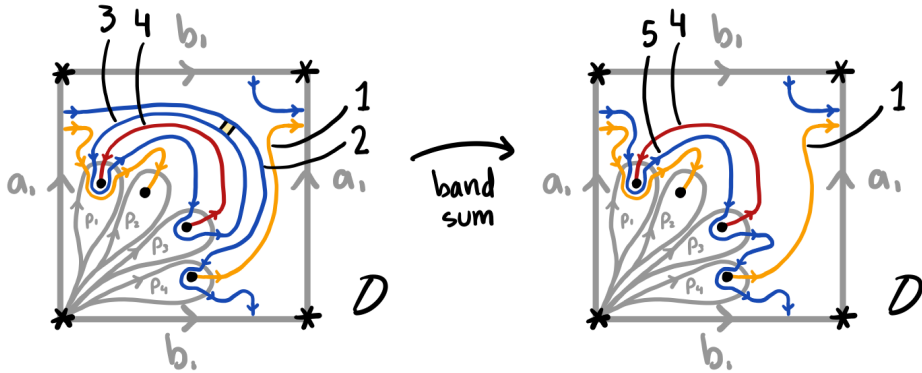


Figure 22: Running example. The curves 2 and 3 on the left are banded together to form the new curve 5 on the right. Now we are finished! In the notation of Example 2.6, the curve 5 is  $C_1$ , the arc 4 is  $S_1$ , and the arc 1 is  $S_2$ .

two banded together curves/arcs. Any other occurrences of the involved labels are modified as well. See Figure 21 for a sequence of folds in our running example, and Figure 22 for the resulting band sum. See Example 2.11 for a different example containing more complicated band sums.

Assuming for a moment that all the required bands do exist, in the final diagram there will be one curve/arc in the diagram for each edge of  $R_n$ , since the graph  $\Gamma$  folds down to  $R_n$ . It then follows that there will be exactly  $b$  arcs and  $g$  closed curves in the final diagram  $\mathcal{D}$ . The preceding discussion then applies to see that the resulting closed curves form a cut system and the diagram  $\mathcal{D}$  indeed does provide a topological realization of  $\phi$ .

**Third stage (existence of bands)** We now tackle the problem of the existence of the bands, which will follow from the following claim. Note that in the claim we are not assuming edges have the same color or label, even though our definition of folding requires edges to have the same color. Furthermore, we prove the existence of some “extra” orientation-reversing bands, which we do not need for the second stage of our proof, but we *do* need as part of our inductive argument for the claim. Thus we prove that bands exist in a more general setting, which will imply that the specific bands we want in the previous part of the proof do indeed exist.

**Claim** Given two incident edges  $e_1$  and  $e_2$  in the graph  $\Gamma$  at any state of the above procedure, there exists a band between the corresponding curves/arcs that label  $e_1$  and  $e_2$  which is disjoint from all other curves/arcs in the diagram  $\mathcal{D}$ . In particular:

- (1) If the incident edges are oriented so that they both go into their shared vertex, as in Figure 23(I), then there exists an orientation-preserving band from the right of the curve/arc labeling  $e_1$  to the right of the curve/arc labeling  $e_2$ .
- (2) If the incident edges are oriented so that they both go out of their shared vertex, as in Figure 23(II), then there exists an orientation-preserving band from the left of the curve/arc labeling  $e_1$  to the left of the curve/arc labeling  $e_2$ .

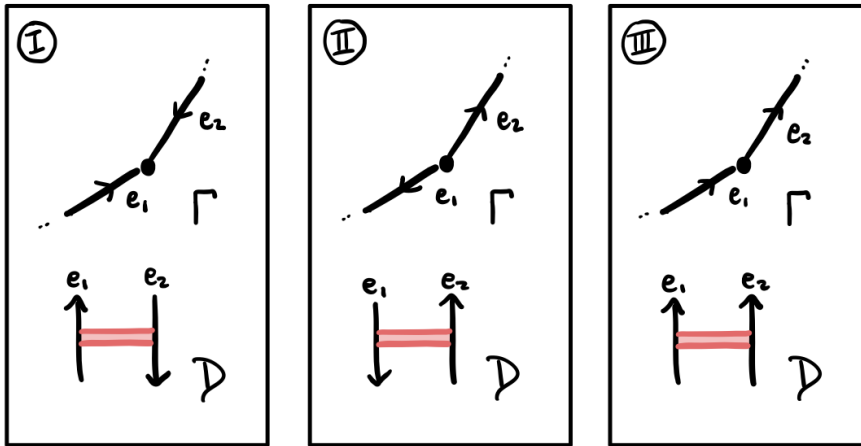


Figure 23: We consider three cases for how the edges  $e_1$  and  $e_2$  are oriented. Here  $e_1$  and  $e_2$  are the names of the edges in the graph  $\Gamma$ , not labels or colors, and we use the same notation to refer to the curves/arcs in  $\mathcal{D}$  that correspond to the edges in  $\Gamma$ .

- (3) If the incident edges are oriented so that one goes into their shared vertex and one goes out, as in Figure 23(III), then there exists an orientation-reversing band from the right of the curve/arc labeling  $e_1$  to the left of the curve/arc labeling  $e_2$ .

Here we are fixing some conventions on how the orientations of incident edges in  $\Gamma$  correspond to orientations in  $\mathcal{D}$ , which can be done without loss of generality and in such a way that the cases (I), (II) and (III) are compatible with each other. Also note that in some cases, both an orientation-preserving and an orientation-reversing band might exist between the corresponding curves/arcs. The statement of the claim only contains the existence of those bands which are necessary for the proof.

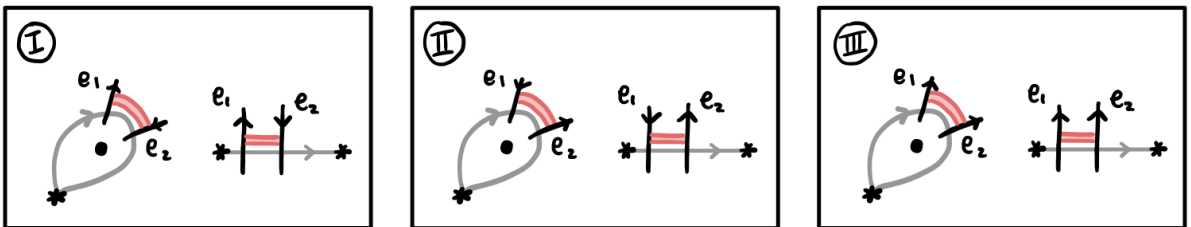


Figure 24: Some orientation checks for the base case of the claim, in which the edges  $e_1$  and  $e_2$  are in the same circle in the graph  $\Gamma$ . We abuse notation and use  $e_1$  and  $e_2$  to also refer to the oriented dashes in  $\mathcal{D}$  that correspond to these edges in  $\Gamma$ . These figures show  $e_1$  and  $e_2$  in the diagram  $\mathcal{D}$ , where  $*$  is the basepoint and the gray, unlabeled curves/arcs represent the generators  $a_i, b_i$  and  $p_i$ , in the notation of the previous stages. Here we check whether a disjoint orientation-preserving band (in pink) between  $e_1$  and  $e_2$  can be found in cases (I) and (II), and whether a disjoint orientation-reversing band (in pink) between  $e_1$  and  $e_2$  can be found in case (III).

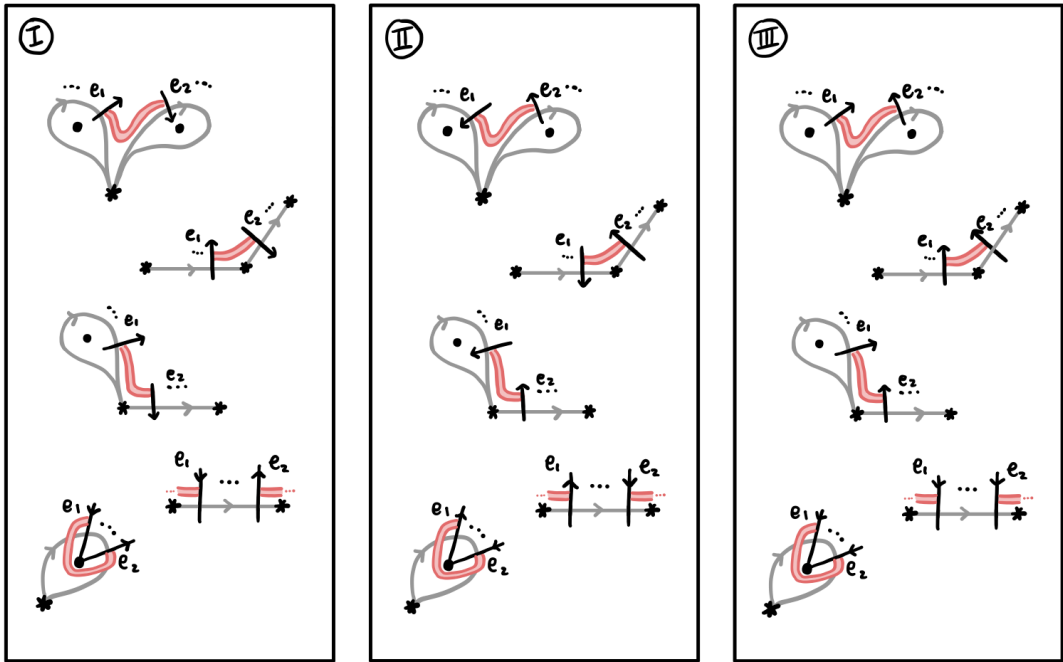


Figure 25: Some orientation checks for the base case of the claim, in which the edges  $e_1$  and  $e_2$  are in different circles in the graph  $\Gamma$ . We abuse notation and use  $e_1$  and  $e_2$  to also refer to the oriented dashes in  $\mathcal{D}$  that correspond to these edges in  $\Gamma$ . These figures show  $e_1$  and  $e_2$  in the diagram  $\mathcal{D}$ , where  $*$  is the basepoint and the gray, unlabeled curves/arcs represent the generators  $a_i, b_i$  and  $p_i$ , in the notation of the previous stages. Here we check whether a disjoint orientation-preserving band (in pink) between  $e_1$  and  $e_2$  can be found in cases (I) and (II), and whether a disjoint orientation-reversing band (in pink) between  $e_1$  and  $e_2$  can be found in case (III).

We prove the claim by induction on the number of folds that have been performed in the algorithm, checking throughout that our three cases for orientations hold. We first check that the claim is valid for the preliminary diagram  $\mathcal{D}$  before any folding has been performed. Our initial graph  $\Gamma$  is topologically a wedge of circles, and incident edges can either be in the same circle or in different circles.

If the edges  $e_1$  and  $e_2$  are in the same circle in  $\Gamma$ , then these correspond to oriented dashes in  $\mathcal{D}$  that are right next to each other (except in one case mentioned below). Thus we can draw a band between them which is disjoint from the rest of the diagram. See Figure 24 for some of the orientation checks. If the edges  $e_1$  and  $e_2$  are in different circles in  $\Gamma$ , then they are both connected to the central vertex. They are therefore labeled by the “outermost” curves/arcs in  $\mathcal{D}$  and are connected to the basepoint by an arc which is disjoint from the other curves/arcs. Thickening and joining these arcs then gives a band. See Figure 25 for some of the orientation checks. (The case where the edges are in the same circle, but both connected to the central vertex and not sharing a second vertex is included in Figure 25 rather than Figure 24. Specifically, see the bottom-most two pictures in Figure 25.)

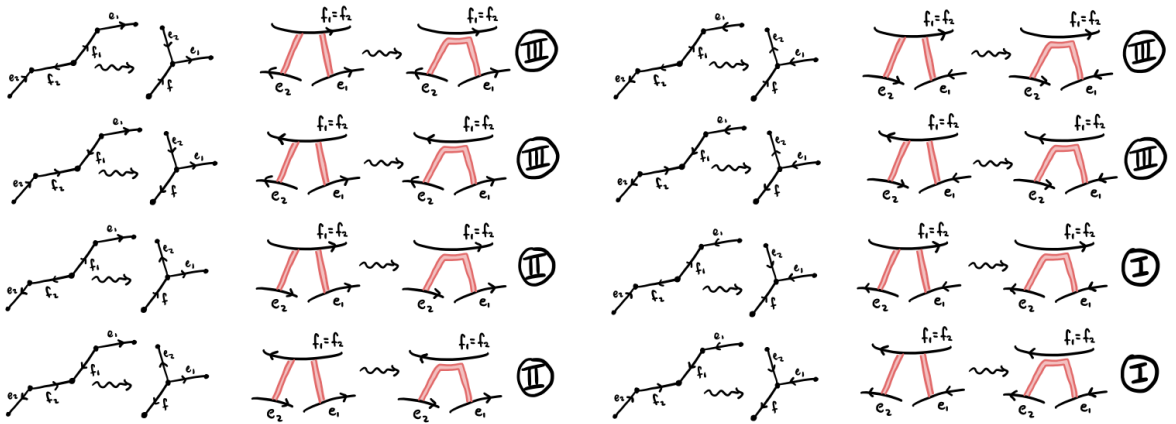


Figure 26: Orientation checks for the inductive step of the claim when the edges  $f_1$  and  $f_2$  have the same label. We abuse notation and use  $e_1, e_2, f_1, f_2$  and  $f$  to also refer to the curves/arcs in the diagram that correspond to these edges in  $\Gamma$ . These figures show the possible orientations for  $e_1, e_2, f_1, f_2$  and  $f$  relative to each other in  $\Gamma$  (on the left), and how in the corresponding diagrams (on the right), a disjoint orientation-preserving band between  $e_1$  and  $e_2$  can be found inductively when they are oriented as in cases (I) and (II), and a disjoint orientation-reversing band between  $e_1$  and  $e_2$  can be found inductively when they are oriented as in case (III). These pictures suffice for all cases in which  $e_1$  and  $e_2$  are connected to the nonshared vertex of  $f_1$  and  $f_2$ , respectively, and are not incident before the fold.

For the inductive step, we verify that the validity of the claim is preserved after folding has occurred. Let  $\mathcal{D}$  be the diagram before the fold,  $\mathcal{D}'$  be the diagram after the fold,  $f_1$  and  $f_2$  be the edges to be folded, and  $f$  be the new folded edge. We need to show that any curves/arcs that label edges that are newly incident after the fold still have a band between them. For ease of explanation, we will slightly abuse notation and use  $e_1, e_2, f_1, f_2$  and  $f$  to also refer to the curves/arcs in the diagram that are labeled by these edges.

We first handle the case where the two folded edges  $f_1$  and  $f_2$  have the same label. Note that, in this case,  $\mathcal{D} = \mathcal{D}'$ . Since  $f_1$  and  $f_2$  have the same label, they correspond to the same curve/arc in  $\mathcal{D}$ , and existing bands will suffice in all cases except those in which  $e_1$  and  $e_2$  are connected to the nonshared vertex of  $f_1$  and  $f_2$ , respectively, and are not incident before the fold. In these cases, observe that a band from  $e_1$  to  $e_2$  can be created by taking the existing band from  $e_1$  to  $f_1$ , following along  $f_1 = f_2 = f$ , and continuing along the existing band from  $f_2$  to  $e_2$ . See Figure 26 for the orientation checks.

