

A POLYNOMIAL UPPER BOUND ON REIDEMEISTER MOVES FOR EACH LINK TYPE

MARC LACKENBY

ABSTRACT. For each link type K in the 3-sphere, we show that there is a polynomial p_K such that any two diagrams of K with c_1 and c_2 crossings differ by at most $p_K(c_1) + p_K(c_2)$ Reidemeister moves. As a consequence, the problem of recognising whether a given link diagram represents K is in the complexity class NP and hence can be completed deterministically in exponential time. We calculate this polynomial p_K explicitly for various classes of links.

CONTENTS

1. Introduction	2
2. Normal Surfaces	6
3. Hierarchies	27
4. Handle structures	37
5. The basic theory of arc presentations	40
6. Admissible partial hierarchies	47
7. Extending the partial hierarchy	59
8. Branched surfaces	62
9. Modifications to an admissible surface	65
10. Parallelism and envelopes	70
11. The Euclidean subsurface	85
12. The thin part of the surface	103
13. The Euler characteristic argument	107
14. Modifications in the case of a broken section	109
15. Realising the hierarchy with polynomially bounded binding weight	113
16. Isotoping a link through a handle structure	117
17. From the isotopy to Reidemeister moves	120
18. Explicit polynomials	126
19. The figure-eight knot	130
20. Torus knots	134
References	135

The author was partially supported by EPSRC grant EP/Y004256/1. For the purpose of open access, the author has applied a CC BY public copyright licence to any author accepted manuscript arising from this submission.

1. INTRODUCTION

Haken, Hemion and Matveev showed that the equivalence problem for knots and links is soluble [9, 12, 18]. But the complexity class of this problem is not at all well understood. In particular, it is not known whether there is an efficient algorithm to decide whether two links are the same. One of the most simple and attractive ways of approaching this problem is to find an upper bound on the number of Reidemeister moves required to pass between two diagrams of a link that is a computable function of the number of crossings in each diagram. In [5], Coward and the author provided such a bound and thereby gave a conceptually simple algorithm to solve the link equivalence problem. However, the bound was huge: a tower of exponentials that is exponentially high. In this paper, we show that, if we fix the link type, one can actually find an upper bound that is a polynomial function of the number of crossings.

Theorem 1.1. *For each link type K in the 3-sphere, there is a polynomial p_K with following property. If D_1 and D_2 are two diagrams for K with c_1 and c_2 crossings respectively, then they differ by a sequence of at most $p_K(c_1) + p_K(c_2)$ Reidemeister moves.*

This has the following algorithmic consequence. For a link type K , the K -recognition problem is the question whether a given link diagram represents K .

Theorem 1.2. *For each link type K , the K -recognition problem lies in NP.*

By definition, a problem is in the complexity class NP if, whenever an instance of the problem has an affirmative answer, there is a certificate that provides this affirmation that can be checked in polynomial time. In the case of the K -recognition problem, this certificate takes a particularly simple form. It consists of a fixed diagram of K , together with a sequence of Reidemeister moves taking the given input diagram to this fixed diagram. We note that NP problems can be solved deterministically in exponential time, simply by checking all possible certificates of polynomially bounded length.

Corollary 1.3. *For each link type K , the K -recognition problem can be solved deterministically in exponential time, as a function of the number of crossings in a given link diagram.*

Theorem 1.1 asserts the existence of the polynomial p_K . The following result shows that it is algorithmically computable.

Theorem 1.4. *There is an algorithm that takes, as its input, a diagram of a link K and outputs the polynomial p_K .*

We implement this procedure when K is the figure-eight knot, and obtain the following.

Theorem 1.5. *For K the figure-eight knot, we may set $p_K(c) = (10^{108}c)^{15460896}$.*

At the end of the paper, we will investigate torus links, and we will find the following explicit polynomial.

Theorem 1.6. *When K is a torus knot, we may set $p_K(c) = (10^{11}c)^{299666}$.*

Again this result has the following algorithmic consequence, which was recently proved by Baldwin and Sivek [1] using other techniques.

Theorem 1.7. *Deciding whether a diagram represents a torus knot is in NP.*

Proof. We are given a diagram D for a link K with c crossings. We need to show that when K is a torus knot, then there is a certificate that establishes this. Suppose that K is the (p, q) -torus knot, where $0 < |p| \leq q$. Note that K has a standard diagram D' with $c' = q(|p| - 1)$ crossings, and this realises the crossing number of K , by a theorem of Murasugi [19]. Our certificate is a sequence of Reidemeister moves, with length bounded by a polynomial function of c , taking D to D' . By Theorem 1.7, the number of Reidemeister moves required to take D to D' is at most $(10^{11}c)^{299666} + (10^{11}c')^{299666}$. Since $c' \leq c$, we deduce that the number of Reidemeister moves is at most $2(10^{11}c)^{299666}$. \square

In general, the degree of p_K is quite easy to compute. Without loss of generality, $p_K(c) = (ac)^d$, for constants a and d . When K is non-split, the degree d is solely a function of the length of a certain type of hierarchy for the exterior of K . But the coefficient a is related to finer details of the hierarchy and also the link.

It remains a very interesting question whether there is a *universal* polynomial p that works for every link K . If there were such a polynomial, then this would imply that the link equivalence problem lies in NP. That would represent a very significant improvement on the known computational complexity of this important problem. (See [15] for the best known upper bound on its computational complexity.)

1.1. Outline of the proof. This paper can be viewed as a continuation of the paper [17]. There, an explicit polynomial upper bound on Reidemeister moves was provided for the unknot. A variety of techniques were used: Dynnikov's work on arc presentations of the unknot [8], normal surface theory, and a detailed analysis of surfaces carried by Euclidean branched surfaces. An outline of the argument there is as follows. Given a diagram D of the unknot, convert it to a rectangular diagram. This specifies an arc presentation for the unknot, in the following sense. An open book decomposition for the 3-sphere is fixed, with binding circle an unknot and with pages that are open discs. A link is an arc presentation if it intersects binding circle at finitely many points, and elsewhere intersects each open page in either the empty set or a properly embedded arc. (For the terminology related to rectangular diagrams and arc presentations, see [6], [8], or Section 5).

The unknot bounds a spanning disc, which can be placed in a form that respects the arc presentation, called admissible form. A good measure of the disc's complexity is the number of times that it intersects the binding circle, which we term its *binding weight*. Now, arising naturally from an arc presentation of a link, there is a polyhedral structure of the link's exterior. By using well-known estimates on the complexity of normal surfaces in such a polyhedral structure, due to Hass and Lagarias [10], we obtain an exponential upper bound on the disc's binding weight. In [8], Dynnikov described a collection of moves that can be used to reduce the binding weight of the spanning disc without increasing the arc index of the presentation. Using this, it is straightforward to find an exponential bound on the number of Reidemeister moves. But one can improve this to a polynomial by using normal surface theory in a deeper way. If the spanning disc has high binding weight, then one can find many parts of the surface that are normally parallel. Using this observation, it was shown in [17] that one can reduce the binding weight of the spanning disc by a definite factor using a single move. This then leads to the polynomial bound on Reidemeister moves.

Many of these ideas can be translated to the setting of general non-split links. The main thing that is absent is a spanning disc. Instead, we use an entire hierarchy for the link exterior. We show that, after polynomially many moves on the arc presentation, all the surfaces in this hierarchy can be arranged to have polynomially bounded binding weight. The hierarchy specifies a handle structure on the exterior of K . By reattaching a regular neighbourhood of K with a simple handle structure, we obtain a handle structure on the 3-sphere, and it is easy to arrange that each component of K runs through a 0-handle and a 1-handle. We pick a fixed diagram D' for K . The aim is to convert D to D' using a controlled number of Reidemeister moves. We pick a fixed copy of D' drawn on a specified 0-handle. This specifies a realisation K' of the knot within this 0-handle. The next step in the argument is to pick a fixed ambient isotopy within this handle structure taking K to K' . This is achieved using a series of moves, for example, sliding part of the knot across a 2-handle. Each 2-handle is embedded within S^3 in a distorted way, but this distortion is controlled. So, when the knot is slid across this 2-handle, one can bound the resulting number of Reidemeister moves in the diagrams obtained by projecting onto the horizontal plane. In this way, we obtain an upper bound on the number of Reidemeister moves required to pass between the given diagram D and the fixed diagram D' . This is a polynomial function $p_K(c(D))$, where $c(D)$ is the crossing number of D . Therefore, if one is given two diagrams D_1 and D_2 of K with crossing number c_1 and c_2 respectively, the number of moves required to pass from one to the other via D' is at most $p_K(c_1) + p_K(c_2)$.

1.2. Structure of the paper. In Section 2, we give an introduction to some aspects of normal surface theory, and we prove some new results. Possibly the most significant of these is Theorem 2.11, which asserts that JSJ surfaces in a compact orientable irreducible 3-manifold with incompressible boundary can always be realised as fundamental normal surfaces. The main aim of this section is to show that many surfaces in a hierarchy are *exponentially controlled*. This means that for any triangulation of the manifold with t tetrahedra, the surface can be arranged to intersect the 1-skeleton at most ab^t times, where a and b are constants depending only on the 3-manifold and the surface, but not the triangulation. The results in Section 2 are then used in Section 3, where the main aim is to show that the exterior of every non-split link, other than the unknot, admits an exponentially controlled hierarchy. This means that, for every manifold appearing in the hierarchy with its inherited boundary pattern, the next surface in the hierarchy is exponentially controlled. We therefore fix such a hierarchy H for our given link K . Although Sections 2 and 3 are focused on triangulations, it is also convenient to use handle structures. We therefore give a very brief introduction to handle structures in Section 4.