Finally, assume that the two edges  $f_1$  and  $f_2$  have different labels. In this case, the diagram  $\mathcal{D}'$  differs from the diagram  $\mathcal{D}$  by an orientation-preserving band sum between  $f_1$  and  $f_2$ , which merges these curves/arcs into the same component  $f$ . Therefore, as in the previous case, existing bands will suffice in all cases except those in which  $e_1$  and  $e_2$  are connected to the nonshared vertex of  $f_1$  and  $f_2$ , respectively, and are not incident before the fold. In these cases, observe that a band from  $e_1$  to  $e_2$  can be created by

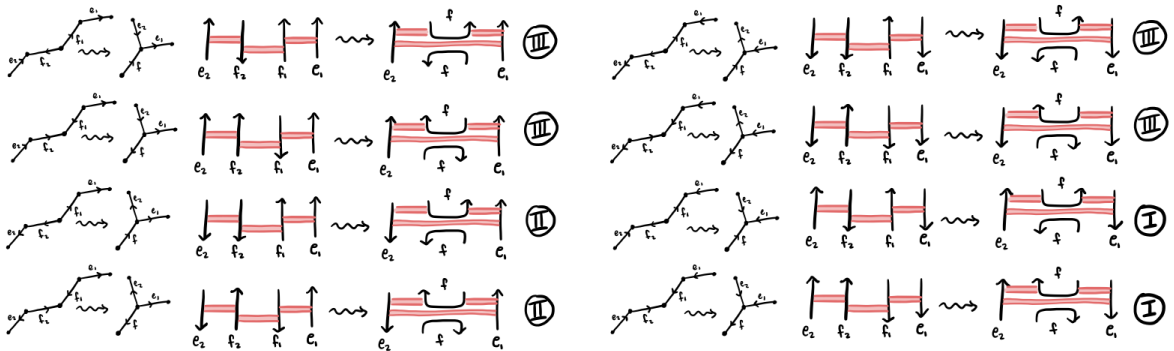


Figure 27: Orientation checks for the inductive step of the claim when the edges  $f_1$  and  $f_2$  have different labels. We abuse notation and use  $e_1, e_2, f_1, f_2$  and  $f$  to also refer to the curves/arcs in the diagram that correspond to these edges in  $\Gamma$ . These figures show the possible orientations for  $e_1, e_2, f_1, f_2$  and  $f$  relative to each other in  $\Gamma$  (on the left), and how in the corresponding diagrams (on the right), a disjoint orientation-preserving band between  $e_1$  and  $e_2$  can be found inductively when they are oriented as in cases (I) and (II), and a disjoint orientation-reversing band between  $e_1$  and  $e_2$  can be found inductively when they are oriented as in case (III). These pictures suffice for all cases in which  $e_1$  and  $e_2$  are connected to the nonshared vertex of  $f_1$  and  $f_2$ , respectively, and are not incident before the fold.

taking the existing band from  $e_1$  to  $f_1$ , following through the “tunnel” created by the band sum, and continuing along the existing band from  $f_2$  to  $e_2$ . See Figure 27 for the orientation checks. □

**Example 2.11** (more complicated band sums) In Figure 28 we present an example containing more complicated band sums (compared to our running example). Here our surface is  $S^2$ . In the top box we start with a preliminary diagram  $\mathcal{D}$  coming from a given  $\phi$  (and a choice of cancellation), and from this we produce a graph  $\Gamma$ . The middle box shows a sequence of Stallings folds, which results in three band sums (corresponding to the highlighted folds). The bottom box shows the result of the band sums in the diagram  $\mathcal{D}$ . Note that in this example there is a necessary order for the band sums: 1 and 3 cannot be banded together until 2 and 5, and then 6 and 4, are banded together.

### 3 Closed 3-manifolds and bridge split links

In this section and the next, we translate fundamental topological theorems into the algebraic setup we described in Section 2, and obtain correspondences between topology and algebra. In each subsection we take as input a topological theorem and show how this translates to algebra. Because we pass through diagrams in between topology and algebra and multiple notions of equivalence are involved, the proofs are rather technical. We include full details in Section 3A and omit some details further on, as each subsection builds from the previous and the proofs follow similarly.

As a warm-up, we begin in Section 3A with the case of closed 3-manifolds where our topological input theorem is the Reidemeister–Singer theorem. Then in Section 3B we show how link theory in 3-manifolds

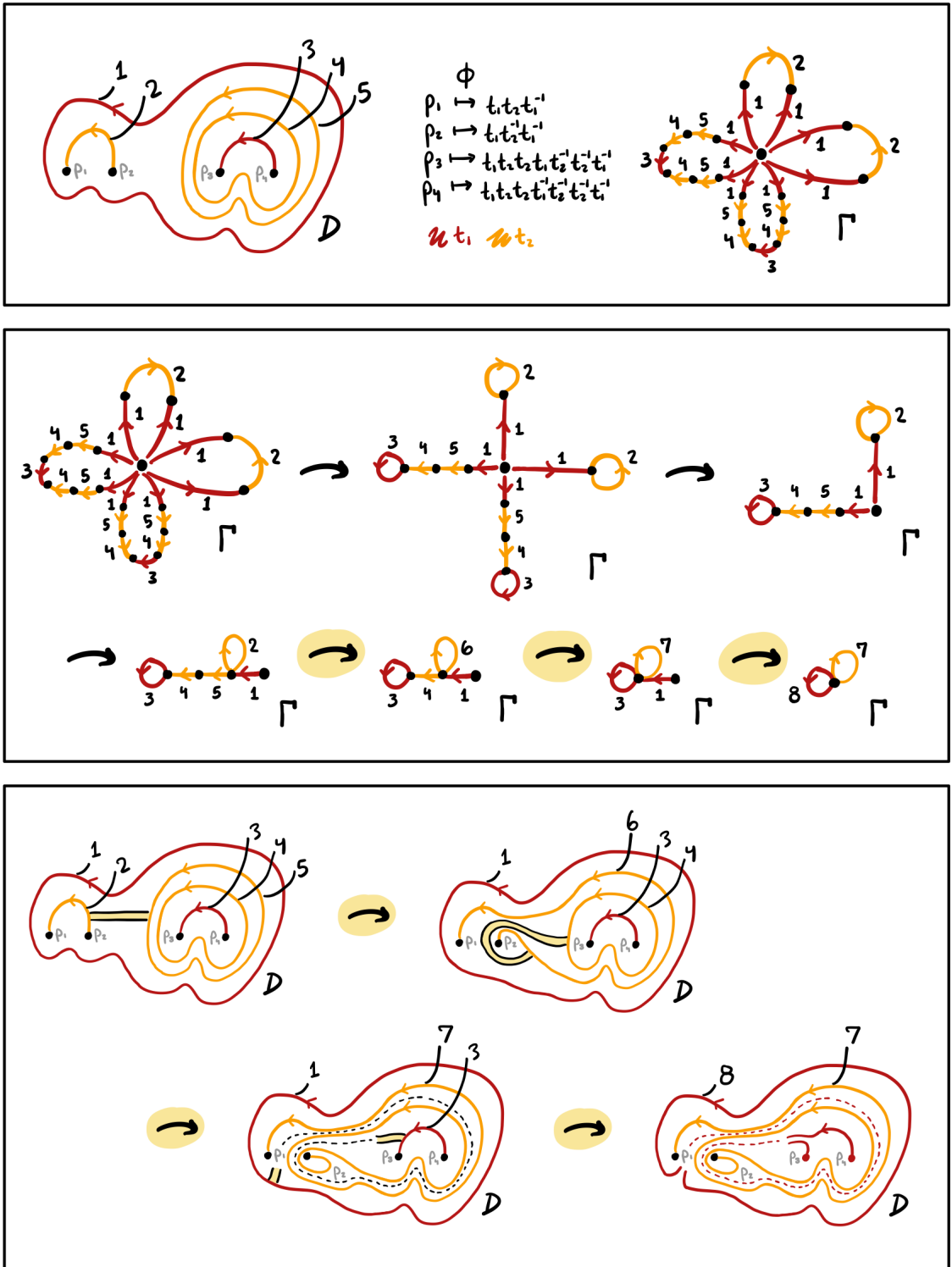


Figure 28: An example containing more complicated band sums.

can be translated into the algebra of bounding homomorphisms up to stabilization, starting from the observation that a pair of bounding homomorphisms determines a link in bridge position in a Heegaard split 3-manifold. We then state a correspondence theorem in this setting.

In Section 4A we recall the 4-dimensional story of group trisections of closed 4-manifolds. Then in Section 4B we consider the case of surfaces inside 4-manifolds, where a triple of bounding homomorphisms with a pairwise freeness condition determines a bridge trisected surface in a trisected 4-manifold.

### 3A Closed 3-manifolds

This section is heavily inspired by Jaco’s announcement [20] of a result similar to Theorem 3.1. Our topological input theorem here is the Reidemeister–Singer theorem.

**Topological input theorem** [36; 39] *Any two Heegaard splittings of a fixed 3-manifold become isotopic after some number of stabilizations.*

Let  $\text{Man}^3$  denote the set of all closed, connected, oriented 3-manifolds considered up to orientation-preserving diffeomorphism (or equivalently homeomorphism). Let  $\text{Alg}^3$  denote the set of pairs of homomorphisms  $(\phi_1, \phi_2)$  where  $\phi_i: \pi_1(\Sigma_g, *) \twoheadrightarrow F_g$  for  $i = 1, 2$  are surjections. We will call such pairs  $(\phi_1, \phi_2)$  *splitting homomorphisms* [40; 19]. Given a single such surjection  $\phi$ , which is a bounding homomorphism for the special case where  $b = 0$ , using Theorem 2.10 we obtain a handlebody, which we will denote by  $H(\phi)$ , such that the following diagram commutes for an isomorphism  $\psi$  as in Lemma 2.4:

$$\begin{array}{ccc}
 & & \pi_1(H(\phi), *) \\
 & \nearrow \iota & \downarrow \cong \psi \\
 \pi_1(\Sigma_g, *) & & F_g \\
 & \searrow \phi & \\
 & & 
 \end{array}$$

By Lemma 2.3,  $H(\phi)$  is the unique handlebody bounding  $\Sigma_g$  with the property that there exists a vertical isomorphism  $\psi$  in this diagram making it commute.

Therefore, given  $(\phi_1, \phi_2) \in \text{Alg}^3$  we can form two handlebodies  $H(\phi_1), H(\phi_2)$  with boundary  $\Sigma_g$ , and thus we obtain a compact 3-manifold  $M(\phi_1, \phi_2) = H(\phi_1) \cup_{\Sigma_g} -H(\phi_2)$ , which is given the orientation that naturally results from gluing the orientations of  $H(\phi_1)$  and  $-H(\phi_2)$ . We thus have a map

$$M: \text{Alg}^3 \rightarrow \text{Man}^3, \quad (\phi_1, \phi_2) \mapsto M(\phi_1, \phi_2).$$

We will define three relations on  $\text{Alg}^3$ , denoted by  $\sim_h, \sim_m$  and  $\sim_s$ , such that this map  $M$  descends to the quotient of  $\text{Alg}^3$  by these relations. Both of the relations  $\sim_h$  and  $\sim_m$  are actually equivalence relations on the set  $\text{Alg}^3$ , while  $\sim_s$  is not. We will abuse notation and also denote by  $\sim_s$  the equivalence relation on  $\text{Alg}^3$  generated by the relation  $\sim_s$  (that is, the smallest equivalence relation containing  $\sim_s$ ). The  $h$  here stands for “handleslide”, the  $m$  for “mapping class”, and the  $s$  for “stabilization”. The proof of Theorem 3.2 motivates this choice of notation.

We say  $(\phi_1, \phi_2) \sim_h (\phi'_1, \phi'_2)$  if, for  $i = 1, 2$ , there exist isomorphisms  $h_i : F_g \rightarrow F_g$  such that the following diagram commutes:

$$\begin{array}{ccc} & \phi_i \twoheadrightarrow & F_g \\ \pi_1(\Sigma_g, *) & & \cong \downarrow h_i \\ & \phi'_i \twoheadrightarrow & F_g \end{array}$$

We call an automorphism  $m : \pi_1(\Sigma_g, *) \rightarrow \pi_1(\Sigma_g, *)$  *orientation-preserving* if the induced automorphism  $H_2(\pi_1(\Sigma_g, *); \mathbb{Z}) \rightarrow H_2(\pi_1(\Sigma_g, *); \mathbb{Z})$  is the identity. We write  $(\phi_1, \phi_2) \sim_m (\phi'_1, \phi'_2)$  if there exists an orientation-preserving isomorphism  $m : \pi_1(\Sigma_g, *) \rightarrow \pi_1(\Sigma_g, *)$  such that, for  $i = 1, 2$ , the following diagram commutes:

$$\begin{array}{ccc} \pi_1(\Sigma_g, *) & & \phi_i \twoheadrightarrow \\ \cong \downarrow m & & F_g \\ \pi_1(\Sigma_g, *) & & \phi'_i \twoheadrightarrow \end{array}$$

Next we define  $\sim_s$ . We note that  $\phi'_i$  will be a map from  $\pi_1(\Sigma_g, *) \twoheadrightarrow F_g$  while  $\phi_i$  will be a map from  $\pi_1(\Sigma_{g+1}, *) \twoheadrightarrow F_{g+1}$ . Let  $a_i, b_i$  be the generators of  $\pi_1(\Sigma_g, *)$  (and, abusing notation,  $\pi_1(\Sigma_{g+1}, *)$ ), and  $h_i$  be the generators of  $F_g$  (and, abusing notation,  $F_{g+1}$ ). We say  $(\phi_1, \phi_2) \sim_s (\phi'_1, \phi'_2)$  if  $\phi_i(a_j) = \phi'_i(a_j)$  and  $\phi_i(b_j) = \phi'_i(b_j)$  for  $i = 1, 2$  and  $j = 1, \dots, g$  (where we are identifying  $F_g$  naturally as a subset of  $F_{g+1}$ ), and the rest of the generators are mapped as follows:

$$\phi_1(a_{g+1}) = h_{g+1}, \quad \phi_1(b_{g+1}) = 1, \quad \phi_2(a_{g+1}) = 1, \quad \phi_2(b_{g+1}) = h_{g+1}.$$

Let  $\sim$  denote the equivalence relation on  $\text{Alg}^3$  generated by  $\sim_h, \sim_m$  and  $\sim_s$ . Now we proceed to the main result of this section. A result that is similar in spirit was announced in [20].

**Theorem 3.1** *The map  $M : \text{Alg}^3 \rightarrow \text{Man}^3$  descends to  $\text{Alg}^3/\sim$  and the resulting map is a bijection.*

**Proof** We consider an intermediate set  $\text{Diag}^3$  whose elements are *Heegaard diagrams*, that is, tuples  $(\Sigma_g, \alpha, \beta)$  where  $\alpha$  and  $\beta$  are cut systems on  $\Sigma_g$  (which are only considered up to isotopy). We will refer to these simply as *diagrams*. Then the map  $M$  factors as

$$\begin{array}{ccc} \text{Alg}^3 & \xrightarrow{M} & \text{Man}^3 \\ & \searrow D & \nearrow R \\ & \text{Diag}^3 & \end{array}$$

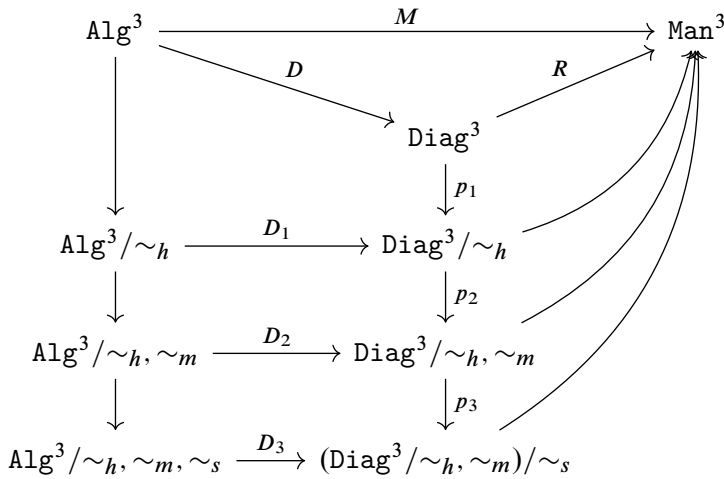
The map  $R : \text{Diag}^3 \rightarrow \text{Man}^3$  is the topological realization of a diagram  $(\Sigma_g, \alpha, \beta)$ , where we cross  $\Sigma_g$  with an interval, glue disks on the respective sides to  $\alpha$  and  $\beta$ , and then glue 3-balls to the resulting sphere boundary components. The map  $D : \text{Alg}^3 \rightarrow \text{Diag}^3$  is the construction of  $M$  using [Theorem 2.10](#), but where we stop at just a diagram (rather than realizing the manifold) with  $\alpha$  corresponding to  $\phi_1$

|   |   |
|---|---|
| $\sim_h$  | an equivalence relation on $\text{Alg}^3$ (as defined above)                    |
| $\sim_m$  | an equivalence relation on $\text{Alg}^3$ (as defined above)                    |
| $\sim_s$  | an equivalence relation on $\text{Alg}^3$ (as defined above)                    |
| $(\phi_1, \phi_2)$                              | an element of $\text{Alg}^3$  |
| $[\phi_1, \phi_2]$                              | an equivalence class in $\text{Alg}^3/\sim_h$                                   |
| $\llbracket \phi_1, \phi_2 \rrbracket$          | an equivalence class in $\text{Alg}^3/\sim_h, \sim_m$                           |
| <hr/>   |   |
| $\sim_h$  | an equivalence relation on $\text{Diag}^3$ (generated by handleslides)          |
| $\sim_m$  | an equivalence relation on $\text{Diag}^3$ (generated by mapping classes)       |
| $\sim_s$  | an equivalence relation on $\text{Diag}^3/\sim_h$ (generated by stabilizations) |
| $(\Sigma_g, \alpha, \beta)$                     | an element of $\text{Diag}^3$   |
| $[\Sigma_g, \alpha, \beta]$                     | an equivalence class in $\text{Diag}^3/\sim_h$                                  |
| $\llbracket \Sigma_g, \alpha, \beta \rrbracket$ | an equivalence class in $\text{Diag}^3/\sim_h, \sim_m$                          |

Table 1: A summary of the notation used throughout the proof of [Theorem 3.1](#).

and  $\beta$  corresponding to  $\phi_2$ . We use the notation  $D(\phi_1)$  and  $D(\phi_2)$  to denote these cut systems, so that  $D: (\phi_1, \phi_2) \mapsto (\Sigma_g, D(\phi_1), D(\phi_2))$ .