Section 5 introduces our other main tools: arc presentations and rectangular diagrams. It is primarily expository, but contains a few new results. We mostly discuss admissible form for a surface, and the resulting singular foliation that such a surface inherits. In Section 6, we consider a partial hierarchy for a link exterior where each surface is admissible form. We define a polyhedral decomposition, a handle structure and a triangulation for the exterior of this partial hierarchy. In Section 7, we consider the next surface S in the hierarchy. This can be arranged to be normal, with an exponential upper bound on its weight, and a further modification makes it admissible.

In Section 8, we introduce branched surfaces. We show that the admissible surface S is naturally carried by a branched surface B . Each patch of B comes from parallel tiles of S .

Section 9 mostly discusses the modifications to an admissible surface S that were introduced by Dynnikov [8], and which drew on work of Bennequin [2] and Birman-Menasco [3]. Here, one considers the singular foliation on S . There are primarily three types of singularity: interior vertices (where the interior of the surface intersects the binding circle transversely), interior saddles (where the interior of the surface is arranged like a saddle with respect to the circle-valued function on the complement of the binding circle) and poles (where the surface has a local maximum or minimum with respect to this height function). However, there are two further types of singularity on ∂S : boundary vertices and boundary saddles. Coming out of each saddle, there are four leaves of this foliation. Such leaves are called *separatrices*. They typically end on a boundary or interior vertex. The *valence* of a vertex is the number of separatrices emanating from it. The moves discussed by Dynnikov arise when there is an interior vertex of valence at most 3 or a boundary vertex of valence 1. He showed that when K is the unknot in an arc presentation, then any admissible spanning disc always has such a vertex. The moves then modify the link and the spanning disc, and result in a reduction of the disc's binding weight. In Section 9, we recall these moves, but we introduce one further modification. In the case where there was a boundary vertex of valence 1, Dynnikov was able to slide the two arcs of K incident to this vertex across a subset of the disc, and thereby reduce the arc index of K . Such a move is not appropriate in our context. The boundary of our surface S runs over previous surfaces in the hierarchy, and so one cannot simply move ∂S across a portion of S and keep it properly embedded. Instead, we introduce an alternative move called a wedge insertion.

Each of the above moves may only reduce the binding weight of S by 1 or 2. As in [17], a crucial part of our argument is to perform such moves but with a much greater reduction in binding weight. This is possible when there is a collection of vertices of S , each with low valence, and with parallel stars. Here, the *star* of a vertex is the union of the leaves of the singular foliation emanating from it. We define parallelism in terms of the branched surface B . Therefore in Section 10, we revisit the moves from Section 9 in the context of many low valence vertices with parallel stars. These modifications change the surface and also the branched surface B . It is important that they do not make B more complicated. Therefore, we define a notion of complexity for B and show that it does not increase. In fact, we have to introduce a second branched surface B' , in order to control what happens during wedge insertion. We package B and B' , along with the admissible surface, in a structure that we call an admissible envelope.

The above argument works well when there are ‘many’ vertices with low valence and having parallel stars. Here, ‘low’ means ‘at most 3’ for interior vertices. The definition of ‘low’ valence for boundary vertices is slightly more complicated. Our task is to show that there are many such vertices, unless S has polynomially bounded binding weight. A fairly simple Euler characteristic argument, presented in Section 13, establishes that there are these vertices unless there are many vertices with valence 4 in the double S_+ of S . If there are such 4-valent vertices in S_+ then their incident tiles patch together to form a subsurface of S_+ with a natural

Euclidean metric. In Section 11, we consider this subsurface. In [17], this subsurface was also considered. It was possible to show that if the Euclidean subsurface was sufficiently big, then it contained a large rectangle or annulus with that the property that opposite sides of the rectangle or annulus are parallel. It was then argued that this rectangle or annulus could be patched together to form a normal torus that was a normal summand for (some multiple of) the spanning disc. This contradicted an initial assumption that the surface had no such summand. Here, we need to make a similar argument, but a somewhat more technical one. In [17], it was possible to hypothesise that the spanning disc had particularly nice properties as a normal surface, whereas in our situation all we have is that S is exponentially controlled. So we need to present a more technical variation of the argument in [17]. However, the conclusion is the same, that the Euclidean subsurface of S_+ cannot be too large, and hence there must in fact be many vertices of S with parallel stars and low valence.

We piece together these various parts of the argument in Section 15. We are given a diagram D for K with crossing number c . An isotopy of D makes it into a rectangular diagram with arc index n at most $(81/20)c$. It then determines an arc presentation for K . We have an exponentially controlled hierarchy H for the exterior of K . In Theorem 15.1, we establish the existence of a polynomial q so that H may be arranged to have binding weight at most $q(n)$, and this is achieved using at most $n^2q(n)$ exchange moves and cyclic permutations of K . This is achieved by arranging each of the surfaces of the hierarchy in turn to have polynomially bounded binding weight.

This hierarchy H induces a handle structure on the exterior of K . We then attach a regular neighbourhood of K , to obtain a handle structure on the 3-sphere. We have arranged that all the 2-handles are thin regular neighbourhoods of admissible surfaces, and that these have polynomially bounded binding weight. In Sections 16, 17 and 18, we isotope the link K through this handle structure until it lies within a 0-handle, and also so that its projection to the horizontal plane in \mathbb{R}^3 is some fixed diagram D' for K . By considering the projection of K to this plane, we need at most polynomially many Reidemeister moves to convert the given diagram D to D' . This then completes the proof of Theorem 1.1.

In Section 19, we run through the details of this construction for the figure-eight knot. We find an explicit hierarchy H that is exponentially-controlled, and from this it is quite straightforward to compute the polynomial q given above. We also bound the number of Reidemeister moves needed when isotoping K through the handle structure. This leads to the proof of Theorem 1.5, where an explicit polynomial upper bound for Reidemeister moves is given. A similar analysis is given in Section 20 in the case of torus knots.

1.3. Notation. If M is a compact manifold and X is a closed subset of M , the manifold $M \setminus\setminus X$ obtained by cutting M along X is the following compactification of $M - X$. We assign M some Riemannian metric, and then $M \setminus\setminus X$ is the completion of $M - X$ with respect to its path metric. The inclusion $M - X \rightarrow M$ induces a map $M \setminus\setminus X \rightarrow M$.

2. NORMAL SURFACES

2.1. Definition. Let \mathcal{T} be a triangulation of a compact 3-manifold M . Then a surface S properly embedded in M is *normal* with respect to \mathcal{T} if it intersects each

tetrahedron in a collection of triangles and squares, as shown in Figure 1. Triangles and squares are known as *elementary normal discs*.

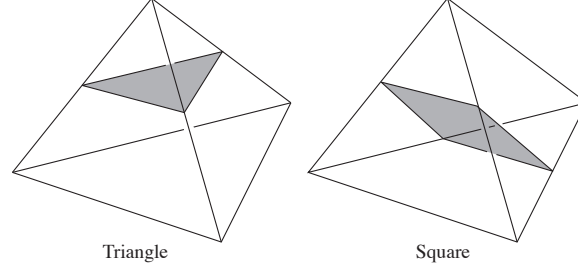


FIGURE 1. Elementary normal discs in a tetrahedron

A normal surface S can be encoded by a collection of non-negative integers as follows. One simply counts the number of triangles and squares of each type in each tetrahedron. There are 7 types of triangle and square in each tetrahedron. So, if \mathcal{T} contains t tetrahedra, then we obtain a list of $7t$ non-negative integers. Each integer is a *co-ordinate* of S and the complete list is called the *normal surface vector* for S and is denoted (S) .

2.2. The weight of a surface. Let M be a compact 3-manifold with a triangulation \mathcal{T} . Let S be a properly embedded surface that is disjoint from the vertices of \mathcal{T} and that intersects the edges transversely. The *weight* $w(S)$ of S is its number of points of intersection with the edges.

2.3. Boundary patterns. We will need to conduct normal surface theory in the presence of boundary patterns. Therefore, in this subsection, we recall some of the terminology relating to boundary patterns.

A *boundary pattern* P for a compact orientable 3-manifold M is a (possibly empty) collection of disjoint simple closed curves and trivalent graphs embedded in ∂M .

A surface embedded in M is *clean* if it is disjoint from P .

We say (M, P) is *boundary-irreducible* if for any clean disc D properly embedded in M , there is a clean disc D' in ∂M such that $\partial D' = \partial D$. This is equivalent to the requirement that $\partial M - P$ is incompressible.

Let S be a properly embedded surface in M transverse to P . A *boundary-compression disc* for S (with respect to the pattern P) is a clean disc D embedded in M such that

- (1) $D \cap S$ is an arc α in ∂D ;
- (2) $\partial D \setminus \alpha = D \cap \partial M$;
- (3) α does not separate off a component of $S \setminus \alpha$ that is a clean disc.

The surface S is *boundary-compressible* (with respect to the pattern) if it admits a boundary-compression disc; otherwise it is *boundary-incompressible*.

We say that a *pattern-isotopy* of a properly embedded surface S in M is an ambient isotopy that preserves the boundary pattern throughout. When two surfaces differ by such an isotopy, we say that they are *pattern-isotopic*.

Two surfaces properly embedded in M are *strongly equivalent* if there is a homeomorphism of M , whose restriction to ∂M is pattern-isotopic to the identity, taking one surface to the other.

2.4. Boundary patterns and normal surfaces. It is well known that incompressible boundary-incompressible surfaces may be placed in normal form, under rather mild hypotheses. The following version of this result, which is Corollary 3.3.25 in [18], achieves this in the presence of boundary patterns.

Proposition 2.1. *Let (M, P) be a compact orientable irreducible boundary-irreducible 3-manifold with boundary pattern. Let \mathcal{T} be a triangulation of M in which P is simplicial. Let S be an incompressible, boundary-incompressible surface properly embedded in (M, P) , other than a 2-sphere or a boundary-parallel disc intersecting P in zero or two points. Then S is pattern-isotopic to a surface S' that is normal with respect to \mathcal{T} . Moreover, we can ensure that the weight of S' is at most the weight of S .*

2.5. Annuli and tori in M . We present some definitions from [18].

Let M be a compact orientable 3-manifold with a boundary pattern P .

An annulus A in ∂M is *almost clean* if its intersection with P is a (possibly empty) collection of disjoint core curves in the interior of A .

An annulus A in ∂M is a *necklace annulus* if P is disjoint from ∂A and separates the components of ∂A and there is an arc in A joining the two components of ∂A such that $A \cap P$ is a single point.

A clean annulus properly embedded in M is *essential* if it is incompressible and boundary-incompressible (with respect to the pattern) and is not parallel to an almost clean annulus in ∂M . Note that when (M, P) is irreducible and boundary-irreducible, then a clean annulus is essential if and only if it is incompressible and not parallel to an almost clean annulus in ∂M .

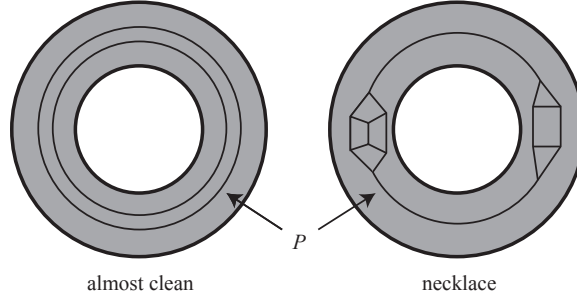


FIGURE 2. Almost clean annulus and necklace annulus in ∂M

A torus T properly embedded in (M, P) is *essential* if it is incompressible and not parallel to a torus in ∂M . (In this definition, it is not relevant whether this torus in ∂M is clean or not.)

We say that (M, P) is *simple* if it is irreducible and boundary-irreducible and contains no properly embedded essential tori and no properly embedded essential clean annuli.

2.6. Normal summation and fundamental surfaces. Let M be a compact orientable 3-manifold with a triangulation \mathcal{T} .

We say that a properly embedded normal surface S is the *sum* of two properly embedded normal surfaces S_1 and S_2 if their normal surface vectors satisfy $(S) = (S_1) + (S_2)$. We write $S = S_1 + S_2$.

A normal surface S is a *fundamental surface* if it cannot be written as the sum of two non-empty normal surfaces.

The following was proved by Hass and Lagarias in [10].

Proposition 2.2. *Let M be a compact 3-manifold and let \mathcal{T} be a triangulation of M with t tetrahedra. Then, for any fundamental normal surface S , each coordinate of S is at most $t2^{7t+2}$.*

We note the following immediate consequence.

Corollary 2.3. *Let M , \mathcal{T} and t be as above. Then for any fundamental normal surface S , $w(S) \leq t^2 2^{7t+6}$.*

2.7. A topological interpretation of normal summation. There is a topological way of viewing normal summation, which is as follows.

Suppose that S is a properly embedded normal surface that is a sum of normal surfaces S_1 and S_2 . We may perform a small pattern-isotopy to S_1 , keeping it normal throughout, so that afterwards S_1 and S_2 intersect transversely in a collection of properly embedded simple closed curves and arcs. Then S is obtained from $S_1 \cup S_2$ by removing a regular neighbourhood of $S_1 \cap S_2$, and then attaching annuli, Möbius bands and discs which join S_1 and S_2 together in this regular neighbourhood. Along each curve and arc of $S_1 \cap S_2$, there are two possible ways to perform this smoothing, but only one correctly gives the normal surface S . This is called a *regular switch*. If the smoothing is performed the other way, it is called an *irregular switch*. In a regular neighbourhood of each component of $S_1 \cap S_2$, there is an I -bundle over this component properly embedded in $M \setminus S$ called a *trace surface*. It is an annulus, disc or Möbius band. A *patch* is a component of $(S_1 \cup S_2) \setminus (S_1 \cap S_2)$.

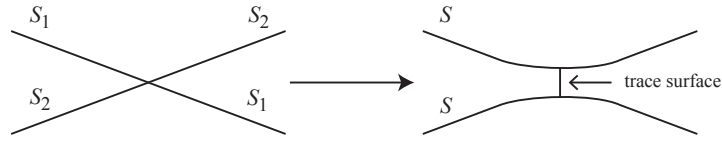


FIGURE 3. Normal summation

A normal sum $S = S_1 + S_2$ of two surfaces is *in reduced form* if S cannot be written as a sum $S'_1 + S'_2$ where each S'_i is pattern-isotopic to S_i and $S'_1 \cap S'_2$ has fewer components than $S_1 \cap S_2$.

There is a well-established theory of summation of normal surfaces, and we record some of its main results for our future reference. These are essentially proved in 4.1.8, 4.1.36, 6.3.20 and 6.3.21 in [18].

Theorem 2.4. *Let (M, P) be a compact orientable irreducible boundary-irreducible 3-manifold with boundary pattern. Let \mathcal{T} be a triangulation of M in which P is simplicial. Let S be a properly embedded, incompressible, boundary-incompressible, normal surface that has least possible weight, among all normal surfaces that are*

pattern-isotopic to it. Assume either that S is connected or that no component of S is a sphere, clean disc, projective plane, boundary-parallel torus, or clean inessential annulus. Suppose that S can be expressed as a normal sum $S_1 + S_2$, where S_1 and S_2 are in general position and non-empty, and the summation is in reduced form. Then the following hold.

- (1) No patch is a disc disjoint from ∂M or a disc that intersects ∂M in a single clean arc.
- (2) Every patch is incompressible and boundary-incompressible (with respect to the pattern P).
- (3) The trace surfaces are incompressible in $M \setminus S$.
- (4) S_1 and S_2 are incompressible and boundary-incompressible (with respect to the pattern P).
- (5) No component of S_1 or S_2 is a sphere, clean disc or projective plane.
- (6) If S is orientable, then no component of S_1 or S_2 is a clean annulus that is parallel to an almost clean annulus in ∂M .
- (7) If S is orientable and has least weight up to strong equivalence, then no component of S_1 or S_2 can be parallel to a toral component T of ∂M such that $|S \cap T| \leq 1$.
- (8) If S is orientable, then no component of S_1 or S_2 can be parallel to a toral component T of ∂M disjoint from P .

Note that (3) is not explicitly stated in [18] but it is an immediate consequence of (1) and (2). This is because a compression disc for a trace surface must be parallel to the boundary of the surface, and hence can be used to create a compression disc for an incident patch. Then by (2) this patch is a disc disjoint from ∂M , contradicting (1).

Proposition 4.1.36 in [18], which gives (4) and (5), has the assumption that S is connected. However, this assumption is only required to rule out the possibility that S has a component that is a sphere, clean disc or projective plane. We therefore make this alternative hypothesis in the above theorem. Similarly, Propositions 6.3.20 and 6.3.21 in [18], which give (6) and (7), also assume that S is connected. However, this assumption is made in order to apply Proposition 4.1.36 in [18] and to exclude the possibility that a component of S_1 or S_2 is a boundary-parallel toral component of S or a clean inessential annulus component of S . However, we assume that S has no such component.

Note also that (8) is slightly different from Proposition 6.3.21 in [18]. There, it is assumed that $|T \cap S| \leq 1$, when aiming to establish that no component of S_1 or S_2 can be parallel to a toral component T of ∂M . It is also assumed that S has least weight up to strong equivalence. However, we note that when T is disjoint from P , then the argument of Proposition 6.3.21 in [18] implies that there is a pattern isotopy that reduces the weight of S . The argument reduces to the case where S_1 and S_2 are arranged near T as shown in Figure 4, with S'_2 a component of S_2 that is parallel to T , and the summation $S_1 + S_2$ is in the same direction along each curve of $S_1 \cap S'_2$. Hence, there is an isotopy that takes S to $S_1 + (S_2 - S'_2)$, thereby reducing its weight. This isotopy may move the curves $S \cap T$, but this is nevertheless a pattern isotopy because T is disjoint from P .

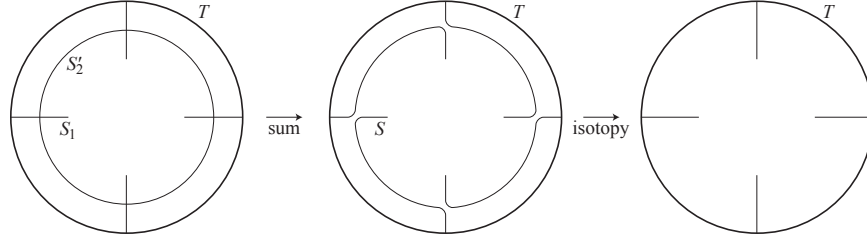


FIGURE 4. Isotopy of S when a summand has a boundary-parallel torus component S'_2

2.8. Components of least weight surfaces.

Theorem 2.5. *Let (M, P) be a compact orientable irreducible boundary-irreducible 3-manifold with boundary pattern. Let \mathcal{T} be a triangulation of M in which P is simplicial. Let S be an incompressible, boundary-incompressible, orientable normal surface properly embedded in (M, P) . Then S has least weight in its pattern-isotopy class if and only if each component of S does.*

Proof. Suppose that S is pattern-isotopic to a normal surface S' with smaller weight. Then some component of S' has smaller weight than the corresponding component of S . Hence, we deduce that this component of S did not have least weight.

Suppose now that S has least weight. Suppose that some component S_1 of S does not have least weight. Let S_2 be the remaining components of S . Let S'_1 and S'_2 be normal surfaces that have least weight among surfaces pattern-isotopic to S_1 and S_2 . Hence, $w(S'_1) < w(S_1)$ and $w(S'_2) \leq w(S_2)$. Now, by Proposition 3.7 in [13], we may form the normal sum $S' = S'_1 + S'_2$ and this surface S' is pattern-isotopic to S . But $w(S') = w(S'_1) + w(S'_2) < w(S_1) + w(S_2) = w(S)$, which contradicts the assumption that S had least weight.