The following commutative diagram is a guide to the logic of the proof:



The goal is to define a bijection  $(\text{Alg}^3/\sim_h, \sim_m, \sim_s) \rightarrow \text{Man}^3$ , so we must show that this map, which passes through an intermediate set of diagrams, is well defined, injective and surjective. We do this by descending by quotients on the algebraic and diagrammatic sides, and checking each time that the relevant map factors through and a bijection between the quotients is achieved.

In [Table 1](#) we summarize the notation used throughout the proof, with precise definitions for the relations on the diagrammatic side following. We abuse notation and use the symbols  $\sim_h, \sim_m, \sim_s$  to denote equivalence relations on both the algebraic and diagrammatic sides.

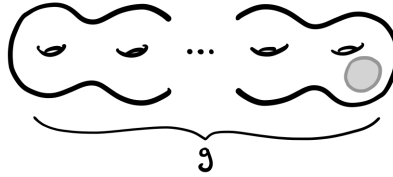


Figure 29: Our standard model for the closed genus  $g$  surface, with a fixed disk for stabilization indicated.

Given two diagrams  $(\Sigma_g, \alpha, \beta)$  and  $(\Sigma_g, \alpha', \beta')$ , we write  $(\Sigma_g, \alpha, \beta) \sim_h (\Sigma_g, \alpha', \beta')$  if there is a sequence of handleslides from the curves  $\alpha$  to  $\alpha'$  and similarly from  $\beta$  to  $\beta'$ . We write  $(\Sigma_g, \alpha, \beta) \sim_m (\Sigma_g, \alpha', \beta')$  if there exists a single mapping class  $\Sigma_g \rightarrow \Sigma_g$  taking  $\alpha$  to  $\alpha'$  and  $\beta$  to  $\beta'$  simultaneously. Let  $[\![\Sigma_g, \alpha, \beta]\!]$  denote an equivalence class of a diagram  $(\Sigma_g, \alpha, \beta)$  under the equivalence relation generated by  $\sim_h$  and  $\sim_m$ .

Recall that stabilizing a Heegaard diagram entails performing a connect sum with the standard genus 1 diagram of  $S^3$ . In order to connect sum in a controlled manner, recall that we have fixed a standard model for the closed genus  $g$  surface, and we additionally fix a disk on this model where the connect sum will be performed. See Figure 29. Because we will mod out by handleslides and mapping classes first, we can assume our Heegaard diagram looks like the standard model. We write  $[\![\Sigma_{g+1}, \alpha, \beta]\!] \sim_s [\![\Sigma_g, \alpha', \beta']\!]$  if we obtain  $\alpha$  and  $\beta$  on  $\Sigma_{g+1}$  from  $\alpha'$  and  $\beta'$  on  $\Sigma_g$  by

- (1) choosing an isotopy of  $\alpha'$  and  $\beta'$  such that they do not intersect the connect sum disk, and
- (2) modifying  $\Sigma_g$  to be  $\Sigma_{g+1}$  (using the fixed disk for the connect sum) and adding the two new curves in the standard genus 1 diagram of  $S^3$  to  $\alpha'$  and  $\beta'$ .

(Note that this description incorporates both stabilization and destabilization, depending on which diagram is seen as the original and which as the modified one.) This operation is not well defined in  $\text{Diag}^3$  because of the choice of isotopy; for instance, see Figure 30. However once we quotient by  $\sim_h$  this is well defined, as we are able to use handleslides to “move” the curve over the attached handle. See Figure 31. Thus we write  $(\text{Diag}^3 / \sim_h, \sim_m) / \sim_s$  rather than  $\text{Diag}^3 / \sim_h, \sim_m, \sim_s$  because unique stabilizations only occur after modding out by handleslides, and additionally we wish to assume our Heegaard diagram looks like our standard model equipped with our fixed disk.

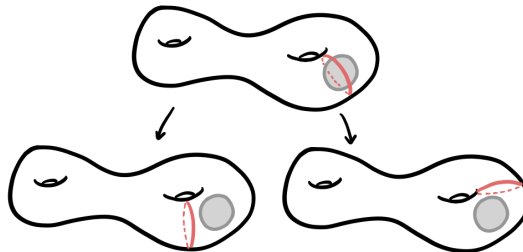


Figure 30: A curve which intersects the disk for stabilization, and two choices of isotopy for moving the curve off of the disk.

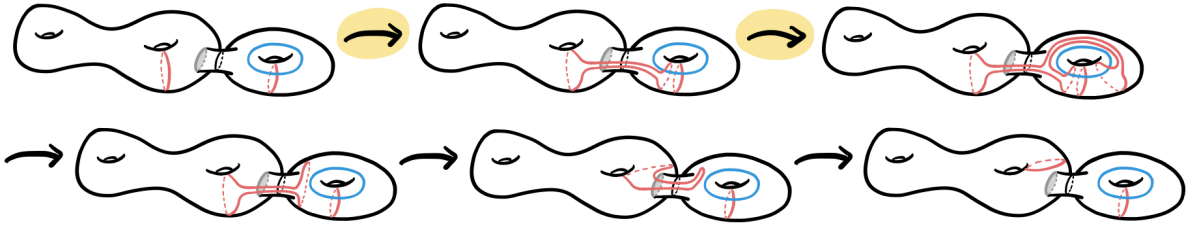


Figure 31: Using handleslides, denoted by the highlighted arrows, to move the curve over the attached handle.

**Claim 1** The map  $p_1 \circ D$  factors through  $\text{Alg}^3/\sim_h$  and the resulting map  $D_1$  is bijective.

Recall that two handlebodies  $H_1$  and  $H_2$  bounding a given surface are equivalent if and only if their diagrams differ by handleslides [22]. Given two splitting homomorphisms  $(\phi_1, \phi_2), (\phi'_1, \phi'_2) \in \text{Alg}^3$  with  $(\phi_1, \phi_2) \sim_h (\phi'_1, \phi'_2)$ , by Lemma 2.3 we see that  $H(\phi_1) = H(\phi'_1)$  and  $H(\phi_2) = H(\phi'_2)$ . Therefore  $(\Sigma_g, D(\phi_1), D(\phi_2)) \sim_h (\Sigma_g, D(\phi'_1), D(\phi'_2))$  and thus the map factors through  $\text{Alg}^3/\sim_h$ , as desired.

To see that the map  $D_1: \text{Alg}^3/\sim_h \rightarrow \text{Diag}^3/\sim_h$  is injective, suppose that  $D_1([\phi_1, \phi_2]) = D_1([\phi'_1, \phi'_2])$  with  $[\phi_1, \phi_2], [\phi'_1, \phi'_2] \in \text{Alg}^3/\sim_h$ . Then  $H(\phi_1) = H(\phi'_1)$  and  $H(\phi_2) = H(\phi'_2)$  and therefore, by the definition of equivalence of handlebodies, we have for  $i = 1, 2$  the commutative diagram

$$\begin{array}{ccc} & & \pi_1(H(\phi_i), *) \\ & \nearrow & \cong \downarrow h_i \\ \pi_1(\Sigma_g, *) & & \pi_1(H(\phi'_i), *) \\ & \searrow & \end{array}$$

for some isomorphisms  $h_i: \pi_1(H(\phi_i), *) \rightarrow \pi_1(H(\phi'_i), *)$ , where the other maps are induced by inclusion. From this, it follows that  $(\phi_1, \phi_2) \sim_h (\phi'_1, \phi'_2)$ . Thus  $D_1$  is injective.

To see that the map  $D_1: \text{Alg}^3/\sim_h \rightarrow \text{Diag}^3/\sim_h$  is surjective, we define a section

$$\sigma: \text{Diag}^3/\sim_h \rightarrow \text{Alg}^3/\sim_h.$$

Let  $[\Sigma_g, \alpha, \beta]$  denote the equivalence class of  $(\Sigma_g, \alpha, \beta)$  in  $\text{Diag}^3/\sim_h$ . We define  $\sigma([\Sigma_g, \alpha, \beta])$  by taking the diagram  $(\Sigma_g, \alpha, \beta)$  with a particular choice of the curves  $\alpha$  and  $\beta$  (so they are no longer isotopy classes but fixed curves). We then reverse the construction of the map  $D$ . That is, we consider each of the sets of curves  $\alpha$  and  $\beta$  separately and apply the construction just as in Lemma 2.4 to obtain maps  $\phi_1, \phi_2: \pi_1(\Sigma_g, *) \rightarrow F_g$ , where here we have chosen orientations for each of the curves in  $\alpha$  and  $\beta$ . These maps  $\phi_1, \phi_2$  are independent of the choice of representatives of the isotopy classes of the curves in  $\alpha$  and  $\beta$ , as well as the choice of orientations, when we consider the result  $[\phi_1, \phi_2]$  in  $\text{Alg}^3/\sim_h$ , giving a map  $\text{Diag}^3 \rightarrow \text{Alg}^3/\sim_h$ . (Note that we have also implicitly chosen an ordering of the curves in each cut system in this construction; however because there are automorphisms of the free group permuting all of the canonical generators, this choice does not matter.) This map factors through to give a map

$\sigma: \text{Diag}^3/\sim_h \rightarrow \text{Alg}^3/\sim_h$  which is a section to  $D_1$  by Lemma 2.3 together with the fact that two cut systems determine the same handlebody if and only if they differ by handleslides. Thus  $D_1$  is surjective.

**Claim 2** *The map  $p_2 \circ D_1$  factors through  $\text{Alg}^3/\sim_h, \sim_m$  and the resulting map  $D_2$  is bijective.*

For well-definedness, suppose that  $(\phi_1, \phi_2) \sim \dots \sim (\phi'_1, \phi'_2)$ , where each  $\sim$  is either  $\sim_h$  or  $\sim_m$ . We must show that  $(\Sigma_g, D(\phi_1), D(\phi_2)) \sim \dots \sim (\Sigma_g, D(\phi'_1), D(\phi'_2))$  where each  $\sim$  is either  $\sim_h$  or  $\sim_m$ . By the previous step of the proof, we know that every  $\sim_h$  equivalence of splitting homomorphisms produces diagrams that are equivalent with respect to  $\sim_h$ . Assume  $(\phi_1, \phi_2) \sim_m (\phi'_1, \phi'_2)$ . By the Dehn–Nielsen–Baer theorem, there exists an orientation-preserving diffeomorphism  $\Sigma_g \rightarrow \Sigma_g$  that fixes the basepoint and realizes the isomorphism  $\pi_1(\Sigma_g, *) \rightarrow \pi_1(\Sigma_g, *)$  that is contained in the assumption that  $(\phi_1, \phi_2) \sim_m (\phi'_1, \phi'_2)$  [7]. This then implies that  $(\Sigma_g, D(\phi_1), D(\phi_2))$  and  $(\Sigma_g, D(\phi'_1), D(\phi'_2))$  are equivalent using  $\sim_h$  and  $\sim_m$ . Therefore, the resulting map  $D_2$  is well defined.

Assume  $D_2([\phi_1, \phi_2]) = D_2([\phi'_1, \phi'_2])$ , where  $[\phi_1, \phi_2], [\phi'_1, \phi'_2] \in \text{Alg}^3/\sim_h, \sim_m$ . We must show that  $(\phi_1, \phi_2) \sim \dots \sim (\phi'_1, \phi'_2)$ , where each  $\sim$  is either  $\sim_h$  or  $\sim_m$ . By assumption,  $(\Sigma_g, D(\phi_1), D(\phi_2)) \sim \dots \sim (\Sigma_g, D(\phi'_1), D(\phi'_2))$ , where each  $\sim$  is either  $\sim_h$  or  $\sim_m$ . Let  $(\Sigma_g, \alpha, \beta)$  and  $(\Sigma_g, \alpha', \beta')$  be two diagrams in the above chain of relations such that  $(\Sigma_g, \alpha, \beta) \sim_m (\Sigma_g, \alpha', \beta')$ . Then there exists an orientation-preserving diffeomorphism

$$F: \Sigma_g \rightarrow \Sigma_g, \quad \alpha \mapsto \alpha', \quad \beta \mapsto \beta',$$

which we can assume fixes the basepoint  $*$ . Let  $(f_1, f_2) = \sigma(\Sigma_g, \alpha, \beta)$  and  $(f'_1, f'_2) = \sigma(\Sigma_g, \alpha', \beta')$ , where  $\sigma$  is the map from the preceding claim. (Note that  $\sigma$  is technically defined on  $\text{Diag}^3/\sim_h$ , but we can similarly apply the same construction to any specific diagram with curves transverse to the generators  $a_1, \dots, b_g$  which are not considered up to isotopy.) Then we have for  $i = 1, 2$  the commutative diagram

$$\begin{array}{ccc} \pi_1(\Sigma_g, *) & & \\ \cong \downarrow & \nearrow f_i & \\ \pi_1(F, *) & \rightarrow & F_g \\ \cong \downarrow & \nwarrow f'_i & \\ \pi_1(\Sigma_g, *) & & \end{array}$$

for  $i = 1, 2$ , where  $\pi_1(F, *)$  is orientation-preserving. Therefore,  $(f_1, f_2) \sim_m (f'_1, f'_2)$ , so the chain of equivalences from  $(\Sigma_g, D(\phi_1), D(\phi_2))$  to  $(\Sigma_g, D(\phi'_1), D(\phi'_2))$  can be converted to a chain of equivalences from  $(\phi_1, \phi_2)$  to  $(\phi'_1, \phi'_2)$ , and the map  $D_2$  is injective.

It follows similarly that

$$\sigma: \text{Diag}^3/\sim_h \rightarrow \text{Alg}^3/\sim_h$$

descends to a map

$$\sigma: \text{Diag}^3/\sim_h, \sim_m \rightarrow \text{Alg}^3/\sim_h, \sim_m,$$

and that it is a section for  $D_2$ .

**Claim 3** The map  $p_3 \circ D_2$  factors through  $\text{Alg}^3 / \sim_h, \sim_m, \sim_s$  and the resulting map  $D_3$  is bijective.

For well-definedness, suppose  $(\phi_1, \phi_2) \sim_s (\phi'_1, \phi'_2)$ . Then, by construction of the map  $D$ , we will have  $[\Sigma_{g+1}, D(\phi_1), D(\phi_2)] \sim_s [\Sigma_g, D(\phi'_1), D(\phi'_2)]$ . Well-definedness therefore follows.

Similarly, by construction of  $D$ , if

$$[\Sigma_{g+1}, D(\phi_1), D(\phi_2)] \sim_s [\Sigma_g, D(\phi'_1), D(\phi'_2)],$$

then  $[\phi_1, \phi_2] \sim_s [\phi'_1, \phi'_2]$ , so  $D_3$  is injective. If  $[\Sigma_{g+1}, \alpha, \beta] \sim_s [\Sigma_g, \alpha', \beta']$ , then, again by construction,  $[\sigma(\Sigma_{g+1}, \alpha, \beta)] \sim_s [\sigma(\Sigma_g, \alpha', \beta')]$ , so  $\sigma$  factors through to give a section of  $D_3$ .

We note that this claim follows more immediately than the previous ones since the definition of the algebraic relation  $\sim_s$  is explicit in the sense that it does not involve any choices, as compared to the definitions of the algebraic relations  $\sim_h$  and  $\sim_m$ . In later sections, we will define other notions of stabilization and they will be similarly explicit.

**Claim 4** The map  $R$  factors through  $(\text{Diag}^3 / \sim_h, \sim_m) / \sim_s$  and the resulting map is bijective.