Note that Proposition 3.7 in [13] was not stated for 3-manifolds with boundary patterns. However, the presence of the simplicial boundary pattern P does not affect the argument there and so Proposition 3.7 in [13] still applies. \square

2.9. Exponentially controlled and weakly fundamental surfaces. We will need to consider two types of normal surface that are less constrained than a fundamental surface. We term these weakly fundamental surfaces and exponentially controlled surfaces. Such surfaces will play an important role in this paper.

Let M be a compact orientable 3-manifold with a boundary pattern P . Let \mathcal{T} be a triangulation of M in which P is simplicial. We say that a normal surface S is *weakly fundamental* if

- (1) no component of S can be expressed as the sum of two non-empty normal surfaces S_1 and S_2 , where S_2 is disjoint from P and is a sphere, projective plane, disc, torus or annulus, and
- (2) no component of S can be expressed as a sum $S_1 + S_2 + S_3$ of non-empty normal surfaces where S_2 and S_3 are Möbius bands disjoint from P .

Proposition 2.6. *Let M be a compact orientable 3-manifold with a boundary pattern P . Let \mathcal{T} be a triangulation of M in which P is simplicial, and let t be the number of tetrahedra in \mathcal{T} . Then, for any weakly fundamental normal surface S*

with no 2-sphere components,

$$w(S) \leq t^2 2^{7t+6} (-\chi(S) + 2|S \cap P| + 2|S|).$$

Proof. Let $\{S_i\}$ be the set of fundamental normal surfaces. We write S as a sum of its components, and then we write each component as a sum of fundamental normal surfaces. We thereby obtain an expression

$$S = \sum_i n_i S_i$$

where each summand is identified with a summand of some component of S . By disregarding the surfaces S_i that do not appear in this sum, we may assume that $n_i > 0$ for all i . Weight, Euler characteristic and the number of intersections with P are all additive with respect to normal summation, and therefore

$$w(S) = \sum_i n_i w(S_i),$$

$$-\chi(S) = -\sum_i n_i \chi(S_i),$$

$$|S \cap P| = \sum_i n_i |S_i \cap P|.$$

Hence,

$$-\chi(S) + 2|S \cap P| = \sum_i n_i (-\chi(S_i) + 2|S_i \cap P|).$$

Since S is weakly fundamental, $-\chi(S_i) + 2|S_i \cap P| \geq 1$, for each i with at most 2 exceptions for each component of S . The exceptions occur when S_i is a torus, projective plane, disc or annulus disjoint from P (in which case at most one S_i can be of this form and this is the only summand for this component of S) or a Möbius band (in which case at most two S_i can be of this form for this component of S). We deduce that

$$\sum_i n_i \leq -\chi(S) + 2|S \cap P| + 2|S|.$$

Corollary 2.3 gives a bound on the weight of each fundamental normal surface. Hence,

$$w(S) \leq t^2 2^{7t+6} (-\chi(S) + 2|S \cap P| + 2|S|),$$

as required. \square

Let M be a compact orientable 3-manifold with a boundary pattern P . For constants c and k , we say that a surface S properly embedded in M is (c, k) -exponentially controlled if, for any triangulation \mathcal{T} of M in which P is simplicial and having t tetrahedra, S is strongly equivalent to a surface S' satisfying $w(S') \leq ck^t$. If a surface is (c, k) -exponentially controlled for some c and k , we say that it is *exponentially controlled*.

An immediate consequence of Proposition 2.6 is the following.

Corollary 2.7. *Let (M, P) be a compact orientable 3-manifold with boundary pattern. Let S be a compact surface properly embedded in M . Suppose that for any triangulation \mathcal{T} of M in which P is simplicial, S is strongly equivalent to a weakly fundamental normal surface. Then S is (c, k) -exponentially controlled, where $c = -\chi(S) + 2|S \cap P| + 2|S|$ and $k = 2^{15}$.*

2.10. Making surfaces weakly fundamental. Many topologically relevant surfaces may be placed in this form, as the following result demonstrates.

Theorem 2.8. *Let (M, P) be a compact orientable irreducible boundary-irreducible 3-manifold with boundary pattern. Let \mathcal{T} be a triangulation of M in which P is simplicial. Suppose that M contains no properly embedded essential tori or essential clean annuli. Let S be an incompressible, boundary-incompressible, orientable surface properly embedded in (M, P) . Suppose that each incompressible toral component of ∂M is disjoint from P . Suppose also that no component of S is a sphere, a clean disc, a clean annulus or a boundary-parallel disc intersecting P twice. Then S is pattern-isotopic to a weakly fundamental normal surface.*

Proof. By Proposition 2.1, S is pattern-isotopic to a normal surface. We may assume that S has smallest weight, up to pattern isotopy. By Theorem 2.5, each component of S has least weight in its pattern-isotopy class. We will show that each component S' of S is weakly fundamental, which will prove the theorem.

Suppose that S' is not weakly fundamental. Assume first that it can be expressed as $S_1 + S_2$, where S_2 is disjoint from P and is a sphere, projective plane, disc, torus or annulus. We may assume that this summation is in reduced form. However, by Theorem 2.4, S_2 cannot be a sphere, projective plane, clean disc, clean annulus or torus. For example, suppose that S_2 is a torus. By Theorem 2.4 (4), it is incompressible. By assumption, it must be parallel to a toral boundary component that is disjoint from P . But this is ruled out by Theorem 2.4 (8).

Suppose that S' can be expressed as $S_1 + S_2 + S_3$ where S_2 and S_3 are Möbius bands disjoint from P . By Theorem 2.4 (4), S_2 and S_3 are incompressible. The orientable double cover of each of the Möbius bands is a properly embedded clean incompressible annulus, which must therefore be parallel to an almost clean annulus in ∂M . Therefore, M is a solid torus and P is a (possibly empty) collection of essential non-meridional curves. We may give M a Seifert fibration in which P is a union of fibres. Now, $[S_2]$ and $[S_3]$ are both non-trivial elements of $H_2(M, \partial M; \mathbb{Z}/2\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z}$ and so their sum is homologically trivial. The sum $S_2 + S_3$ is incompressible and boundary-incompressible by Theorem 2.4 (4). It is also orientable, since it is homologically trivial. It is therefore pattern-isotopic to a surface that is vertical or horizontal, plus possibly some boundary-parallel components. If any component is vertical, then it is a clean annulus, which is inessential, contradicting Theorem 2.4 (6). If any component is boundary-parallel, it is a disc intersecting the pattern twice, since any other connected boundary-parallel surface is either a clean disc or boundary-compressible, contradicting Theorem 2.4 (5) or (4). If any component is horizontal, it is a meridian disc. Each component of $S_2 + S_3$ is therefore a disc. This is impossible since the Euler characteristic of $S_2 + S_3$ is zero. \square

2.11. JSJ surfaces. Let (M, P) be a compact orientable irreducible boundary-irreducible 3-manifold with boundary pattern.

Let F be an incompressible torus or an incompressible boundary-incompressible clean annulus properly embedded in (M, P) . Then F is *rough* if any incompressible torus and any clean incompressible boundary-incompressible annulus in M is pattern-isotopic to a surface disjoint from F . This is Definition 6.4.12 of [18].

A surface F properly embedded in M is *JSJ* if it is either a rough torus not parallel to a clean torus in ∂M , or a rough clean annulus not parallel to an almost clean annulus in ∂M . This is Definition 6.4.29 of [18].

It is possible to show that any collection of JSJ surfaces can be pattern-isotoped to be pairwise disjoint. Hence, the *JSJ surfaces* are obtained by taking one representative of each JSJ surface, up to pattern-isotopy, and arranging them to be disjoint from each other. The resulting surface is in fact well-defined up to pattern-isotopy, when (M, P) is irreducible and boundary-irreducible. (See Theorem 6.4.31 in [18].)

Let F be the JSJ surfaces for (M, P) . Give $M \setminus F$ the boundary pattern consisting of the union of ∂F and $(M \setminus F) \cap P$. Then each component of $M \setminus F$ with this boundary pattern is a *JSJ piece*. (Matveev [18] calls them *JSJ-chambers*.)

We now introduce two types of properly embedded clean essential annulus. A clean essential annulus A properly embedded in M is *longitudinal* if any other clean essential annulus A' can be pattern-isotoped so that $\partial A \cap \partial A' = \emptyset$. Otherwise A is *transverse*. This is Definition 6.4.13 in [18]. Note that any JSJ annulus is always longitudinal and never transverse.

We now recall Theorem 6.4.42, Lemma 6.4.32, Proposition 6.4.35 and Proposition 6.4.41 from [18], which combine to give the following result.

Theorem 2.9. *Let (M, P) be a compact orientable irreducible boundary-irreducible 3-manifold with boundary pattern. Then its JSJ pieces fall into the following types:*

- (1) *simple 3-manifolds with pattern;*
- (2) *Seifert fibred manifolds with pattern consisting of fibres;*
- (3) *I-bundles, with pattern disjoint from the horizontal boundary and intersecting each vertical boundary component in a non-empty collection of parallel copies of its core curve.*

Moreover, if a JSJ piece contains an essential torus or a clean essential longitudinal annulus, it is of the second type. If a JSJ piece contains a clean essential transverse annulus, it is of the third type.

Remark 2.10. Note that a JSJ annulus or torus can be pattern-isotoped off any properly embedded clean Möbius band or Klein bottle F , for the following reason. The orientable double cover \tilde{F} is a properly embedded clean annulus or torus. We claim that this is incompressible. Otherwise it would compress to two clean discs or to a 2-sphere. In the case of two discs, the boundary-irreducibility of (M, P) implies that the boundary of each of these discs would bound a clean disc in ∂M . One of these discs in ∂M can be extended so that its boundary is ∂F . But then the union of this disc and F is a projective plane. The irreducibility of M implies that it is a copy of \mathbb{RP}^3 , which contradicts the fact that M has non-empty boundary. When \tilde{F} compresses to a sphere, this bounds ball in M that does not contain the Klein bottle. Hence, again M is closed, which is a contradiction. This proves the claim. Therefore any JSJ annulus or torus can be pattern-isotoped off \tilde{F} . It is therefore disjoint from the regular neighbourhood $N(F)$ or lies within it. In the latter case, it can also be pattern-isotoped off F . For we may give $N(F)$ the structure of an I -bundle or circle bundle over a Möbius band in which F is a union of fibres. Then the JSJ torus or annulus can also be made a union of fibres and hence is an annulus or torus parallel to \tilde{F} .

2.12. Connected JSJ surfaces are fundamental. Many topologically relevant surfaces can be made fundamental. In particular, this is true of the JSJ annuli and tori, as follows.

Theorem 2.11. *Let (M, P) be a compact orientable irreducible boundary-irreducible 3-manifold with boundary pattern. Let \mathcal{T} be a triangulation of M in which P is simplicial. Then each connected JSJ surface that has least weight in its pattern-isotopy class is fundamental.*

Our proof will use the following notion. We say that two surfaces S_1 and S_2 properly embedded in M are *partially coincident* if $S_1 \cap S_2$ is a compact subsurface of S_1 and S_2 . We say that they have *essential intersection* if no component of $S_1 \cap S_2$ or $(S_1 \cup S_2) \setminus (S_1 \cap S_2)$ is a disc disjoint from ∂M or a disc intersecting ∂M in a single arc.

When at least one of S_1 and S_2 is a torus and they have essential intersection, then $S_1 \cap S_2$ consists of annuli or possibly a torus. When S_1 and S_2 are both clean annuli and they have essential intersection, then either every component of $S_1 \cap S_2$ is an annulus parallel to a core curve of S_1 and S_2 or every component of $S_1 \cap S_2$ is a *square*, which is a disc consisting of an arc in ∂S_i , an essential arc properly embedded in S_i , another arc in ∂S_i and another essential arc properly embedded in S_i .

A *generalised product region* is a connected 3-manifold of the form $(F \times [-1, 1]) / \sim$, where \sim is an equivalence relation with each equivalence classes equal either to a singleton or to an interval of the form $x \times [-1, 1]$ for points $x \in J$, where J is a finite union of closed intervals and circles in ∂F . A *horizontal boundary component* is the image of $F \times \{1\}$ or $F \times \{-1\}$. A *vertical boundary component* is the closure of the image of a component of $(\partial F - J) \times [-1, 1]$. The *pinching locus* is the image of $J \times [-1, 1]$.

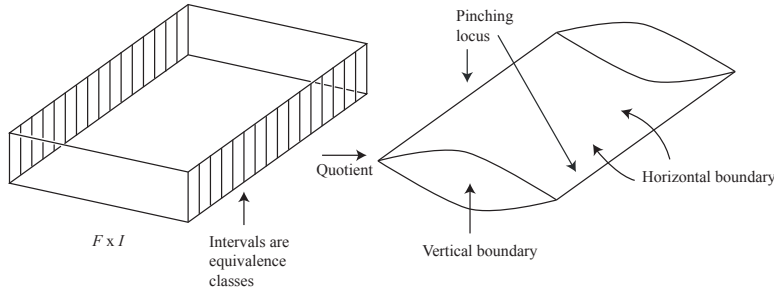


FIGURE 5. A generalised product region is obtained from a product by collapsing certain arcs in the boundary

Lemma 2.12. *Let S_1 and S_2 be surfaces properly embedded in a compact orientable 3-manifold M with boundary pattern P , where each S_i is an incompressible boundary-incompressible torus or clean annulus. Suppose that they are partially coincident but not equal, that they have essential intersection, and that there is a pattern-isotopy taking S_1 off S_2 . Then there is a component X of $M \setminus (S_1 \cup S_2)$ that is a generalised product region X with the following properties:*

- (1) each horizontal boundary component lies in S_1 or in S_2 ;

- (2) each vertical boundary component lies in ∂M and is disjoint from P ;
- (3) the pinching locus is a union of components of $S_1 \cap S_2 \cap X$.

It is possible that other components of $S_1 \cap S_2$ lie in the horizontal boundary. It is also possible that the inclusion of the product region into M is not injective, since components of $S_1 \cap S_2$ in the boundary of the product region may be identified.

Proof. Note that each component F of $S_1 \cap S_2$ is a square or annulus. We say that F is *removable* if the two components of $S_1 - S_2$ incident to ∂F emanate from the same side of S_2 (or equivalently, the two components of $S_2 - S_1$ incident to ∂F emanate from the same side of S_1). Otherwise F is *non-removable*. (See Figure 6.)

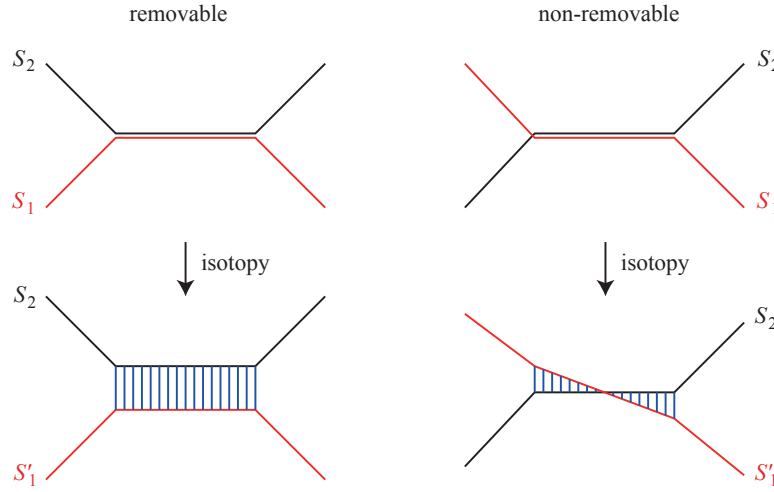


FIGURE 6. Isotoping partially coincident surfaces S_1 and S_2 so that they intersect transversely

We now perform a small isotopy to S_1 so that the resulting surface S'_1 intersects S_2 transversely, where this intersection consists of a single arc or curve in each non-removable component of $S_1 \cap S_2$. Note that we may obtain the original arrangement S_1 and S_2 by collapsing some product regions as follows. For each removable component F of $S_1 \cap S_2$, we have a copy of $F \times [-1, 1]$ in M , with $(F \times [-1, 1]) \cap (S'_1 \cup S_2) = F \times \{-1, 1\}$. If we collapse each interval $\{*\} \times [-1, 1]$ to a point, we obtain the original arrangement of S_1 and S_2 . Similarly, for each non-removable component F of $S_1 \cap S_2$, we have a copy of $(F - (S'_1 \cap S_2)) \times [-1, 1]$ embedded in M and when the copies of $[-1, 1]$ are collapsed to points, we again revert to the original arrangement of S_1 and S_2 .

We now apply a theorem of Waldhausen [22] (in the case where the surface are closed) or Johansson [14] (in the case where the surfaces have boundary). This states that there is a generalised product region arising as a component of $M \setminus (S'_1 \cup S_2)$ where one horizontal boundary component lies in S'_1 , the other horizontal boundary component lies in S_2 , any vertical boundary lies in ∂M and is disjoint from P , and the binding locus consists of components of $S'_1 \cap S_2$.

We now perform the collapses described above. The generalised product region collapses to a collection of generalised product regions for S_1 and S_2 , with components of $S_1 \cap S_2$ attached. Any one of these new generalised product regions

gives the desired component of $M \setminus (S_1 \cup S_2)$. Note that the collapses may create self-identifications on the boundary of the generalised product region. \square

Lemma 2.13. *Let (M, P) be a compact orientable irreducible boundary-irreducible 3-manifold with boundary pattern. Let \mathcal{T} be a triangulation of M in which P is simplicial. Let F be a clean properly embedded surface, with least weight in its pattern-isotopy class, and that is incompressible and boundary-incompressible. Suppose that $F = F_1 + F_2$ is in reduced form.*

- (1) *If A is a trace annulus or disc, then A is incompressible and boundary-incompressible in the manifold $M \setminus F$ with boundary pattern consisting of P and the copies of ∂F .*
- (2) *Suppose that A is a trace annulus. Suppose also that for any other summation $F = F'_1 + F'_2$, then $|F'_1 \cap F'_2| \geq |F_1 \cap F_2|$. Then A is essential in $M \setminus F$.*
- (3) *If F is a JSJ annulus, and A is a trace disc, then the JSJ piece of M containing A is an I -bundle chamber, in which F is a vertical boundary-component.*

Proof. (1) Each trace annulus and disc is incompressible by Theorem 2.4. Suppose that A is boundary-compressible and hence parallel to an annulus or disc in F . If we were to perform an irregular switch along A but regular switches along all other trace discs, annuli and Möbius bands, the result would be a copy of F as well as a compressible annulus or torus. After performing an irregular switch, there is always an isotopy that strictly reduces the weight of the surface. Hence, we deduce that F did not have least weight in its pattern-isotopy class, which is contrary to hypothesis.

(2) Note first that F is connected, because otherwise there is a summation $F = F'_1 + F'_2$ with $F'_1 \cap F'_2 = \emptyset$, contradicting our minimality assumption. Suppose that A is an inessential trace annulus and hence, by (1), parallel to an almost clean annulus in $M \setminus F$. Thus, A separates off a solid torus V in $M \setminus F$ in which the pattern is a non-empty collection of disjoint longitudes. There may be other trace annuli in V , but they are all incompressible and boundary-incompressible by (1). Hence, there is a trace annulus in V that is outermost in V . By considering this trace annulus instead, we may assume that the interior of V contains no trace annuli. Let F'_1 be $(F_1 \setminus V) \cup (F_2 \cap V)$ and define F'_2 similarly. Then $F = F'_1 + F'_2$ and $|F'_1 \cap F'_2| < |F_1 \cap F_2|$, contrary to assumption.