Every closed, orientable 3-manifold admits a Heegaard decomposition (for example, by taking a triangulation and taking the Heegaard splitting surface to be the boundary of a regular neighborhood of the 1-skeleton). Let  $Y$  be a closed, orientable 3-manifold and let  $S \subset Y$  be a Heegaard splitting surface of genus  $g$ . Choose an identification of  $S$  with  $\Sigma_g$ . We have  $Y = H_1 \cup_S H_2$  for two handlebodies  $H_1$  and  $H_2$ , and by choosing collections of  $g$  disjoint properly embedded disks  $\mathbb{D}_1$  and  $\mathbb{D}_2$  in  $H_1$  and  $H_2$ , respectively, that cut  $H_1$  and  $H_2$  into a 3-ball, then looking at  $(S, \partial\mathbb{D}_1, \partial\mathbb{D}_2)$ , we have a diagram whose topological realization is  $Y$ . (Here the topological realization is as before: thicken  $S$ , glue disks to  $\mathbb{D}_1$  and  $\mathbb{D}_2$  on their respective sides, and glue in 3-balls to the resulting spheres.) Using the identification of  $S$  with  $\Sigma_g$ , we obtain a diagram  $(\Sigma_g, \alpha, \beta) \in \text{Diag}^3$  where  $\alpha$  and  $\beta$  are the respective images of  $\partial\mathbb{D}_1$  and  $\partial\mathbb{D}_2$ , and the image of  $(\Sigma_g, \alpha, \beta)$  in  $\text{Man}^3$  is  $Y$ . Therefore the map  $R: \text{Diag}^3 \rightarrow \text{Man}^3$  is surjective.

The factored-through map  $R: (\text{Diag}^3 / \sim_h, \sim_m) / \sim_s \rightarrow \text{Man}^3$  is injective by the Reidemeister–Singer theorem, and hence a bijection. By composing  $D_3$  with this map, we obtain the theorem.  $\square$

### 3B Bridge split links in 3-manifolds

In Section 3A we used that any pair of Heegaard splittings of the same fixed 3-manifold become isotopic after some number of stabilization operations (which corresponds to connect summing with the genus 1 splitting of the 3-sphere). The goal of this section will be translating the corresponding uniqueness up to perturbation statement for bridge splittings of links in 3-manifolds into the algebra of bounding homomorphisms.

**Topological input theorem** [15; 44] *Let  $L$  be a link in a fixed Heegaard split 3-manifold. Then any two bridge splittings of  $L$  become isotopic after some number of perturbations.*

Let  $\text{Man}^{(3,1)}$  denote the set of closed, connected, oriented 3-manifolds  $M$  together with a link  $L \subset M$  modulo orientation-preserving diffeomorphisms preserving the links. Let  $\text{Alg}^{(3,1)}$  denote the set of pairs  $(\phi_1, \phi_2)$  such that  $\phi_1, \phi_2: \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \twoheadrightarrow F_{g+b}$  are bounding homomorphisms. Throughout this section, let  $a_j, b_j$  and  $p_k$  denote the generators of  $\pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *)$ , where  $a_j, b_j$  are the surface generators (for  $j = 1, \dots, g$ ) and  $p_k$  are the puncture generators (for  $k = 1, \dots, 2b$ ), and let  $h_l, t_l$  denote the generators of  $F_{g+b}$  (for  $l = 1, \dots, b$ ).

We define the map  $(M, L): \text{Alg}^{(3,1)} \rightarrow \text{Man}^{(3,1)}$  as follows. Given  $(\phi_1, \phi_2) \in \text{Alg}^{(3,1)}$ , let  $H(\phi_1)$  and  $H(\phi_2)$  be the handlebodies bounding  $\Sigma_g$  that result from the application of [Theorem 2.10](#), and further let  $T(\phi_1) \subset H(\phi_1)$  and  $T(\phi_2) \subset H(\phi_2)$  be the resulting trivial tangles in these handlebodies. We then define  $(M, L)(\phi_1, \phi_2)$  to be the 3-manifold  $H(\phi_1) \cup_{\Sigma_g} H(\phi_2)$  (with the orientation as in [Section 3A](#)) together with the link  $T(\phi_1) \cup_{\{p_1, \dots, p_{2b}\}} T(\phi_2)$ .

We now define the analogues of the equivalence relations  $\sim_h, \sim_m$  and  $\sim_s$  on  $\text{Alg}^3$  in this setting. Given  $(\phi_1, \phi_2), (\phi'_1, \phi'_2) \in \text{Alg}^{(3,1)}$ , we write  $(\phi_1, \phi_2) \sim_h (\phi'_1, \phi'_2)$  if, for  $i = 1, 2$ , there exist isomorphisms  $h_i: F_{g+b} \rightarrow F_{g+b}$  such that the following diagram commutes:

$$\begin{array}{ccc} & & F_{g+b} \\ & \nearrow \phi_i & \\ \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) & & \cong \downarrow h_i \\ & \searrow \phi'_i & F_{g+b} \end{array}$$

Let  $m: \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \rightarrow \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *)$  be an automorphism that preserves the conjugacy classes of  $p_1, \dots, p_{2b}$  setwise. Then  $m$  descends to an automorphism  $\pi_1(\Sigma_g, *) \rightarrow \pi_1(\Sigma_g, *)$ , by the surjective map  $\pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \twoheadrightarrow \pi_1(\Sigma_g, *)$  which sends each  $p_i$  to the identity and is the identity on all of the elements  $a_1, b_1, \dots, a_g, b_g$ . We call  $m$  *orientation-preserving* if this corresponding automorphism  $\pi_1(\Sigma_g, *) \rightarrow \pi_1(\Sigma_g, *)$  is orientation-preserving. Given  $(\phi_1, \phi_2), (\phi'_1, \phi'_2) \in \text{Alg}^{(3,1)}$ , we write  $(\phi_1, \phi_2) \sim_m (\phi'_1, \phi'_2)$  if there exists an orientation-preserving isomorphism

$$m: \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \rightarrow \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *)$$

such that, for  $i = 1, 2$ , the following diagram commutes:

$$\begin{array}{ccc} \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) & & F_{g+b} \\ \cong \downarrow m & \nearrow \phi_i & \\ \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) & \searrow \phi'_i & \end{array}$$

While  $\sim_h$  and  $\sim_m$  as defined in this section are very similar to  $\sim_h$  and  $\sim_m$  as defined in [Section 3A](#), the analogue of  $\sim_s$  is a bit more complicated. We will have one such relation  $\sim_{s_g}$ , which is directly analogous to  $\sim_s$  in [Section 3A](#); namely, it captures the idea of increasing the genus of the Heegaard splitting while leaving everything else fixed. In addition, there are two relations  $\sim_{s_b^1}$  and  $\sim_{s_b^2}$  which will correspond to the idea of modifying a link in bridge position by perturbation.

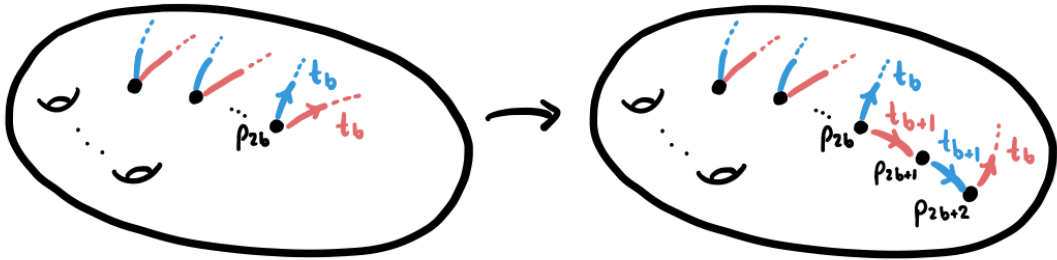


Figure 32: Performing a perturbation, where one of the pink tangle strands is being pulled through the surface. This adds a tangle strand to each side and increases the number of punctures by two. If  $\phi_1$  corresponds to pink and  $\phi_2$  corresponds to blue, then this is a picture of the  $\sim_{s_b^1}$  version of perturbation. Switching the colors would give a picture for the  $\sim_{s_b^2}$  version.

Given  $(\phi_1, \phi_2), (\phi'_1, \phi'_2) \in \text{Alg}^{(3,1)}$ , we now define the relation  $\sim_{s_g}$ . We note that  $\phi'_i$  will be a map from  $\pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \twoheadrightarrow F_{g+b}$  while  $\phi_i$  will be a map from  $\pi_1(\Sigma_{g+1} - \{p_1, \dots, p_{2b}\}, *) \twoheadrightarrow F_{g+1+b}$ . We say  $(\phi_1, \phi_2) \sim_{s_g} (\phi'_1, \phi'_2)$  if  $\phi_i(a_j) = \phi'_i(a_j)$ ,  $\phi_i(b_j) = \phi'_i(b_j)$  and  $\phi_i(p_k) = \phi'_i(p_k)$  for  $i = 1, 2$ ,  $j = 1, \dots, g$  and  $k = 1, \dots, 2b$  (where we are identifying  $F_{g+b}$  naturally as a subset of  $F_{g+1+b}$ , identifying the  $h_i$  generators in  $F_{g+b}$  with  $h_i$  in  $F_{g+1+b}$  and similarly with  $t_i$ ), and the rest of the generators are mapped as follows:

$$\phi_1(a_{g+1}) = h_{g+1}, \quad \phi_1(b_{g+1}) = 1, \quad \phi_2(a_{g+1}) = 1, \quad \phi_2(b_{g+1}) = h_{g+1}.$$

Finally, we define  $\sim_{s_b^1}$  and  $\sim_{s_b^2}$ , which correspond to perturbation of the tangle strands. See Figure 32. Suppose now that  $b > 0$ . There are two such operations since we can either push a tangle strand from the side corresponding to  $\phi_1$  across  $\Sigma_g$ , or we can push a tangle strand corresponding to  $\phi_2$ . The motivation for this comes from investigating Figure 32 and imagining applying the operation  $\sigma$  from the proof of Theorem 3.1 to the before and after parts of the figure. Let  $(\phi_1, \phi_2), (\phi'_1, \phi'_2) \in \text{Alg}^{(3,1)}$ . We note that  $\phi'_i$  will be a map from  $\pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \twoheadrightarrow F_{g+b}$  while  $\phi_i$  will be a map from  $\pi_1(\Sigma_g - \{p_1, \dots, p_{2b}, p_{2b+1}, p_{2b+2}\}, *) \twoheadrightarrow F_{g+b+1}$ . Assume without loss of generality that  $\phi'_1(p_{2b}) = \phi'_2(p_{2b}) = t_b$ . (We can do this because we first mod out by mapping class group elements; see the proof of Theorem 3.1, Claim 1, for details.) We write  $(\phi_1, \phi_2) \sim_{s_b^1} (\phi'_1, \phi'_2)$  if

$$\begin{aligned} \phi_1(p_{2b}) &= t_{b+1}, & \phi_1(p_{2b+1}) &= (t_{b+1})^{-1}, & \phi_1(p_{2b+2}) &= t_b, \\ \phi_2(p_{2b}) &= t_b, & \phi_2(p_{2b+1}) &= t_{b+1}, & \phi_2(p_{2b+2}) &= (t_{b+1})^{-1}, \end{aligned}$$

and  $\phi_i$  and  $\phi'_i$  agree for all other elements in the generating sets (suitably identifying the groups) for  $i = 1, 2$ . We similarly define  $\sim_{s_b^2}$  by swapping the roles of the indices 1 and 2.

Let  $\sim$  denote the equivalence relation on  $\text{Alg}^{(3,1)}$  generated by  $\sim_h, \sim_m, \sim_{s_g}, \sim_{s_b^1}$  and  $\sim_{s_b^2}$ .

**Theorem 3.2** *The map  $(M, L): \text{Alg}^{(3,1)} \rightarrow \text{Man}^{(3,1)}$  descends to  $\text{Alg}^{(3,1)}/\sim$  and the resulting map is a bijection.*

**Proof** As in the proof of [Theorem 3.1](#), we consider an intermediate set  $\text{Diag}^{(3,1)}$  whose elements are *diagrams*, that is, tuples  $(\Sigma_g, \alpha, \beta, S_\alpha, S_\beta)$  where  $\alpha, \beta$  are cut systems on  $\Sigma_g$ , and  $S_\alpha, S_\beta$  are shadow diagrams for trivial tangles with endpoints  $\{p_1, \dots, p_{2b}\}$  (which are all only considered up to isotopy). In other words,  $(\alpha, S_\alpha)$  is a curve-and-arc system for one tangle and handlebody, and  $(\beta, S_\beta)$  is a curve-and-arc system for the other. (Refer to the beginning of [Section 2](#) for the definition of *curve-and-arc system*.) Then the map  $(M, L)$  factors as

$$\begin{array}{ccc} \text{Alg}^{(3,1)} & \xrightarrow{(M,L)} & \text{Man}^{(3,1)} \\ & \searrow D & \nearrow R \\ & & \text{Diag}^{(3,1)} \end{array}$$

The map  $R: \text{Diag}^{(3,1)} \rightarrow \text{Man}^{(3,1)}$  is the topological realization of a diagram  $(\Sigma_g, \alpha, \beta, S_\alpha, S_\beta)$ , where we cross  $\Sigma_g$  with an interval, glue disks on the respective sides to  $\alpha$  and  $\beta$ , glue three balls to the resulting sphere boundary components to obtain two handlebodies, and then push the interiors of the shadow arcs in  $S_\alpha$  and  $S_\beta$  into their respective handlebody to obtain two tangles. The map  $D: \text{Alg}^{(3,1)} \rightarrow \text{Diag}^{(3,1)}$  is the construction of  $(M, L)$  using [Theorem 2.10](#), but where we stop at just a diagram with curves  $\alpha$  and arcs  $S_\alpha$  corresponding to  $\phi_1$  and curves  $\beta$  and arcs  $S_\beta$  corresponding to  $\phi_2$ .

In the proof of [Theorem 3.1](#) (see [Claim 1](#)) we used the fact that two handlebodies bounding a given surface are equivalent if and only if their diagrams differ by handleslides. In this proof the following fact will take its place: two tangles in a handlebody are isotopic (fixing their boundary points) if and only if their curve-and-arc systems are related by a sequence of isotopies and slides. This folklore fact appears, for instance, in [[32](#), Proposition 3.1; [30](#), Proposition 5.2], both citing [[21](#)] for the proof idea. Our topological input theorem can now be translated into the following diagrammatic statement: two diagrams of the same link in a 3-manifold are related by a sequence of perturbations, deperturbations, and moves from the above fact. In other words, these are diagrammatic equivalence relations which have the algebraic counterparts  $\sim_h, \sim_m, \sim_{s_g}, \sim_{s_b^1}$  and  $\sim_{s_b^2}$  as described above.

The rest of the proof then follows in similar fashion as before; mod out the map  $D$  by these diagrammatic equivalence relations, and then show each time that the map factors through and a bijection between quotients is achieved. We leave the details to the reader. Then we use both our topological input theorem from this section and the Reidemeister–Singer theorem (as previously) to achieve the result. □

**Remark 3.3** [Theorem 3.2](#) is in fact a generalization of [Theorem 3.1](#) where we take the links to be empty. Note also that the number of components of the link resulting from a given pair  $(\phi_1, \phi_2)$  can easily be read off from the maps  $\phi_1$  and  $\phi_2$ . We can thereby describe the partition of the set  $\text{Alg}^{(3,1)}$  so that the different equivalence classes correspond in the bijection above to manifolds with links of  $0, 1, 2, \dots$  components (in particular,  $\sim$  respects this partition). In particular, [Theorem 3.1](#) is recovered by restricting to the case where  $b = 0$ .

## 4 Closed 4-manifolds and bridge trisected surfaces

Here we continue the translation of topology into algebra that we started in the previous section, but moving up a dimension. The proofs in this section follow similarly to those in the previous, and thus we omit some of the details.

### 4A Closed 4-manifolds

In 2016 Gay and Kirby introduced *trisectings* of 4-manifolds, which can be seen as 4-dimensional analogues of Heegaard splittings.

**Definition 4.1** (trisection of a 4-manifold [8]) A  $(g; k_1, k_2, k_3)$ -trisection of a smooth, closed, connected, oriented 4-manifold  $X^4$  is a decomposition  $X^4 = X_1 \cup X_2 \cup X_3$  with the following properties:

- (1) Each  $X_i$  is a 4-dimensional 1-handlebody, that is, diffeomorphic to  $\natural^{k_i}(S^1 \times B^3)$ . If  $k_i = 0$ , we interpret this boundary connected sum as  $B^4$ .
- (2) The  $X_i$ 's intersect pairwise in genus  $g$  handlebodies; that is, their pairwise intersections are diffeomorphic to  $H_g := \natural^g(S^1 \times B^2)$ .
- (3) The triple intersection of the  $X_i$ 's is a genus  $g$  oriented surface (denoted by  $\Sigma_g$ ), called the *central surface*.

If  $k_1 = k_2 = k_3 =: k$ , we call the trisection *balanced*, and denote this by  $(g; k)$ .

A trisection is determined by its *spine*, namely the central surface along with the three handlebodies the surface bounds, as there is a unique way, up to diffeomorphism, to fill this in with 4-dimensional 1-handlebodies [28]. Every smooth, closed, connected, oriented 4-manifold admits a trisection, and a trisection of a given 4-manifold is unique up to stabilization (see [8] for a proof of this result and the precise definition of stabilization).