(3) Since A is a trace disc, F is decomposed by $F \cap A$ into two squares. Each of these squares has a product structure as $I \times I$ where $(I \times I) \cap \partial M = I \times \partial I$. Similarly, A has such a structure. So a regular neighbourhood of $F \cup A$ is an I -bundle over a surface S . Note that A is not attached to both sides of F , since this would imply the existence of an essential clean annulus that could not be isotoped off F , contradicting the assumption that F is a JSJ annulus. Hence, S is a pair of pants or a once-punctured Möbius band. Suppose first that a vertical boundary component of this I -bundle is compressible. Then it bounds a copy of $D^2 \times I$ with interior disjoint from the bundle, and we can extend the bundle over this 3-ball. In the case where the base surface S was a once-punctured Möbius band, this implies that the component of $M \setminus F$ containing A was an I -bundle chamber, as required. So suppose that the base surface was a pair of pants. There may have been other trace discs in $D^2 \times I$, but there is one that is outermost, and we hence may assume that

A is outermost in $D^2 \times I$. But then the same patch is incident to both components of $A \cap F$, which is impossible, since one lies in F_1 and one lies in F_2 . Thus, we have reduced to the case where the vertical boundary components of the I -bundle are incompressible. If any is inessential, then extend the I -bundle over the incident solid torus. We end with an I -bundle having essential vertical boundary. This must be part of an I -bundle piece in M . Furthermore, F is a vertical boundary component of this piece, as claimed. \square

Proof of Theorem 2.11. Let T be a connected JSJ surface. Pick a least weight normal representative for T , up to pattern-isotopy. If T is not fundamental, then we may write $T = T_1 + T_2$ where T_1 is connected. We may assume that $|T_1 \cap T_2|$ is minimal among all such summations, and in particular that the sum is in reduced form. By Theorem 2.4, no T_i has positive Euler characteristic. Hence, $\chi(T_1) = \chi(T_2) = 0$. Furthermore, T_1 and T_2 are incompressible and boundary-incompressible. Hence, T_1 is a clean annulus, torus, Klein bottle or Möbius band. Therefore, since T is rough, there is a pattern-isotopy taking T_1 off T . In the case where T_1 is a Möbius band or Klein bottle, this is the content of Remark 2.10.

The summation $T = T_1 + T_2$ places T and T_1 into partially coincident position. Specifically, each patch in T_1 becomes a component of $T \cap T_1$. Each trace annulus, disc or Möbius band becomes a component of $T_1 \setminus T$. Each patch in T_2 becomes a component of $T \setminus T_1$. Note that for each trace annulus or disc in $T_1 \setminus T$, the incident components of $T \setminus T_1$ emanate from opposite sides of the trace disc or annulus. This is a consequence of normal summation.

We claim that the trace annuli and discs are incident to just one side of T . Any trace disc or annulus is incompressible and boundary-incompressible in $M \setminus T$ by Lemma 2.13 (1). If there is a trace annulus A , then this is essential in $M \setminus T$ by Lemma 2.13 (2). So, by Theorem 2.9, the JSJ piece of M containing A is either Seifert fibred or an I -bundle chamber, but in the latter case, the horizontal boundary is one or two copies of T . Thus, in both cases, the JSJ piece of M containing A is Seifert fibred, and A is a union of fibres. If there is a trace disc on one side of T , then Lemma 2.13 (3) implies that this lies in an I -bundle chamber of $M \setminus T$, and this must be a solid torus. Again the JSJ piece of M containing A is Seifert fibred. Note that there cannot be both trace annuli and trace discs, since trace discs would cut the annulus T into discs. Hence if there are trace discs or annuli on both sides of T , then we deduce that on both sides of T there are Seifert fibre spaces. Moreover, as trace annuli are pairwise disjoint, we deduce that these Seifert fibre spaces have isotopic regular fibres. In all case, we deduce that T is not JSJ, which is a contradiction, proving the claim.

Since there is an admissible isotopy taking T off T_1 , Lemma 2.12 gives a generalised product region V . Since its boundary is orientable, it can contain no trace Möbius band components. Consider first the horizontal boundary component F_1 of V lying in T_1 . We claim that this contains exactly one removable component of $T \cap T_1$. If there were more than one such component, then between two of them would be a component of $T_1 \setminus T$, which would be a trace annulus or disc, with the two incident components of $T \setminus T_1$ emanating from the same side, which is impossible. If there was no removable component of $T \cap T_1$ in F_1 , then F_1 would be an entire trace disc or annulus, but then again the two incident components of $T \setminus T_1$, which lie in ∂V , would emanate from the same side of F_1 .

Now consider the horizontal boundary component F of V lying in T . We claim that this contains no removable components of $T \cap T_1$. For if there were such a component, then each component of $T \setminus T_1$ incident to ∂F would have components of $T_1 \setminus T$ emanating from both sides. Hence, there would be trace annuli or discs incident to both sides of T . But we have shown above that this does not arise.

Thus, we have shown that the product region V must have exactly two components of $T_1 \setminus T$ in its boundary, in other words exactly two trace annuli or two trace discs.

Note that the map $V \rightarrow M$ is an injection. This is because the only way that it can fail to be an injection is when the two horizontal boundary components of V contain components of $T_1 \cap T$ that are identified. But we have shown that F contains no such components. So the two horizontal boundary components of V are disjoint apart from along their boundaries. Consider the surfaces T'_1 and T'_2 that are obtained from T_1 and T_2 by swapping the two horizontal boundary components of V . In other words, $T'_1 = (T_1 \setminus F_1) \cup F$ and T'_2 is obtained from T_2 by removing $F \cap T_2$ and attaching the remainder of ∂V . Then T'_1 and T'_2 are normal, and $T = T'_1 + T'_2$, but $|T'_1 \cap T'_2| < |T_1 \cap T_2|$. This contradicts the assumption that the summation $T_1 + T_2$ is in reduced form. \square

Theorem 2.14. *Let (M, P) be a compact orientable irreducible boundary-irreducible 3-manifold with boundary pattern. Let \mathcal{T} be a triangulation of M in which P is simplicial. Then the JSJ surfaces may be realised as a disjoint union of fundamental surfaces.*

Proof. Pick a normal representative for the JSJ surfaces that has least weight in its pattern-isotopy class. By Theorem 2.5, each component has least weight in its pattern-isotopy class. Hence by Theorem 2.11, each component is fundamental. \square

2.13. Normal annuli in Seifert fibre spaces. We will be considering the JSJ tori for the exterior of a non-split link in the 3-sphere. It is possible that certain complementary regions of these tori may be Seifert fibre spaces. These require techniques that are rather different from other types of 3-manifolds. So, in this section, we recall and develop the necessary machinery.

We start by recalling what Seifert fibre spaces can arise. By [4, Proposition 3.2], such Seifert fibre spaces fall into one of the following types:

- (1) a solid torus, which has base space a disc with at most one cone point;
- (2) the product of a planar surface with the circle; this has no exceptional fibres; it arises, for example, as the exterior of the link shown in Figure 7 (i);
- (3) the exterior of a torus link, with zero, one or both of the core curves of the associated solid tori removed, as shown in Figure 7 (ii); its base space is punctured 2-sphere with at most two cone points; moreover, if the space has two exceptional fibres, these have coprime orders.

We will be concerned with annuli properly embedded in such a Seifert fibre space.

Lemma 2.15. *Let M be a Seifert fibre space with non-empty boundary, that embeds in the 3-sphere, other than a solid torus or $S^1 \times S^1 \times I$.*

- (1) *If ∂M is a single torus, then there is an essential annulus A properly embedded in M , such that the exterior of A is two solid tori.*

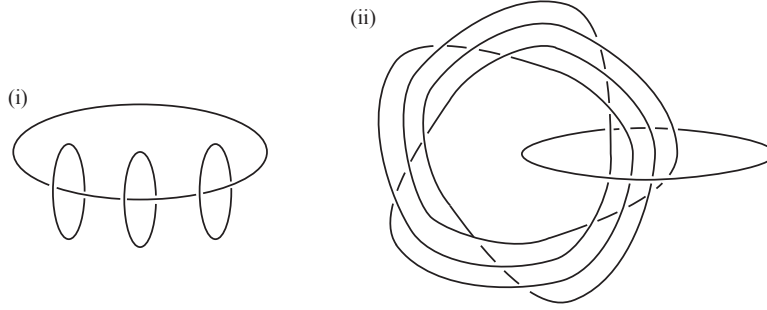
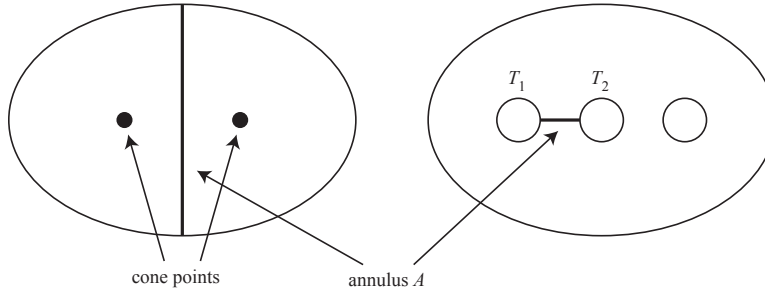


FIGURE 7. Some links with Seifert fibred exteriors

- (2) If ∂M is more than one torus, let T_1 and T_2 be distinct boundary components. Then there is an essential annulus A properly embedded in M , with one boundary component on T_1 and one boundary component on T_2 .

In both cases, A is unique, up to a homeomorphism of M that is isotopic to the identity on ∂M .

Proof. We start by describing A , and then go on to explain why it is unique. When ∂M is a single torus, the base space of the Seifert fibre space is a disc with two cone points. We pick a properly embedded arc in the disc that separates the two cone points. When ∂M is more than one torus, we pick a properly embedded arc in the base space joining the boundary components corresponding to T_1 and T_2 . In each case, the inverse image in M of this arc is the required annulus A .

FIGURE 8. Some possible base spaces of M containing the projection of A

We now show that A is unique, up to a homeomorphism of M that is isotopic to the identity on ∂M . The hypothesis that M embeds in the 3-sphere, has non-empty boundary, and is not a solid torus or $S^1 \times S^1 \times I$ implies that its Seifert fibering is unique up to isotopy [20, 21]. Any properly embedded orientable essential surface in a Seifert fibre space is isotopic to a surface that is horizontal or vertical. If an annulus is horizontal, then the base orbifold must have zero Euler characteristic, but this is ruled out by our hypothesis that M embeds in the 3-sphere and is not $S^1 \times S^1 \times I$. So, we may assume that A is vertical. It therefore projects to an arc in the base orbifold, which avoid the cone points. It is easy to see that there is a unique such arc, up to orientation-preserving homeomorphism of the base space

that preserves each boundary component and each cone point. This implies that A is unique, up to a homeomorphism of M that is isotopic to the identity on ∂M . \square

Proposition 2.16. *Let M and A be as in Lemma 2.15. Let P be a boundary pattern, such that for each component T of ∂M , one of the following holds:*

- (1) T is disjoint from P ;
- (2) $T \cap P$ is a disjoint union of essential simple closed curves;
- (3) $T \setminus P$ is discs.

Suppose that among the annuli in Lemma 2.15, A intersects P as few times as possible. When A is disjoint from P and ∂M is connected, we also require that both components of ∂A lie in the same component of $\partial M \setminus P$. Pick a triangulation of M in which P is simplicial. Then A is strongly equivalent to a weakly fundamental surface.

Proof. Since A is essential, it can be isotoped to a normal surface. We pick a representative for A that has minimal weight, up to strong equivalence. Suppose that A is not weakly fundamental.

Suppose first that A can be expressed as a sum $S_1 + S_2$ of normal surfaces, where S_2 is a sphere, projective plane, torus, clean disc or clean annulus. We may assume that the sum is in reduced form. By Theorem 2.4, these surfaces are both incompressible, and neither has a sphere, projective plane, clean disc or clean inessential annulus component. Hence, $\chi(S_1) = \chi(S_2) = 0$.

Consider first the case when S_2 , say, is a torus. Then, $\partial A = \partial S_1$. So either some component of S_1 is an annulus or two components of S_1 are Möbius bands. But the latter case cannot arise, because each Möbius band would have to be vertical and so would contain a singular fibre of order 2, whereas M contains at most two singular fibres and these have coprime orders. So, S_1 has an essential annulus component. It cannot be boundary-parallel by our assumption on P . Since it has the same boundary as A , it is strongly equivalent to A . But S_1 has smaller weight than A , which is a contradiction. So we may assume that A has no torus summand.

Suppose now that S_2 is a clean essential annulus and that S_1 has an annulus component. Consider first the case where M has base space a disc and two exceptional fibres. Then, by our minimality assumption about $|A \cap P|$, A must be disjoint from P . So, by assumption, both components of ∂A lie in the same component of $\partial M \setminus P$. Hence, ∂S_1 and ∂S_2 also lie in this component. Hence, S_2 is strongly equivalent to A . But it has smaller weight than A , which is a contradiction. So suppose that M has more than one boundary component. Then by assumption, the components of ∂A lie on distinct tori T_1 and T_2 . Now $\partial S_1 \cup \partial S_2$ lies on the same boundary components as ∂A . Since S_2 is clean, the torus or tori T containing ∂S_2 have a clean fibre. By our minimality assumption on $|A \cap P|$, $A \cap T$ is disjoint from P . Hence, $A \cap T$ lies in the same component of $T \setminus P$ as $S_1 \cap T$ and $S_2 \cap T$. If S_2 has one boundary component on T_1 and one boundary component on T_2 , then we may replace A by this surface. So suppose that both boundary components of S_2 lie on some boundary component T of ∂M , then $\partial S_1 \cap (\partial M - T) = \partial A \cap (\partial M - T)$, which is single curve. This curve cannot be the boundary of a Möbius band component of S_1 , since then every other component of S_1 is a torus or an annulus with both boundary curves on T , which implies that $[A \cap T] = [S_1 \cap T] + [S_2 \cap T] = 0 \in H_1(T; \mathbb{Z}_2)$, but this is not the case. So we may

replace A by the annular component of S_1 containing $\partial A \cap (\partial M - T)$ and reduce its weight, which is a contradiction.

Suppose now that S_2 is a clean essential annulus and that S_1 has no annulus or torus components. Then S_1 is a collection of Möbius bands. But M cannot support two or more disjoint Möbius bands. So S_1 is a single Möbius band. If M has a single boundary component, then, as above, S_2 is strongly equivalent to A and has smaller weight, which is a contradiction. So, suppose that M has more than one boundary component, and that one component of ∂A lies on T_1 and the other lies on T_2 . Then ∂S_1 is a single regular fibre on some component (T_1 , say) of ∂M . The curves ∂S_2 are isotopic to regular fibres on ∂M . If both curves of ∂S_2 lie on T_2 , then $[A \cap T_2] = [S_2 \cap T_2] = 0 \in H_1(\partial M; \mathbb{Z}_2)$, which is not the case. If one curve of ∂S_2 lies on T_1 and the other lies on T_2 , then $[A \cap T_1] = [\partial S_1 \cap T_1] + [\partial S_2 \cap T_1] = 0 \in H_1(\partial M; \mathbb{Z}_2)$, which again is not the case. Finally if both curves of ∂S_2 lie on T_1 , then A is disjoint from T_2 , which is not the case. In all cases, we obtain a contradiction.

Suppose now that $A = S_1 + S_2 + S_3$ where S_2 and S_3 are clean Möbius bands. Suppose also that M has a single boundary component. Then S_2 (say) is a vertical surface that projects to an arc in the base space running from the boundary of the base space to a singular point of order 2. So $2S_2$ is a clean essential annulus. By assumption, A is therefore clean and both its boundary components lie in the same component of $\partial M \setminus P$. Hence, the annuli $2S_2$ and $2S_3$ also have boundary lying in this component, and so are strongly equivalent to A . Now $w(S_2 + S_3)$ is the mean of $w(2S_2)$ and $w(2S_3)$. So one of $w(2S_2)$ and $w(2S_3)$ is at most $w(S_2 + S_3)$, the former say. But then $w(2S_2) \leq w(S_2 + S_3) < w(A)$, contradicting the assumption that A has least weight.

Suppose now that $A = S_1 + S_2 + S_3$ where S_2 and S_3 are clean Möbius bands, and that M has at least two boundary components. Now $[\partial S_2]$ represents the homology class of a fibre of T_1 or T_2 . The same is true of $[\partial S_3]$. Hence, $S_2 + S_3$ cannot be a single Möbius band and some tori. It also cannot have more than one Möbius band component, since M does not support two disjoint Möbius bands. So, $S_2 + S_3$ contains a clean annulus component. We may therefore write $A = S'_1 + S'_2$ where S'_2 is a clean annulus, and we have dealt with that case above. \square

Proposition 2.17. *Let M be homeomorphic to $T^2 \times [0, 1]$, with boundary pattern P so that (M, P) is boundary-irreducible. Suppose that M contains no clean essential annulus. Let \mathcal{T} be a triangulation of M in which P is simplicial. Then any incompressible, boundary-incompressible annulus A properly embedded in M , in general position with respect to P and joining the two components of ∂M , is strongly equivalent to a weakly fundamental normal surface.*

Proof. By Proposition 2.1, A is pattern-isotopic to a normal surface. We may assume that A has smallest weight, up to strong equivalence. By Theorem 2.4 (5), no summand of A is a sphere, clean disc or projective plane. By Theorem 2.4 (4) and (7), no summand is a torus. By Theorem 2.4 (4) and (6), no summand is a clean inessential annulus. We are assuming that M contains no clean essential annulus. The manifold $T^2 \times I$ does not contain a properly embedded Möbius band. Hence, A is weakly fundamental. \square

Proposition 2.18. *Let M be homeomorphic to $T^2 \times [0, 1]$, with non-empty boundary pattern P so that (M, P) is boundary-irreducible. Let \mathcal{T} be a triangulation of M in which P is simplicial. Let A be any essential annulus disjoint from the pattern, joining the two boundary components. Then A is strongly equivalent to a fundamental surface.*

Proof. We may assume that A has least weight, up to strong equivalence. Suppose that A is not fundamental, and hence is a normal sum $S_1 + S_2$. We may assume that this summation is in reduced form. Hence, by Theorem 2.4, S_1 are incompressible and boundary-incompressible and have zero Euler characteristic. Now $T^2 \times [0, 1]$ supports no properly embedded Möbius band. Therefore, S_1 and S_2 are unions of clean annuli or tori. For at least one i , $[S_i \cap \partial M]$ is non-trivial in $H_1(\partial M; \mathbb{Z}_2)$. Say that this is true of S_1 . Then S_1 contains a clean annulus joining the two components of ∂M . The components of $\partial M \setminus P$ that it intersects must be the same components that A intersects. These components of $\partial M \setminus P$ are annuli. So, there is a pattern isotopy taking ∂S_1 to ∂A . There is then a homeomorphism equal to the identity on ∂M taking S_1 to A . This contradicts the assumption that A had least weight. \square