**Topological input theorem** [8] *Any pair of trisectings of a fixed 4-manifold become isotopic after some number of stabilizations.*

Technically, there are two notions of stabilization: a balanced one (which increases the genus of the central surface by three) and an unbalanced one (which increases the genus by one). We will assume all trisectings are balanced and work only with balanced stabilizations. We can make this assumption without loss of generality as any unbalanced trisection can be made balanced with some number of unbalanced stabilizations. Note that we could easily expand the algebraic relations included in this section to include those which would correspond to unbalanced stabilizations, but for the sake of simplicity of notation we have not done this.

Let  $\text{Man}^4$  be the set of closed, connected, oriented, smooth 4-manifolds up to orientation-preserving diffeomorphism. Let  $\text{Alg}^4$  denote the set of triples  $(\phi_1, \phi_2, \phi_3)$  of surjective homomorphisms  $\pi_1(\Sigma_g, *) \twoheadrightarrow F_g$  for some integer  $g$  such that  $M(\phi_1, \phi_2)$ ,  $M(\phi_2, \phi_3)$  and  $M(\phi_3, \phi_1)$  are all diffeomorphic to  $\#^k(S^1 \times S^2)$

for some integer  $k$ , where  $M$  is as in Section 3A. By the Poincaré conjecture [35], this is equivalent to the property that the pushout of  $\phi_i$  and  $\phi_j$  for  $i \neq j$  are (necessarily finitely generated) free groups. Note that here, an individual  $\phi$  is a bounding homomorphism for the special case  $b = 0$  which satisfies the additional pushout property, and a triple  $(\phi_1, \phi_2, \phi_3) \in \text{Alg}^4$  is a *group trisection* of the fundamental group of a 4-manifold, as described in [1] and Section 1.

We have a map  $X: \text{Alg}^4 \rightarrow \text{Man}^4$  just as in [1]. Namely, given  $(\phi_1, \phi_2, \phi_3) \in \text{Alg}^4$ , identify the handlebodies  $H(\phi_1), H(\phi_2), H(\phi_3)$  along their common boundary  $\Sigma_g$ . Now the three 3-manifolds  $M(\phi_i, \phi_j) = H(\phi_i) \cup_{\Sigma_g} H(\phi_j)$  for  $i \neq j$  are all diffeomorphic to some  $\#^k(S^1 \times S^2)$  by the assumption that the pairwise pushouts of the  $\phi_i$  and  $\phi_j$  are free groups. Therefore we may glue in three 4-dimensional 1-handlebodies (uniquely by [28]) to obtain a smooth, closed 4-manifold  $X(\phi_1, \phi_2, \phi_3)$ . We orient  $M(\phi_1, \phi_2) \subset X(\phi_1, \phi_2, \phi_3)$  as in Section 3A and we orient  $X$  so that the orientation restricted to the 4-dimensional 1-handlebody that is glued to  $M(\phi_1, \phi_2)$  induces this orientation on  $M(\phi_1, \phi_2)$ .

We define the relations  $\sim_h$  and  $\sim_m$  in an analogous fashion as was done for  $\text{Alg}^3$  in Section 3A. The stabilization relation  $\sim_s$  on  $\text{Alg}^4$  is defined as follows. Here  $\phi'_i$  will be a map from  $\pi_1(\Sigma_g, *) \twoheadrightarrow F_g$  while  $\phi_i$  will be a map from  $\pi_1(\Sigma_{g+3}, *) \twoheadrightarrow F_{g+3}$ . Let  $a_i, b_i$  be the generators of  $\pi_1(\Sigma_g, *)$  (and, abusing notation,  $\pi_1(\Sigma_{g+3}, *)$ ), and  $h_i$  be the generators of  $F_g$  (and, abusing notation,  $F_{g+3}$ ). We say  $(\phi_1, \phi_2, \phi_3) \sim_s (\phi'_1, \phi'_2, \phi'_3)$  if  $\phi_i(a_j) = \phi'_i(a_j)$  and  $\phi_i(b_j) = \phi'_i(b_j)$  for  $i = 1, 2, 3$  and  $j = 1, \dots, g$  (where we are identifying  $F_g$  naturally as a subset of  $F_{g+3}$ ), and the rest of the generators are mapped as follows:

$$\begin{aligned} \phi_1(a_{g+1}) &= h_{g+1}, & \phi_2(a_{g+1}) &= h_{g+1}, & \phi_3(a_{g+1}) &= 1, \\ \phi_1(b_{g+1}) &= 1, & \phi_2(b_{g+1}) &= 1, & \phi_3(b_{g+1}) &= h_{g+1}, \\ \phi_1(a_{g+2}) &= h_{g+2}, & \phi_2(a_{g+2}) &= 1, & \phi_3(a_{g+2}) &= h_{g+2}, \\ \phi_1(b_{g+2}) &= 1, & \phi_2(b_{g+2}) &= h_{g+2}, & \phi_3(b_{g+2}) &= 1, \\ \phi_1(a_{g+3}) &= 1, & \phi_2(a_{g+3}) &= h_{g+3}, & \phi_3(a_{g+3}) &= h_{g+3}, \\ \phi_1(b_{g+3}) &= h_{g+3}, & \phi_2(b_{g+3}) &= 1, & \phi_3(b_{g+3}) &= 1. \end{aligned}$$

This is an algebraic analogue of the topological operation of stabilizing a trisection, just as in Section 3A where we presented the analogous notion for stabilizations of Heegaard splittings. As before, we let the equivalence relation  $\sim_s$  be the symmetrization of the relation  $\sim_s$ .

As we now have three maps, we must define one more relation, which does not have a counterpart in Section 3, corresponding to cyclically permuting the “colors” of the curves on a trisection diagram (that is, cyclically permuting the roles of the cut system curves  $\alpha, \beta$  and  $\gamma$ ). We say  $(\phi_1, \phi_2, \phi_3) \sim_c (\phi'_1, \phi'_2, \phi'_3)$  if  $\phi_1 = \phi'_2, \phi_2 = \phi'_3$  and  $\phi_3 = \phi'_1$ .

Let  $\sim$  denote the equivalence relation on  $\text{Alg}^4$  generated by  $\sim_h, \sim_m, \sim_s$  and  $\sim_c$ . A result very similar to the following theorem is stated in [1, Theorem 5].

**Theorem 4.2** (compare with [1]) *The map  $X: \text{Alg}^4 \rightarrow \text{Man}^4$  descends to  $\text{Alg}^4/\sim$  and the resulting map is a bijection.*

**Proof** The proof proceeds analogously to the discussion of the Heegaard splittings of closed 3-manifolds, where the role of the Reidemeister–Singer theorem is taken on by the uniqueness of trisections up to stabilization from [8]. □

**Remark 4.3** We can algorithmically determine if the pushout of such a pair  $\phi_i$  and  $\phi_j$  is in fact a free group. Namely, we can construct the corresponding 3-manifold and algorithmically check if it is a (possibly empty) connected sum of copies of  $S^1 \times S^2$  (see for example [27]). We are unaware if there is a more direct algebraic method to verify this condition.

### 4B Bridge trisected surfaces in 4-manifolds

Finally we turn to the setting of knotted surfaces in 4-manifolds. For us, a *knotted surface* is a closed surface smoothly embedded in a smooth, closed, connected, oriented 4-manifold. In particular, our surfaces are not necessarily orientable or connected.

In 2018 Meier and Zupan showed that knotted surfaces in trisected 4-manifolds can always be isotoped into a compatibly trisected surface. This inherited decomposition is called a *bridge trisection*.

**Definition 4.4** (bridge trisection of a knotted surface [31; 32]) A  $(g; k_1, k_2, k_3; b; c_1, c_2, c_3)$ -*bridge trisection* of a knotted surface  $S$  in a 4-manifold  $X^4$  is a decomposition

$$(X^4, S) = (X_1, \mathcal{D}_1) \cup (X_2, \mathcal{D}_2) \cup (X_3, \mathcal{D}_3)$$

with the following properties:

- (1) The decomposition  $X^4 = X_1 \cup X_2 \cup X_3$  is a  $(g; k_1, k_2, k_3)$ -trisection of  $X^4$ .
- (2) Each  $\mathcal{D}_i$  is a boundary parallel collection of  $c_i$  disks in  $X_i$ .
- (3) The  $\mathcal{D}_i$ 's intersect pairwise in trivial  $b$ -strand tangles (denoted by  $T_b$ ) in the handlebodies which are the pairwise intersections of the  $X_i$ 's.

Just as before, if  $k_1 = k_2 = k_3 =: k$ , we replace these parameters with just one  $k$  and call the trisection *balanced*. We assume again that all trisections are balanced (with respect to the  $k_i$  parameter; the  $c_i$  may be different).

Pairwise, the tangles  $T_b$  form unlinks. A bridge trisection is determined by these three tangles, as there is a unique way to smoothly cap off unlinks in  $\#^k(S^1 \times S^2)$  with disks [32]. A knotted surface in a trisected 4-manifold is said to be in *bridge position* if its intersection with the trisection of the 4-manifold results in a bridge trisection. Knotted surfaces can always be isotoped to be in bridge position, and this is unique up to perturbation (for the proof of this result and the precise definition of perturbation, see [31; 32; 18]).

**Topological input theorem** [31; 32; 18] *Any pair of bridge trisections for a smoothly embedded surface in a fixed underlying trisection of a 4-manifold become isotopic after some number of perturbations.*

Throughout this section, let  $a_j, b_j$  and  $p_k$  denote the generators of  $\pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *)$ , where  $a_j, b_j$  are the surface generators (for  $j = 1, \dots, g$ ) and  $p_k$  are the puncture generators (for  $k = 1, \dots, 2b$ ), and let  $h_j, t_l$  denote the generators of  $F_{g+b}$  (for  $l = 1, \dots, b$ ). Given a bounding homomorphism

$$\phi: \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \twoheadrightarrow \langle t_1, \dots, t_b, h_1, \dots, h_g \rangle,$$

we have an associated homomorphism, which we call the *associated closed bounding homomorphism* for  $\phi$ , denoted by

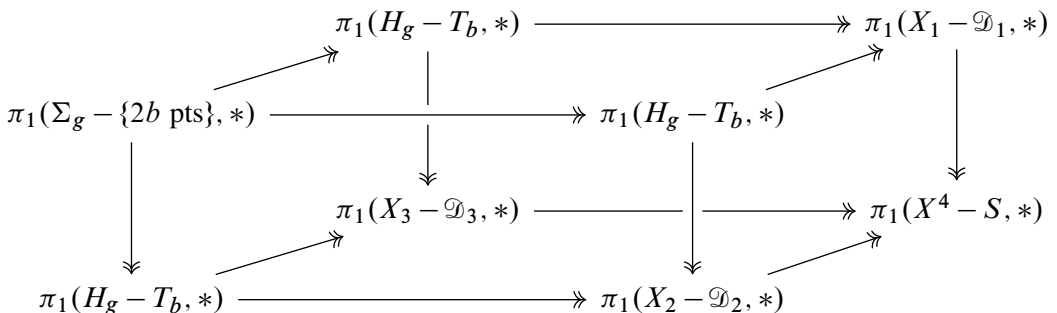
$$\bar{\phi}: \pi_1(\Sigma_g, *) \twoheadrightarrow \langle h_1, \dots, h_g \rangle,$$

which is given by postcomposing  $\phi$  by the map quotienting out all of the  $t_l$  (and again calling the images of the  $h_j$  by  $h_j$ ) and then sending all  $a_j$  and  $b_j$  (now thought of as in  $\pi_1(\Sigma_g, *)$ ) to the resulting elements of the free group generated by the  $h_j$ . Note that since  $\phi$  is a bounding homomorphism,  $\bar{\phi}$  is also.

Let  $\text{Man}^{(4,2)}$  denote the set of closed, connected, oriented, smooth 4-manifolds  $X$  together with a union of closed (potentially nonorientable or disconnected) surfaces  $S \subset X$  modulo orientation-preserving diffeomorphisms preserving the surfaces setwise. Let  $\text{Alg}^{(4,2)}$  denote the set of triples  $(\phi_1, \phi_2, \phi_3)$  of bounding homomorphisms  $\phi_1, \phi_2, \phi_3: \pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \twoheadrightarrow F_{g+b}$  such that

- (1) the pushouts of pairs of the associated closed bounding homomorphisms  $\bar{\phi}_i, \bar{\phi}_j$  for  $i \neq j$  are all free groups, and
- (2) the pushouts of pairs of the bounding homomorphisms  $\phi_i, \phi_j$  are all free groups of rank equal to the sum of the rank of the pushout of  $\bar{\phi}_i$  and  $\bar{\phi}_j$  plus the number of components of  $(\phi_1, \phi_2)$  as in Remark 3.3.

In other words, an element in  $\text{Alg}^{(4,2)}$  is a *group trisection* of a knotted surface group, which could also be described with the following commutative diagram, where every face is a pushout and every homomorphism is surjective (analogous to that in Section 1):



We will need the following lemma in the construction of the map  $(X, S): \text{Alg}^{(4,2)} \rightarrow \text{Man}^{(4,2)}$ .

**Lemma 4.5** (the free group characterizes unlinks) *The fundamental group detects the unlink in the connected sum  $\#^k(S^1 \times S^2)$ . That is, if  $L = L_1 \sqcup \cdots \sqcup L_n \hookrightarrow \#^k(S^1 \times S^2)$  is an  $n$ -component link with  $\pi_1(S^3 - L, *) \cong F_{n+k}$  a free group on  $n + k$  generators, then  $L$  is the  $n$ -component unlink.*

**Proof** This proof is a generalization of the argument in [17, Theorem 1] for unlinks in  $S^3$ . Any abelian subgroup of a free group is cyclic, so the meridian-longitude generators of the torus around a link component span an abelian group in the free  $\pi_1(S^1 \times S^2 - L, *)$ . Since the meridians generate the first homology of a link complement, we know that the longitudes have to be nullhomotopic in the link exterior. Now apply the loop theorem to obtain disjointly embedded disks recognizing the split unlink.  $\square$

**Remark 4.6** There exist nontrivial links  $L \subset \#^k(S^1 \times S^2)$  whose complement has free fundamental group (but not free of the same rank as the group of the unlink). For example, the core curve of one of the solid tori in the standard genus 1 Heegaard splitting of  $S^1 \times S^2$  has complement homotopy equivalent to the other solid torus; that is, the fundamental group of its complement is free on one generator. On the other hand, the fundamental group of the complement of the unknot in  $S^1 \times S^2$  is free on two generators (the meridian of the unknot and the generator of  $\pi_1(S^1 \times S^2, *)$ ).

**Remark 4.7** An analogous statement to Lemma 4.5 is false for higher dimensional links. Cochran [6] exhibited links of 2-spheres in 4-spheres whose groups are free (but not free on their meridians).

Now we define the map  $(X, S): \text{Alg}^{(4,2)} \rightarrow \text{Man}^{(4,2)}$  as follows. Given  $(\phi_1, \phi_2, \phi_3) \in \text{Alg}^{(4,2)}$ , let  $H(\phi_1)$ ,  $H(\phi_2)$  and  $H(\phi_3)$  be the handlebodies bounding  $\Sigma_g$  that result from the application of Theorem 2.10, and further let  $T(\phi_i) \subset H(\phi_i)$  for  $i = 1, 2, 3$  be the resulting trivial tangles in these handlebodies. We glue the handlebodies  $H(\phi_1)$ ,  $H(\phi_2)$  and  $H(\phi_3)$  together along the common surface  $\Sigma_g$  and obtain a closed, oriented, smooth 4-manifold  $X = X(\bar{\phi}_1, \bar{\phi}_2, \bar{\phi}_3)$  as in Section 4A. (Here we have used the condition that the pairwise pushouts of the associated closed bounding homomorphisms are free groups to ensure that we obtain  $\#^k(S^1 \times S^2)$  as the result of gluing two of the handlebodies together, and thus we can cap off with 4-dimensional 1-handlebodies.) The surface  $S$  in  $X$  is obtained by considering the unions of the tangles  $T(\phi_i) \cup_{\{p_1, \dots, p_{2b}\}} T(\phi_j) \subset H(\phi_i) \cup_{\Sigma_g} H(\phi_j)$  for  $i \neq j$ . Since the pushout of  $\phi_i$  and  $\phi_j$  is a free group of the appropriate rank, by Lemma 4.5 we know that the unions of these tangles are all unlinks. Therefore, we can take disjoint bridge disks bounding  $T(\phi_i) \cup_{\{p_1, \dots, p_{2b}\}} T(\phi_j)$  in  $H(\phi_i) \cup_{\Sigma_g} H(\phi_j)$  and push these into the 4-dimensional 1-handlebody bounding  $H(\phi_i) \cup_{\Sigma_g} H(\phi_j)$  in  $X$ . Then the union of these three sets of disks (with one set in each 4-dimensional 1-handlebody) is the knotted surface  $S$ .

We now define the analogues of the various relations from Sections 3B and 4A in this setting. The definitions of  $\sim_h$  and  $\sim_m$  are exactly analogous to the definitions in Section 3B. Just as stabilization appeared in Section 3B as several different relations — namely,  $\sim_{s_g}$  for changing the genus of the surface and  $\sim_{s_b^1}, \sim_{s_b^2}$  for perturbing each of the two tangles — in our current setting, stabilization will also

manifest as several different relations. Here we will have  $\sim_{s_g}$  corresponding to changing the genus of the surface, and  $\sim_{s_b^1}$ ,  $\sim_{s_b^2}$  and  $\sim_{s_b^3}$  corresponding to changing the number of strands in the tangles.

Given  $(\phi_1, \phi_2, \phi_3), (\phi'_1, \phi'_2, \phi'_3) \in \text{Alg}^{(4,2)}$ , we now define  $(\phi_1, \phi_2, \phi_3) \sim_{s_g} (\phi'_1, \phi'_2, \phi'_3)$ . We note that  $\phi'_i$  will be a map from  $\pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \twoheadrightarrow F_{g+b}$  while  $\phi_i$  will be a map from

$$\pi_1(\Sigma_{g+1} - \{p_1, \dots, p_{2b}\}, *) \twoheadrightarrow F_{g+3+b}.$$

We say  $(\phi_1, \phi_2, \phi_3) \sim_{s_g} (\phi'_1, \phi'_2, \phi'_3)$  if  $\phi_i(a_j) = \phi'_i(a_j)$ ,  $\phi_i(b_j) = \phi'_i(b_j)$  and  $\phi_i(p_k) = \phi'_i(p_k)$  for  $i = 1, 2, 3$ ,  $j = 1, \dots, g$  and  $k = 1, \dots, 2b$  (where we are identifying  $F_{g+b}$  naturally as a subset of  $F_{g+3+b}$ , identifying the  $h_i$  generators in  $F_{g+b}$  with  $h_i$  in  $F_{g+3+b}$  and similarly with  $t_i$ ), and the rest of the generators are mapped just as in the definition of  $\sim_s$  in Section 4A.

Now we discuss the algebraic version of perturbation in this setting, which is analogous to the definitions of  $\sim_{s_b^1}$ ,  $\sim_{s_b^2}$  in Section 3B. Suppose that  $b > 0$ . This definition is motivated by the pictures of perturbation shown in Figures 33 and 34, where two arcs of different colors are banded together to create three new arcs, one of each color. There are three such operations to consider depending on how we cyclically permute the colors. For each operation, there are two cases under which the operation can occur: either the two arcs to be banded together share an endpoint, or they do not.

Let  $(\phi_1, \phi_2, \phi_3), (\phi'_1, \phi'_2, \phi'_3) \in \text{Alg}^{(4,2)}$ . In both cases  $\phi'_i$  will be a map from  $\pi_1(\Sigma_g - \{p_1, \dots, p_{2b}\}, *) \twoheadrightarrow F_{g+b}$  while  $\phi_i$  will be a map from  $\pi_1(\Sigma_g - \{p_1, \dots, p_{2b}, p_{2b+1}, p_{2b+2}\}, *) \twoheadrightarrow F_{g+b+1}$ . If the arcs to be banded together share an endpoint, then assume without loss of generality that  $\phi'_1(p_{2b}) = \phi'_2(p_{2b}) = t_b$ . (We can do this because we first mod out by mapping class group elements; see the proof of Theorem 3.1, Claim 1, for details.) We write  $(\phi_1, \phi_2, \phi_3) \sim_{s_b^1} (\phi'_1, \phi'_2, \phi'_3)$  if

$$\begin{aligned} \phi_1(p_{2b}) &= t_{b+1}, & \phi_1(p_{2b+1}) &= (t_{b+1})^{-1}, & \phi_1(p_{2b+2}) &= t_b, \\ \phi_2(p_{2b}) &= t_{b+1}, & \phi_2(p_{2b+1}) &= (t_{b+1})^{-1}, & \phi_2(p_{2b+2}) &= t_b, \\ & & \phi_3(p_{2b+1}) &= t_{b+1}, & \phi_3(p_{2b+2}) &= (t_{b+1})^{-1}, \end{aligned}$$

and  $\phi_i$  and  $\phi'_i$  agree for all other elements in the generating sets (suitably identifying the groups) for  $i = 1, 2, 3$ . See Figure 33. We similarly define  $\sim_{s_b^2}$  and  $\sim_{s_b^3}$  in this case by swapping the roles of the indices.

If the arcs to be banded together do not share an endpoint, then assume without loss of generality that  $\phi'_1(p_{2b-1}) = t_b$  and  $\phi'_2(p_{2b}) = t_b$ . We write  $(\phi_1, \phi_2, \phi_3) \sim_{s_b^1} (\phi'_1, \phi'_2, \phi'_3)$  if

$$\begin{aligned} \phi_1(p_{2b-1}) &= t_{b+1}, & \phi_1(p_{2b+1}) &= (t_{b+1})^{-1}, & \phi_1(p_{2b+2}) &= t_b, \\ \phi_2(p_{2b}) &= t_{b+1}, & \phi_2(p_{2b+1}) &= (t_{b+1})^{-1}, & \phi_2(p_{2b+2}) &= t_b, \\ & & \phi_3(p_{2b+1}) &= t_{b+1}, & \phi_3(p_{2b+2}) &= (t_{b+1})^{-1}, \end{aligned}$$

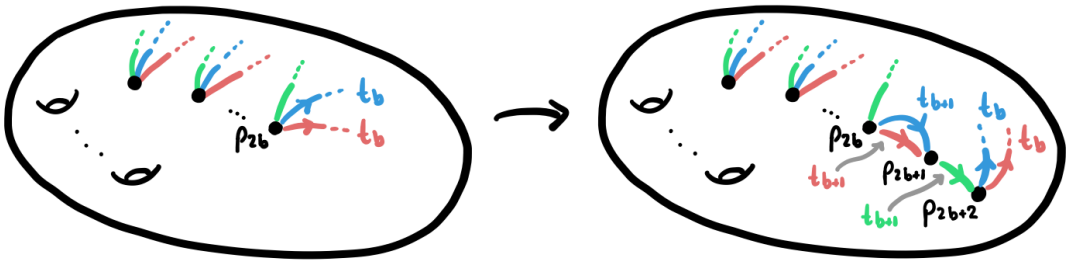


Figure 33: Performing a perturbation, where two arcs of different colors with the same endpoint are banded together. This creates three new tangle strands, one of each color, and increases the number of punctures by two. If  $\phi_1$  corresponds to pink,  $\phi_2$  corresponds to blue, and  $\phi_3$  corresponds to green, then this is a picture of the  $\sim_{s_b^1}$  version of perturbation. Cyclically permuting the colors gives a picture for the  $\sim_{s_b^2}$  and  $\sim_{s_b^3}$  version.

and  $\phi_i$  and  $\phi'_i$  agree for all other elements in the generating sets (suitably identifying the groups) for  $i = 1, 2, 3$ . See Figure 34. We similarly define  $\sim_{s_b^2}$  and  $\sim_{s_b^3}$  in this case by swapping the roles of the indices.

Finally, we must again include a relation corresponding to cyclically permuting the “colors” of the curves and arcs on a trisection diagram (that is, cyclically permuting the roles of the three curve-and-arc systems). We say  $(\phi_1, \phi_2, \phi_3) \sim_c (\phi'_1, \phi'_2, \phi'_3)$  if  $\phi_1 = \phi'_2, \phi_2 = \phi'_3$  and  $\phi_3 = \phi'_1$ . Then let  $\sim$  denote the equivalence relation on  $\text{Alg}^{(4,2)}$  generated by  $\sim_h, \sim_m, \sim_{s_g}, \sim_{s_b^1}, \sim_{s_b^2}, \sim_{s_b^3}$  and  $\sim_c$ .

**Theorem 4.8** *The map  $(X, S): \text{Alg}^{(4,2)} \rightarrow \text{Man}^{(4,2)}$  descends to  $\text{Alg}^{(4,2)}/\sim$  and the resulting map is a bijection.*

**Proof** As in the proofs of Theorems 3.1, 3.2 and 4.2, we consider an intermediate set  $\text{Diag}^{(4,2)}$  whose elements are *trisection diagrams*, that is, tuples  $(\Sigma_g, \alpha, \beta, \gamma, S_\alpha, S_\beta, S_\gamma)$  where  $\alpha, \beta, \gamma$  are cut systems

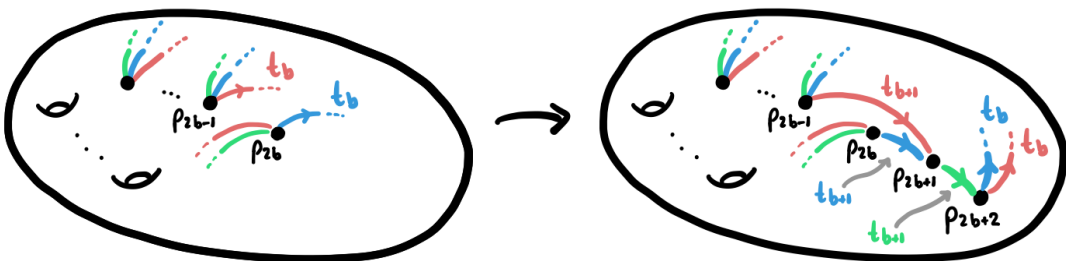


Figure 34: Performing a perturbation, where two arcs of different colors with different endpoints are banded together. This creates three new tangle strands, one of each color, and increases the number of punctures by two. If  $\phi_1$  corresponds to pink,  $\phi_2$  corresponds to blue, and  $\phi_3$  corresponds to green, then this is a picture of the  $\sim_{s_b^1}$  version of perturbation. Cyclically permuting the colors gives a picture for the  $\sim_{s_b^2}$  and  $\sim_{s_b^3}$  version.

on  $\Sigma_g$  and  $S_\alpha, S_\beta, S_\gamma$  are shadow diagrams for trivial tangles with endpoints  $\{p_1, \dots, p_{2b}\}$  (which are all only considered up to isotopy). In other words,  $(\alpha, S_\alpha)$  is a curve-and-arc system for one tangle and handlebody, and similarly for the other two pairs. (Refer to the beginning of Section 2 for the definition of *curve-and-arc system*.) Then the map  $(X, S)$  factors as

$$\begin{array}{ccc} \text{Alg}^{(4,2)} & \xrightarrow{(X,S)} & \text{Man}^{(4,2)} \\ & \searrow D \quad \nearrow R & \\ & \text{Diag}^{(4,2)} & \end{array}$$

The map  $R: \text{Diag}^{(4,2)} \rightarrow \text{Man}^{(4,2)}$  is the topological realization of a diagram  $(\Sigma_g, \alpha, \beta, \gamma, S_\alpha, S_\beta, S_\gamma)$ , where we cross  $\Sigma_g$  with a disk, glue disks on the respective sides to  $\alpha, \beta$  and  $\gamma$ , glue 3-balls to the resulting sphere boundary components to obtain three handlebodies, and then push the interiors of the shadow arcs in  $S_\alpha, S_\beta$  and  $S_\gamma$  into their respective handlebody to obtain three tangles which are pairwise unlinks. Since the pairwise unions of the handlebodies are diffeomorphic to  $\#^k(S^1 \times S^2)$ , we can fill these in uniquely with 4-dimensional 1-handlebodies. Then cap off the unlinks with disks in the pairwise unions of the handlebodies, and then push these disks into the 4-dimensional 1-handlebodies bounded by these unions to create a knotted surface. The map  $D: \text{Alg}^{(4,2)} \rightarrow \text{Diag}^{(4,2)}$  is the construction of  $(X, S)$  using Theorem 2.10, but where we stop at just a diagram with curves  $\alpha$  and arcs  $S_\alpha$  corresponding to  $\phi_1$ , curves  $\beta$  and arcs  $S_\beta$  corresponding to  $\phi_2$ , and curves  $\gamma$  and arcs  $S_\gamma$  corresponding to  $\phi_3$ .

Our topological input theorem can now be translated into the following diagrammatic statement: two diagrams of the same knotted surface in a 4-manifold are related by a sequence of isotopies, slides, perturbations and deperturbations, as in the proof of Theorem 3.2, and, additionally, cyclically permuting the colors. In other words, these are diagrammatic equivalence relations which have the algebraic counterparts  $\sim_h, \sim_m, \sim_{s_g}, \sim_{s_b^1}, \sim_{s_b^2}, \sim_{s_b^3}$  and  $\sim_c$  as described above.

The rest of the proof then follows in similar fashion as before; mod out the map  $D$  by these diagrammatic equivalence relations, and show each time that the map factors through and a bijection between quotients is achieved. Then by our topological input theorem from this section, along with that from Section 4A, we achieve the result. □

One consequence of Theorem 4.8 is the following corollary. By the Gordon–Luecke theorem, (one-dimensional) knots in  $S^3$  are determined by the oriented homeomorphism type of their complements [14], but the same is not true for knotted surfaces in  $S^4$ ; see for instance [12; 42; 26]. However, the extra information contained in a group trisection is enough to distinguish knotted surfaces.

**Corollary 4.9** *Although smoothly knotted surfaces in the 4-sphere cannot always be distinguished by their fundamental groups (or even their complements), they **can** always be distinguished by the group trisections of their fundamental groups.*

## 5 Examples and consequences

In this section we discuss some examples and consequences of our results, including examples of nonequivalent group trisections of the same group, an algebraic version of the smooth unknotting conjecture, a group-theoretic characterization of knot groups, and musings on algorithmic decidability. We will frequently make use of the following definition.

**Definition 5.1** (fundamental group of pairs or triples of maps) The *fundamental group* of  $(\phi_1, \phi_2)$  (in  $\text{Alg}^3$  or  $\text{Alg}^{(3,1)}$ ) is their pushout. Similarly, the *fundamental group* of  $(\phi_1, \phi_2, \phi_3)$  (in  $\text{Alg}^4$  or  $\text{Alg}^{(4,2)}$ ) is the group that results from pushing out the three homomorphisms.

Alternatively, the fundamental group of  $(\phi_1, \phi_2)$  or  $(\phi_1, \phi_2, \phi_3)$  is the fundamental group of the corresponding topological space constructed from these maps.

### 5A Nonequivalent group trisections of the same group

Here we present a few examples of nonequivalent group trisections of the same group. Specifically, we provide nonequivalent triples  $(\phi_1, \phi_2, \phi_3)$  in  $\text{Alg}^{(4,2)}/\sim$  with the same fundamental group, where the nonequivalence follows from known topological results about the knotted surfaces realizing these groups as their fundamental group. We leave many details of the calculations to the reader, but see [3; 37] for a more thorough treatment of the following examples, as well as [23, Section 4.1] for a description of how to calculate fundamental groups from tri-plane diagrams.

Note that in the case of closed 4-manifolds, there are as many (nonequivalent) group trisections of the trivial group as there are exotic smooth structures on simply connected 4-manifolds. As a given 4-manifold with an embedded surface can admit nonequivalent group trisections of the same group, which correspond to different surfaces, the theory of group trisections for (4-manifold, knotted surface) pairs is even richer than that for 4-manifolds alone.

**5A.1 Unknotted  $\mathbb{RP}^2$ s in  $S^4$**  For our first example of nonequivalent group trisections of knotted surface groups, consider the following.

**Corollary 5.2** *There exist two elements of  $\text{Alg}^{(4,2)}$  which are not equivalent in the quotient  $\text{Alg}^{(4,2)}/\sim$ , but both push out to  $\mathbb{Z}/2\mathbb{Z}$ . Under topological realization these elements correspond to unknotted  $\mathbb{RP}^2$ s in  $S^4$ .*

**Proof** As shown in [31, Figure 15], we can represent two unknotted  $\mathbb{RP}^2$ s in  $S^4$ , with Euler numbers  $\pm 2$ , by the tri-plane diagrams in Figures 35 and 36. Both  $\mathbb{RP}^2$ s produce group trisections of  $\mathbb{Z}/2\mathbb{Z}$ , but by **Theorem 4.8** these group trisections cannot be equivalent as the surfaces are not isotopic. Here we include some details of the construction of these group trisections for the sake of illustration, but see [3,

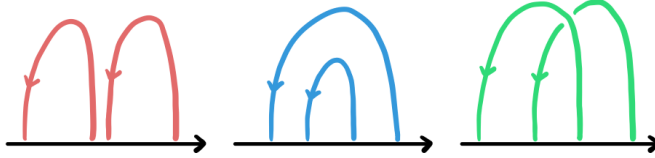


Figure 35: A tri-plane diagram for the unknotted  $\mathbb{R}P^2$  in  $S^4$  with Euler number  $-2$ .

Section 4.2.3] for a full treatment, including an example of the use of [Theorem 2.10](#) to recover the  $\gamma$  (green) tangle from the group trisection.

As  $\mathbb{R}P^2$  is nonorientable, it is not possible to consistently orient the tri-plane diagrams, but we choose arbitrary orientations for each tangle separately in order to write down the group trisection maps; the choice of orientations here will not matter. For both  $\mathbb{R}P^2$ s, presentations for the groups making up the initial three epimorphisms of the group trisection are as follows, where (abusing notation)  $T_2^\alpha$ ,  $T_2^\beta$  and  $T_2^\gamma$  are the trivial tangles as shown in Figures 35 and 36:

$$\begin{aligned} \pi_1(\Sigma_0 - \{p_1, \dots, p_4\}, *) &= \langle p_1, p_2, p_3, p_4 \mid p_1 p_2 p_3 p_4 = 1 \rangle, \\ \pi_1(H_0 - T_2^\alpha, *) &= \langle x_1, x_2 \rangle, \\ \pi_1(H_0 - T_2^\beta, *) &= \langle y_1, y_2 \rangle, \\ \pi_1(H_0 - T_2^\gamma, *) &= \langle z_1, z_2 \rangle. \end{aligned}$$

Below we write down the initial three maps (in other words, the elements in  $\text{Alg}^{(4,2)}$ ) for the  $\mathbb{R}P^2$  with Euler number  $-2$ , corresponding to the  $\alpha$  tangle (left/red),  $\beta$  tangle (center/blue) and  $\gamma$  tangle (right/green):

$$\begin{aligned} \alpha: \quad & p_1 \mapsto x_1, \quad p_2 \mapsto x_1^{-1}, \quad p_3 \mapsto x_2, \quad p_4 \mapsto x_2^{-1}, \\ \beta: \quad & p_1 \mapsto y_1, \quad p_2 \mapsto y_2, \quad p_3 \mapsto y_2^{-1}, \quad p_4 \mapsto y_1^{-1}, \\ \gamma: \quad & p_1 \mapsto z_1, \quad p_2 \mapsto z_2, \quad p_3 \mapsto z_1^{-1}, \quad p_4 \mapsto z_1 z_2^{-1} z_1^{-1}. \end{aligned}$$

The only difference between the initial three maps for the  $\mathbb{R}P^2$  with Euler number  $+2$  and those above is a slight change in the map for the  $\gamma$  tangle (the maps for the  $\alpha$  and  $\beta$  tangles are unchanged), corresponding to the crossing change between the two tangles:

$$\gamma: \quad p_1 \mapsto z_1, \quad p_2 \mapsto z_2, \quad p_3 \mapsto z_2 z_1^{-1} z_2^{-1}, \quad p_4 \mapsto z_2^{-1}.$$

As the initial maps determine the entire pushout cube, it is sufficient to provide these maps; to recover the other maps in the cube, push out repeatedly until the entire cube is formed. □

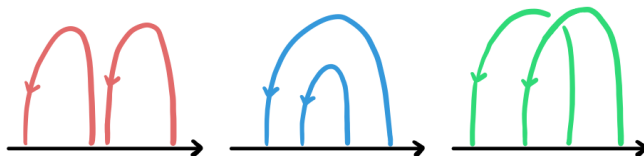


Figure 36: A tri-plane diagram for the unknotted  $\mathbb{R}P^2$  in  $S^4$  with Euler number  $+2$ .

**5A.2 Twist-spun torus knots in  $S^4$**  A well-known family of twist-spun torus knots gives the following corollary of a combination of [11] and Theorem 4.8.

**Corollary 5.3** *There exists a collection of three elements of  $\text{Alg}^{(4,2)}$  which are not equivalent in the quotient  $\text{Alg}^{(4,2)}/\sim$ , but push out to the same group. Under topological realization these elements correspond to knotted spheres in  $S^4$  which have isomorphic knotted surface groups.*

**Proof** Under topological realization, these elements of  $\text{Alg}^{(4,2)}$  correspond to the knotted spheres  $S^2 \hookrightarrow S^4$  listed below:

- $\tau^2 T_{3,5}$ , the 2-twist spin of the (3, 5)-torus knot;
- $\tau^3 T_{2,5}$ , the 3-twist spin of the (2, 5)-torus knot;
- $\tau^5 T_{2,3}$ , the 5-twist spin of the (2, 3)-torus knot.

Presentations for the corresponding knotted surface groups are given by, for instance,

$$\pi_1(S^4 - \tau^2 T_{3,5}, *) \cong \langle x, y \mid x^3 = y^5, [a^2, x] = 1, [a^2, y] = 1 \rangle,$$

where  $a = a(x, y)$  is a meridian of  $T_{3,5}$  expressed as a word in the nonmeridional generators  $x, y$  coming from the genus 1 Heegaard decomposition of the 3-sphere. Recall that this Heegaard splitting leads to the presentation  $\pi_1(S^3 - T_{3,5}, *) \cong \langle x, y \mid x^3 = y^5 \rangle$  of the group of the torus knot complement. If we write the first homology of the complement multiplicatively generated by  $H_1(S^4 - \tau^2 T_{3,5}) \cong \langle t \rangle$ , then under abelianization the generators of the knot group map to  $x \mapsto t^5, y \mapsto t^3$ , and we pick our orientations so that the meridian maps to  $a \mapsto t$ .

To write down the presentations for the other knot groups, we will abuse notation and reuse the letters  $x, y$  for the generators of the torus knot complement and  $a$  for a choice of meridian. With this,

$$\pi_1(S^4 - \tau^3 T_{2,5}, *) \cong \langle x, y \mid x^2 = y^5, [a^3, x] = 1, [a^3, y] = 1 \rangle,$$

$$\pi_1(S^4 - \tau^5 T_{2,3}, *) \cong \langle x, y \mid x^2 = y^3, [a^5, x] = 1, [a^5, y] = 1 \rangle.$$

However all three of these groups  $\pi_1(S^4 - \tau^i T_{j,k}, *)$  are abstractly isomorphic to a direct product  $\mathbb{Z} \times \text{Dod}^*$  of the integers with the binary dodecahedral group  $\text{Dod}^*$ ; see [43].

See [37, Section 16] for explicit constructions of group trisections of these groups which come from the tri-plane diagrams described in [31, Figure 20]. These group trisections are not equivalent, which follows from Theorem 4.8 together with Gordon’s observation [11] that  $\tau^2 T_{3,5}, \tau^3 T_{2,5}$  and  $\tau^5 T_{2,3}$  are all distinct, nonisotopic 2-knots. Gordon shows that the minimal exponent of a meridian’s power which is central in the groups  $\pi_1(S^4 - \tau^i T_{j,k}, *)$  has to divide the twisting parameter. Pick a generator  $p$  of the punctured sphere group and follow it through the group trisection cube; we call its image in the final knotted sphere group  $p$  as well. Observe that the image of  $p$  is a meridional generator, and thus the smallest power of  $p$  that lands these elements in the center is an invariant distinguishing the group, and consequently the group trisections. □

**Remark 5.4** Corollary 5.3 generalizes. Take coprime integers  $p, q, r \geq 2$  and consider the following knotted 2-spheres:

- $\tau^p T_{q,r}$ , the  $p$ -twist spin of the  $(q, r)$ -torus knot;
- $\tau^q T_{p,r}$ , the  $q$ -twist spin of the  $(p, r)$ -torus knot;
- $\tau^r T_{p,q}$ , the  $r$ -twist spin of the  $(p, q)$ -torus knot.

Since  $p, q, r$  are pairwise coprime, all three of these knotted spheres  $\tau^i T_{j,k} : S^2 \hookrightarrow S^4$  for  $\{i, j, k\} = \{p, q, r\}$  have the same fundamental group for their complement. By Zeeman [43], all of them are fibered by the punctured bounded Brieskorn sphere  $\Sigma(p, q, r)$ , but these fibrations have monodromies with different periods. If the fiber is a homology sphere, one can show that the resulting groups of the 2-knots are abstractly isomorphic. The example for  $(p, q, r) = (2, 3, 5)$  is also discussed by Boyle [4], where he starts with Zeeman's observation that these twist spun 2-knots are fibered by a punctured Poincaré homology sphere  $\Sigma(2, 3, 5)$ . Even though they share the same group, these are nonisotopic 2-knots and one way to distinguish them is to look at the smallest power of a meridian which lies in the center of the knot group. This family of knots is further discussed in [11].

**Remark 5.5** In [42], Suciu constructs an infinite family of ribbon 2-knots  $R_i : S^2 \hookrightarrow S^4$  in the 4-sphere, each of which has  $\pi_1(S^4 - R_i, *)$  isomorphic to the trefoil group. The 2-knots are pairwise nonisotopic, and can be distinguished by the  $\mathbb{Z}\pi_1$ -module structure of  $\pi_2$  of their complements. By bridge trisecting these examples, one can construct an infinite family of elements in  $\text{Alg}^{(4,2)}$  satisfying the same properties as in Corollary 5.3. See [37, Section 17].

## 5B Algebraic version of the smooth unknotting conjecture

The smooth unknotting conjecture posits that an embedded sphere in  $S^4$  is smoothly unknotted if and only if the group of its complement is infinite cyclic. Here we use our correspondence in Theorem 4.8 to give an algebraic statement equivalent to the unknotting conjecture. This is in the spirit of a result in [1] which gives a similar group-theoretic conjecture that is equivalent to the smooth 4-dimensional Poincaré conjecture (following the tradition of [40]).

Note that given  $(\phi_1, \phi_2, \phi_3) \in \text{Alg}^{(4,2)}$ , we can tell just from the maps  $(\phi_1, \phi_2, \phi_3)$  whether the resulting surface is connected. Similarly, we can determine the Euler characteristic of the resulting surface directly from  $(\phi_1, \phi_2, \phi_3)$ . We say  $(\phi_1, \phi_2, \phi_3)$  is *spherical* if its corresponding surface is a sphere. Just as mentioned in [1], it follows from Theorem 4.2 that there is a purely group-theoretic condition on whether or not  $X(\bar{\phi}_1, \bar{\phi}_2, \bar{\phi}_3)$ , the 4-manifold built from the associated closed bounding homomorphisms  $\bar{\phi}_i$ , is orientation-preserving diffeomorphic to the 4-sphere with the standard orientation. We say in this case that  $(\phi_1, \phi_2, \phi_3)$  *represents*  $S^4$ . Then from Theorem 4.8, it follows that the group-theoretic conjecture below is equivalent to the smooth unknotting conjecture.

**Conjecture 5.6** For every spherical  $(\phi_1, \phi_2, \phi_3) \in \text{Alg}^{(4,2)}$  which represents  $S^4$ , we have  $(\phi_1, \phi_2, \phi_3) \sim (s_1, s_2, s_3)$ , where  $(s_1, s_2, s_3)$  is the group trisection given by maps

$$s_1, s_2, s_3 : \pi_1(S^2 - \{p_1, p_2\}, *) \twoheadrightarrow \mathbb{Z},$$

with  $s_1 = s_2 = s_3$  and  $s_1(p_1) = 1$  (this corresponds to the unknotted sphere in  $S^4$ ).

### 5C A group-theoretic characterization of knot groups

All knots discussed here are smooth. This subsection is motivated by the question: Exactly which groups arise as fundamental groups of codimension-2 knot complements in  $S^n$ ? An algebraic characterization of such groups was given by Kervaire; namely, such groups are exactly the finitely presentable groups  $G$  with  $H_1(G) = \mathbb{Z}$ ,  $H_2(G) = 0$ , and such that  $G$  can be normally generated by a single element [33]. When  $n = 3$ , Artin gave a characterization for fundamental groups of knot complements in terms of group presentations [2]. When  $n = 4$ , Kamada [24] and González-Acuña [9] have given similar characterizations also in terms of group presentations.

The methods of this paper give alternative group-theoretic characterizations of fundamental groups of knot complements in  $S^n$  for  $n = 3, 4$ , in a unified fashion. We say  $(\phi_1, \phi_2) \in \text{Alg}^{(3,1)}$  is *connected* if the resulting link is connected; note that this could instead be phrased entirely in terms of the maps  $\phi_1$  and  $\phi_2$ . We say  $(\phi_1, \phi_2)$  *represents*  $S^3$  if the resulting 3-manifold is diffeomorphic to the 3-sphere; note that this too can be phrased in an entirely algebraic way due to [Theorem 3.1](#). (See the previous subsection for the corresponding definition when  $n = 4$ .) Then directly from [Theorems 3.2](#) and [4.8](#) we have the following characterizations.

**Corollary 5.7** A group  $G$  is the fundamental group of a knot complement in  $S^3$  if and only if there exists some connected  $(\phi_1, \phi_2) \in \text{Alg}^{(3,1)}$  such that  $(\phi_1, \phi_2)$  represents  $S^3$ , and the fundamental group of  $(\phi_1, \phi_2)$  (recall [Definition 5.1](#)) is  $G$ .

Similarly, a group  $G$  is the fundamental group of a knot complement in  $S^4$  if and only if there exists some spherical  $(\phi_1, \phi_2, \phi_3) \in \text{Alg}^{(4,2)}$  such that  $(\phi_1, \phi_2, \phi_3)$  represents  $S^4$ , and the fundamental group of  $(\phi_1, \phi_2, \phi_3)$  (recall [Definition 5.1](#)) is  $G$ .

### 5D Algorithmic decidability

Recognizing whether a given finitely presented group is the group of the complement of a knotted surface in the 4-sphere is not algorithmic; this is the case  $\mathcal{A} = \mathcal{G}$ ,  $\mathcal{B} = \mathcal{H}_2$  in [10, Theorem 1.1]. It is undecidable whether a given finite presentation describes the fundamental group of the complement of a codimension-2 knot in  $S^n$  for  $n \geq 3$  [13]. The following then is a corollary of a combination of [10, Theorem 1.1] and [Theorem 4.8](#).

**Corollary 5.8** *Given a group  $G$ , it is an undecidable problem to find a group trisection  $(\phi_1, \phi_2, \phi_3)$  in  $\text{Alg}^{(4,2)}$  representing  $S^4$  with  $b > 0$  and fundamental group  $G$  (recall [Definition 5.1](#)), or to show that none exists (because this would decide whether  $G$  is a knotted surface group).*

Recall that there *is* an algorithm for deciding whether the pushouts appearing in group trisections are free groups (see [Remark 4.3](#)).

**Question 5.9** There are many undecidability results in group theory (see for example [\[34\]](#)). Can the above techniques be used to show the existence or nonexistence of an algorithm to recognize if a closed smooth 4-manifold is diffeomorphic to  $S^4$ ? Can the above techniques be used to show the existence or nonexistence of an algorithm to recognize if a smooth embedding of a 2-sphere in  $S^4$  is unknotted? In a different direction, can our framework be useful for determining if there is a polynomial-time algorithm for deciding whether a knot in  $S^3$  is the unknot?

**Question 5.10** Are there extensions of the bijections between sets of the form  $\text{Alg}^n/\sim$  and sets of the form  $\text{Man}^n$  (as in [Sections 3](#) and [4](#)) for  $n > 4$ ? A negative answer says that somehow the free groups and surface groups are responsible for the unique character of low-dimensional topology.

## References

- [1] **A Abrams, D T Gay, R Kirby**, *Group trisections and smooth 4-manifolds*, *Geom. Topol.* 22 (2018) 1537–1545 [MR](#)
- [2] **E Artin**, *Theorie der Zöpfe*, *Abh. Math. Sem. Univ. Hamburg* 4 (1925) 47–72 [MR](#)
- [3] **S Blackwell**, *Combinatorial and group theoretic approaches to trisected surfaces in 4-manifolds*, PhD thesis, University of Georgia (2022) [MR](#) Available at <https://www.proquest.com/docview/2709946407>
- [4] **J Boyle**, *Classifying 1-handles attached to knotted surfaces*, *Trans. Amer. Math. Soc.* 306 (1988) 475–487 [MR](#)
- [5] **M Clay**, *Free groups and folding*, from “Office hours with a geometric group theorist” (M Clay, D Margalit, editors), Princeton Univ. Press (2017) 66–84 [MR](#)
- [6] **T Cochran**, *Ribbon knots in  $S^4$* , *J. London Math. Soc.* 28 (1983) 563–576 [MR](#)
- [7] **B Farb, D Margalit**, *A primer on mapping class groups*, Princeton Mathematical Series 49, Princeton Univ. Press (2012) [MR](#)
- [8] **D Gay, R Kirby**, *Trisecting 4-manifolds*, *Geom. Topol.* 20 (2016) 3097–3132 [MR](#)
- [9] **F González-Acuña**, *A characterization of 2-knot groups*, *Rev. Mat. Iberoamericana* 10 (1994) 221–228 [MR](#)
- [10] **F González-Acuña, C M Gordon, J Simon**, *Unsolvable problems about higher-dimensional knots and related groups*, *Enseign. Math.* 56 (2010) 143–171 [MR](#)
- [11] **C M Gordon**, *Some higher-dimensional knots with the same homotopy groups*, *Quart. J. Math. Oxford Ser.* 24 (1973) 411–422 [MR](#)

- [12] **C M Gordon**, *Knots in the 4-sphere*, Comment. Math. Helv. 51 (1976) 585–596 [MR](#)
- [13] **C M Gordon**, *Some embedding theorems and undecidability questions for groups*, from “Combinatorial and geometric group theory”, London Math. Soc. Lecture Note Ser. 204, Cambridge Univ. Press (1995) 105–110 [MR](#)
- [14] **C M Gordon, J Luecke**, *Knots are determined by their complements*, J. Amer. Math. Soc. 2 (1989) 371–415 [MR](#)
- [15] **C Hayashi**, *Stable equivalence of Heegaard splittings of 1-submanifolds in 3-manifolds*, Kobe J. Math. 15 (1998) 147–156 [MR](#)
- [16] **J Hempel**, *3-manifolds*, Annals of Mathematics Studies 86, Princeton Univ. Press (1976) [MR](#)
- [17] **J A Hillman**, *Alexander ideals of links*, Lecture Notes in Math. 895, Springer (1981) [MR](#)
- [18] **M C Hughes, S Kim, M Miller**, *Isotopies of surfaces in 4-manifolds via banded unlink diagrams*, Geom. Topol. 24 (2020) 1519–1569 [MR](#)
- [19] **W Jaco**, *Heegaard splittings and splitting homomorphisms*, Trans. Amer. Math. Soc. 144 (1969) 365–379 [MR](#)
- [20] **W Jaco**, *Stable equivalence of splitting homomorphisms*, from “Topology of manifolds”, Markham, Chicago, IL (1970) 153–156 [MR](#)
- [21] **K Johannson**, *Topology and Combinatorics of 3-manifolds*, Lecture Notes in Math. 1599, Springer (1995) [MR](#)
- [22] **J Johnson**, *Notes on Heegaard splittings*, unpublished (2006)
- [23] **J Joseph, J Meier, M Miller, A Zupan**, *Bridge trisections and classical knotted surface theory*, Pacific J. Math. 319 (2022) 343–369 [MR](#)
- [24] **S Kamada**, *A characterization of groups of closed orientable surfaces in 4-space*, Topology 33 (1994) 113–122 [MR](#)
- [25] **S Kamada**, *Braid and knot theory in dimension four*, Mathematical Surveys and Monographs 95, Amer. Math. Soc., Providence, RI (2002) [MR](#)
- [26] **T Kanenobu, K-i Kazama**, *The peripheral subgroup and the second homology of the group of a knotted torus in  $S^4$* , Osaka J. Math. 31 (1994) 907–921 [MR](#)
- [27] **G Kuperberg**, *Algorithmic homeomorphism of 3-manifolds as a corollary of geometrization*, Pacific J. Math. 301 (2019) 189–241 [MR](#)
- [28] **F Laudenbach, V Poénaru**, *A note on 4-dimensional handlebodies*, Bull. Soc. Math. France 100 (1972) 337–344 [MR](#)
- [29] **C J Leininger, A W Reid**, *The co-rank conjecture for 3-manifold groups*, Algebr. Geom. Topol. 2 (2002) 37–50 [MR](#)
- [30] **J Meier**, *Filling braided links with trisected surfaces*, Algebr. Geom. Topol. 24 (2024) 803–895 [MR](#)
- [31] **J Meier, A Zupan**, *Bridge trisections of knotted surfaces in  $S^4$* , Trans. Amer. Math. Soc. 369 (2017) 7343–7386 [MR](#)
- [32] **J Meier, A Zupan**, *Bridge trisections of knotted surfaces in 4-manifolds*, Proc. Natl. Acad. Sci. USA 115 (2018) 10880–10886 [MR](#)
- [33] **F Michel, C Weber**, *Michel Kervaire work on knots in higher dimensions*, preprint (2014) [arXiv 1409.0704](#)

- [34] **C F Miller, III**, *Decision problems for groups—survey and reflections*, from “Algorithms and classification in combinatorial group theory”, Math. Sci. Res. Inst. Publ. 23, Springer (1992) 1–59 [MR](#)
- [35] **G Perelman**, *Finite extinction time for the solutions to the Ricci flow on certain three-manifolds*, preprint (2003) [arXiv math/0307245](#)
- [36] **K Reidemeister**, *Zur dreidimensionalen Topologie*, Abh. Math. Sem. Univ. Hamburg 9 (1933) 189–194 [MR](#)
- [37] **B M Ruppik**, *Casson–Whitney unknotting, deep slice knots and group trisections of knotted surface type*, PhD thesis, Max Planck Institute for Mathematics and the University of Bonn (2022) Available at <https://nbn-resolving.org/urn:nbn:de:hbz:5-66903>
- [38] **H Schubert**, *Über eine numerische Knoteninvariante*, Math. Z. 61 (1954) 245–288 [MR](#)
- [39] **J Singer**, *Three-dimensional manifolds and their Heegaard diagrams*, Trans. Amer. Math. Soc. 35 (1933) 88–111 [MR](#)
- [40] **J Stallings**, *How not to prove the Poincaré conjecture*, from “Topology seminar”, Ann. of Math. Stud. 60, Princeton Univ. Press (1966) 83–88 [MR](#)
- [41] **J R Stallings**, *Topology of finite graphs*, Invent. Math. 71 (1983) 551–565 [MR](#)
- [42] **A I Suci**, *Infinitely many ribbon knots with the same fundamental group*, Math. Proc. Cambridge Philos. Soc. 98 (1985) 481–492 [MR](#)
- [43] **E C Zeeman**, *Twisting spun knots*, Trans. Amer. Math. Soc. 115 (1965) 471–495 [MR](#)
- [44] **A Zupan**, *Bridge and pants complexities of knots*, J. Lond. Math. Soc. 87 (2013) 43–68 [MR](#)

SB: *Department of Mathematics, University of Virginia  
Charlottesville, VA, United States*

RK: *Department of Mathematics, University of California, Berkeley  
Berkeley, CA, United States*

MK: *Department of Mathematics, University of Chicago  
Chicago, IL, United States*

VL: *Department of Mathematics, University of Connecticut  
Storrs, CT, United States*

BR: *Heinrich-Heine-Universität  
Düsseldorf, Germany*

[blackwell@virginia.edu](mailto:blackwell@virginia.edu), [kirby@math.berkeley.edu](mailto:kirby@math.berkeley.edu), [michael.r.klug@gmail.com](mailto:michael.r.klug@gmail.com),  
[vincent.2.longo@uconn.edu](mailto:vincent.2.longo@uconn.edu), [benjamin.ruppik@hhu.de](mailto:benjamin.ruppik@hhu.de)

<https://seblackwell.com>, <https://math.berkeley.edu/~kirby>,  
<https://mathematics.uchicago.edu/people/profile/michael-klug>,  
<https://math.uconn.edu/person/vincent-longo>, <https://bruppik.de>

Received: 9 August 2023      Revised: 24 April 2024

# ALGEBRAIC & GEOMETRIC TOPOLOGY

[msp.org/agt](https://msp.org/agt)

## EDITORS

### PRINCIPAL ACADEMIC EDITORS

John Etnyre  
[etnyre@math.gatech.edu](mailto:etnyre@math.gatech.edu)  
Georgia Institute of Technology

Kathryn Hess  
[kathryn.hess@epfl.ch](mailto:kathryn.hess@epfl.ch)  
École Polytechnique Fédérale de Lausanne

### BOARD OF EDITORS

|                        |   |                   |   |
|------------------------|---|-------------------|---|
| Julie Bergner          | University of Virginia<br><a href="mailto:jeb2md@eservices.virginia.edu">jeb2md@eservices.virginia.edu</a>          | Thomas Koberda    | University of Virginia<br><a href="mailto:thomas.koberda@virginia.edu">thomas.koberda@virginia.edu</a>  |
| Steven Boyer           | Université du Québec à Montréal<br><a href="mailto:cohf@math.rochester.edu">cohf@math.rochester.edu</a>             | Markus Land       | LMU München<br><a href="mailto:markus.land@math.lmu.de">markus.land@math.lmu.de</a>                     |
| Tara E Brendle         | University of Glasgow<br><a href="mailto:tara.brendle@glasgow.ac.uk">tara.brendle@glasgow.ac.uk</a>                 | Christine Lescop  | Université Joseph Fourier<br><a href="mailto:lescop@ujf-grenoble.fr">lescop@ujf-grenoble.fr</a>         |
| Indira Chatterji       | CNRS & Univ. Côte d'Azur (Nice)<br><a href="mailto:indira.chatterji@math.cnrs.fr">indira.chatterji@math.cnrs.fr</a> | Norihiko Minami   | Yamato University<br><a href="mailto:minami.norihiko@yamato-u.ac.jp">minami.norihiko@yamato-u.ac.jp</a> |
| Octav Cornea           | Université de Montréal<br><a href="mailto:cornea@dms.umontreal.ca">cornea@dms.umontreal.ca</a>                      | Andrés Navas      | Universidad de Santiago de Chile<br><a href="mailto:andres.navas@usach.cl">andres.navas@usach.cl</a>    |
| Alexander Dranishnikov | University of Florida<br><a href="mailto:dranish@math.ufl.edu">dranish@math.ufl.edu</a>                             | Jessica S Purcell | Monash University<br><a href="mailto:jessica.purcell@monash.edu">jessica.purcell@monash.edu</a>         |
| Tobias Ekholm          | Uppsala University, Sweden<br><a href="mailto:tobias.ekholm@math.uu.se">tobias.ekholm@math.uu.se</a>                | Birgit Richter    | Universität Hamburg<br><a href="mailto:birgit.richter@uni-hamburg.de">birgit.richter@uni-hamburg.de</a> |
| Mario Eudave-Muñoz     | Univ. Nacional Autónoma de México<br><a href="mailto:mario@matem.unam.mx">mario@matem.unam.mx</a>                   | Jérôme Scherer    | École Polytech. Féd. de Lausanne<br><a href="mailto:jerome.scherer@epfl.ch">jerome.scherer@epfl.ch</a>  |
| David Futер            | Temple University<br><a href="mailto:dfuter@temple.edu">dfuter@temple.edu</a>                                       | Vesna Stojanoska  | Univ. of Illinois at Urbana-Champaign<br><a href="mailto:vesna@illinois.edu">vesna@illinois.edu</a>     |
| John Greenlees         | University of Warwick<br><a href="mailto:john.greenlees@warwick.ac.uk">john.greenlees@warwick.ac.uk</a>             | Zoltán Szabó      | Princeton University<br><a href="mailto:szabo@math.princeton.edu">szabo@math.princeton.edu</a>          |
| Matthew Hedden         | Michigan State University<br><a href="mailto:mhedden@math.msu.edu">mhedden@math.msu.edu</a>                         | Maggy Tomova      | University of Iowa<br><a href="mailto:maggy-tomova@uiowa.edu">maggy-tomova@uiowa.edu</a>                |
| Kristen Hendricks      | Rutgers University<br><a href="mailto:kristen.hendricks@rutgers.edu">kristen.hendricks@rutgers.edu</a>              | Daniel T Wise     | McGill University, Canada<br><a href="mailto:daniel.wise@mcgill.ca">daniel.wise@mcgill.ca</a>           |
| Hans-Werner Henn       | Université Louis Pasteur<br><a href="mailto:henn@math.u-strasbg.fr">henn@math.u-strasbg.fr</a>                      | Lior Yanovski     | Hebrew University of Jerusalem<br><a href="mailto:lior.yanovski@gmail.com">lior.yanovski@gmail.com</a>  |
| Daniel Isaksen         | Wayne State University<br><a href="mailto:isaksen@math.wayne.edu">isaksen@math.wayne.edu</a>                        |                   |   |

---

See inside back cover or [msp.org/agt](https://msp.org/agt) for submission instructions.


The subscription price for 2025 is US \$760/year for the electronic version, and \$1110/year (+\$75, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP. Algebraic & Geometric Topology is indexed by [Mathematical Reviews](#), [Zentralblatt MATH](#), [Current Mathematical Publications](#) and the [Science Citation Index](#).

Algebraic & Geometric Topology (ISSN 1472-2747 printed, 1472-2739 electronic) is published 9 times per year and continuously online, by Mathematical Sciences Publishers, 2000 Allston Way # 59, Berkeley, CA 94701-4004. Periodical rate postage paid at Oakland, CA 94615-9651, and additional mailing offices. POSTMASTER: send address changes to Mathematical Sciences Publishers, 2000 Allston Way # 59, Berkeley, CA 94701-4004.

---

AGT peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

 **mathematical sciences publishers**

nonprofit scientific publishing

<https://msp.org/>

© 2025 Mathematical Sciences Publishers

# ALGEBRAIC & GEOMETRIC TOPOLOGY

Volume 25 Issue 8 (pages 4437–5174) 2025

---

|   |      |
|---|------|
| Hierarchies for relatively hyperbolic virtually special groups  | 4437 |
| EDUARD EINSTEIN   |      |
| Intersection norms on surfaces and Birkhoff sections for geodesic flows   | 4499 |
| MARCOS COSSARINI and PIERRE DEHORNOY  |      |
| Thin knots and the cabling conjecture   | 4547 |
| ROBERT DEYESO III   |      |
| Linear linkless embeddings: proof of a conjecture by Sachs  | 4585 |
| LYNN STANFIELD  |      |
| Constructing rational homology 3-spheres that bound rational homology 4-balls   | 4599 |
| LISA LOKTEVA  |      |
| Homological stability for the ribbon Higman–Thompson groups   | 4633 |
| RACHEL SKIPPER and XIAOLEI WU   |      |
| A group-theoretic framework for low-dimensional topology, or: how not to study low-dimensional topology?                    | 4667 |
| SARAH BLACKWELL, ROBION KIRBY, MICHAEL KLUG, VINCENT LONGO and BENJAMIN RUPPIK  |      |
| Classification of genus-two surfaces in $S^3$   | 4719 |
| FILIPPO BARONI  |      |
| Meromorphic projective structures: signed spaces, grafting and monodromy  | 4787 |
| SPANDAN GHOSH and SUBHOJOY GUPTA  |      |
| A new twist on modular links from an old perspective  | 4827 |
| KHANH LE  |      |
| Flat fully augmented links are determined by their complements  | 4839 |
| CHRISTIAN MILLICHAP and ROLLAND TRAPP   |      |
| BNSR-invariants of surface Houghton groups  | 4897 |
| NOAH TORGERSON and JEREMY WEST  |      |
| Topological symmetry groups of the generalized Petersen graphs  | 4921 |
| ANGELYNN ÁLVAREZ, ERICA FLAPAN, MARK HUNNELL, JOHN HUTCHENS, EMILLE LAWRENCE, PAUL LEWIS, CANDICE PRICE and RUTH VANDERPOOL |      |
| Crushing surfaces of positive genus   | 4949 |
| BENJAMIN A BURTON, THIAGO DE PAIVA, ALEXANDER HE and CONNIE ON YU HUI   |      |
| Annular links from Thompson’s group $T$   | 5013 |
| LOUISA LILES  |      |
| Realizing pairs of multicurves as cylinders on translation surfaces   | 5031 |
| JULIET AYGUN, JANET BARKDOLL, AARON CALDERON, JENAVIE LORMAN and THEODORE SANDSTROM   |      |
| Involutive Khovanov homology and equivariant knots  | 5059 |
| TAKETO SANO   |      |
| A diagrammatic computation of abelian link invariants   | 5113 |
| DAVID CIMASONI, LIVIO FERRETTI and JESSICA LIU  |      |
| Equivariant double-slice genus, stabilization, and equivariant stabilization  | 5137 |
| MALCOLM GABBARD   |      |
| Coarse and bi-Lipschitz embeddability of subspaces of the Gromov–Hausdorff space into Hilbert spaces                        | 5153 |
| NICOLÒ ZAVA   |      |