2.14. Boundary parallel surfaces.

Lemma 2.19. *Let (M, P) be a compact orientable irreducible boundary-irreducible 3-manifold with boundary pattern. Let S' be a union of edges of P , circles of P , and components of $\partial M \setminus P$ that are not necessarily disjoint in ∂M , and let S be an incompressible boundary-incompressible surface that is obtained from S' by taking the frontier of a regular neighbourhood and that contains no disc component intersecting P zero or two times. Then S is $(6, 2)$ -exponentially controlled.*

Proof. Let \mathcal{T} be any triangulation of M in which P is simplicial. Then a representative for S intersects each edge of \mathcal{T} at most twice. There are at most $6t$ edges, where t is the number of tetrahedra of \mathcal{T} . Hence, the weight of S is at most $12t$. It need not be normal, but it can be isotoped to a normal surface without increasing its weight, by Proposition 2.1. Therefore, its weight is at most $12t \leq 6 \cdot 2^t$. Hence, it is $(6, 2)$ -exponentially controlled. \square

2.15. Sections of circle bundles. In this subsection, we consider the case where M is a circle bundle over a surface with non-empty boundary. Our goal will be to show that, for any triangulation of M , there is a surface isotopic to a section of the bundle having controlled weight. However, we need somewhat more than this. The bundle M will be decomposed into solid tori, each of which is a union of circle fibres. Moreover, the intersection of any two such solid tori will consist of a (possibly empty) union of annuli, which again is a union of fibres. We want to ensure that the surface intersects each solid torus in a collection of meridian discs. In order to ensure this, we will need to generalise some of the results earlier in this section.

The first thing that we do is to slightly relax the notion of an exponentially controlled surface. Given a compact orientable 3-manifold M with boundary pattern P , we say that a surface S properly embedded in M is (c, k) -nearly exponentially controlled if, given any triangulation of M with t tetrahedra in which P is simplicial, there is an orientation-preserving homeomorphism of M that preserves each edge, vertex and circle of P , and that takes S to a surface with weight at most ck^t .

The next thing that we need to do is to consider a more general notion of weight for a surface. Suppose that E_1 and E_2 are sets of edges of the triangulation, where $E_1 \subset E_2$. The *prioritised weight* of a properly embedded surface S is the triple

$$(|S \cap E_1|, |S \cap E_2|, w(S)).$$

Two such triples are compared using lexicographical ordering.

In our setting, each solid torus will be simplicial in the triangulation, and E_2 will be the edges in the boundary of these solid tori, and E_1 will be the set of edges that lie either in three such solid tori or in two such solid tori and ∂M .

We note that there is the following version of Theorem 2.4 that refers to prioritised weight.

Theorem 2.20. *Let M , P and \mathcal{T} be as in Theorem 2.4. Let $E_1 \subset E_2$ be collections of edges of the triangulation. Let S be a properly embedded, connected, incompressible, boundary-incompressible, normal surface that has least possible prioritised weight, among all normal surfaces that are pattern-isotopic to it. Suppose that S can be expressed as a normal sum $S_1 + S_2$, where S_1 and S_2 are in general position and non-empty, and the summation is in reduced form. Then the conclusions of Theorem 2.4 all hold.*

The results from [18] used in the proof of Theorem 2.4 do not refer to the prioritised weight, but rather the weight of S . However, we observe that all the arguments used in these proofs in [18] work just as well for prioritised weight, and hence Theorem 2.20 follows.

Lemma 2.21. *Let M be a compact orientable 3-manifold that is a circle-bundle over a surface. Let P be a boundary pattern for M . Let C be a disjoint union of fibres in M , each of which lies in the interior of M or in P , and where each component of ∂M contains at least one component of C . Let A be a collection of annuli in M , each of which is embedded, is a union of fibres and has boundary in C . Suppose that the interior of these annuli are disjoint, and that for each component of C , there are three sheets of A and ∂M coming into it. Suppose that M is triangulated, with P , C and A all simplicial. Let E_1 be the set of edges in C , and let E_2 be the set of edges in $A \cup \partial M$. Let S be a surface properly embedded in M , and suppose that there is a homeomorphism of M that preserves A and that takes S to a section. Assume that S has minimal number of intersections with P among all such surfaces. Then there is an isotopy preserving $A \cup P$ taking S to a normal surface S' with prioritised weight at most that of S .*

Proof. Consider a surface S' that is isotopic to S , via an isotopy preserving $A \cup P$, and that has least prioritised weight among all such surfaces. Note that S' intersects each component of C at least once, since it is isotopic to a section, and hence S' intersects each component of C exactly once, because its prioritised weight has been minimised. Furthermore, for each annulus of A , S' intersects that annulus in a single essential arc.

Suppose that S' is not normal. Then we may apply one of the normalisation moves in the proof of Corollary 3.3.25 in [18]. We claim that each such move gives an isotopy supported away from $A \cup \partial M$. For example, suppose that there is a simple closed curve of intersection between S' and some face of the triangulation. Since S' intersects each annulus of A in a single essential arc, we deduce that the face does not lie in A . Similarly, it does not lie in ∂M . Consider such a simple

closed curve that is innermost in the face. It bounds a disc D in the face. By the incompressibility of S' , ∂D bounds a disc D' in S' , and by the irreducibility of M , $D \cup D'$ bounds a ball. We may isotope D' across this ball and then a bit beyond the face. This procedure removes intersections between D' and the 1-skeleton of the triangulation. However, it has no effect on $A \cap S'$, since if any component of $A \cap S'$ were removed, the resulting surface would be disjoint from some fibre, which is impossible.

Other moves in the proof of Corollary 3.3.25 in [18] remove two points of intersection between some edge and S' . We claim that such an edge cannot lie in $A \cup \partial M$. Certainly, the edge cannot lie in C , since each component of C intersects S' just once. Also, the edge cannot lie in P by our assumption that the number intersection points with P is minimal. On the other hand, if the edge lies in some annular component A' of $(A \cup \partial M) \setminus C$ and is disjoint from P , then the isotopy replaces the essential arc $A' \cap S'$ with a pattern-isotopic arc with smaller weight plus possibly a simple closed curve. This contradicts the assumption that the prioritised weight of S' is minimal.

In all cases, we deduce that the isotopy is supported away from $A \cup \partial M$. The resulting surface is normal. \square

Lemma 2.22. *Let M , P , C , A , E_1 and E_2 be as in Lemma 2.21. Suppose that $M \setminus A$ is a union of solid tori. Let \mathcal{T} be a triangulation of M in which C , P and A are simplicial. Let S be a surface that, after a homeomorphism preserving A , is equal to a section. Let S' be a surface that is the image of S under a homeomorphism of M that preserves each vertex, edge and circle of P , and which has no greater prioritised weight. Then there is a homeomorphism of M preserving each vertex, edge and circle of P and each component of A taking S to S' .*

Proof. The homeomorphism of M taking S to S' preserves each of the fibres in $C \cap \partial M$ and so is isotopic to a fibre-preserving homeomorphism. The surface S intersects each component of C exactly once. Since S' differs from S by a fibre-preserving homeomorphism and an isotopy, it also intersects each component of C an odd number of times. Because the prioritised weight of S' is at most that of C , we deduce that S' intersects each component of C exactly once. Hence, S' intersects each annulus in A and $\partial M \setminus P$ in exactly one essential arc plus possibly some inessential simple closed curves. However, S' cannot intersect any component of $\partial M \setminus P$ in an inessential simple closed curve, since S' is isotopic to a section. Moreover, if S' had an inessential simple closed curve of intersection with some component of A , then we could perform an isotopy to S' supported in the interior of M to reduce its prioritised weight. So, $S' \cap (A \cup \partial M)$ is homeomorphic to $S \cap (A \cup \partial M)$. Furthermore, their regular neighbourhoods in S' and S respectively are homeomorphic. The surfaces S' and S are obtained from these regular neighbourhoods by attaching $S' \setminus N(A \cup \partial M)$ and $S \setminus N(A \cup \partial M)$ respectively. The latter is a disjoint union of discs. Hence, as S' and S have the same Euler characteristic, $S' \setminus N(A \cup \partial M)$ is also discs. Each such disc is a meridian disc for a solid torus component of $M \setminus A$. So, there is a homeomorphism of M preserving each vertex, edge and circle of P and each component of A taking S' to a section. Therefore, there is also such a homeomorphism taking S to S' . \square

Proposition 2.23. *Let M , P , C , A , E_1 and E_2 be as in Lemma 2.21. Suppose that $M \setminus A$ is a union of solid tori. Let \mathcal{T} be a triangulation of M in which C , P and A*

are simplicial. Let S be a properly embedded surface that is isotopic, via an isotopy preserving A , to a section. Suppose that, among such surfaces, S intersects P as few times as possible. Then there is a homeomorphism of M preserving each vertex, edge and circle of P and each component of A taking S to a weakly fundamental normal surface.

Proof. By Lemma 2.21, there is an isotopy preserving $A \cup P$ taking S to a normal surface without increasing its prioritised weight. We now fix S to be such a surface that has least prioritised weight, up to homeomorphism of M preserving each vertex, edge and circle of P and each component of A . Then by Lemma 2.22, S has least prioritised weight, up to a homeomorphism of M that preserves each vertex, edge and circle of P and each component of A . Hence, the same is true of the surface $2S$, which is two parallel copies of S .

Suppose that S is not weakly fundamental. Suppose first that $S = S_1 + S_2$ where S_1 is non-empty and S_2 is a sphere, projective plane, torus, clean disc or clean annulus. The surface $2S$ is a sum $S'_1 + S'_2$ where each S'_i is the (possibly disconnected) orientable double cover of S_i . We may assume that the summation is in reduced form. By Theorem 2.20, S'_1 and S'_2 are both incompressible and boundary-incompressible, have no component that is a 2-sphere, a clean disc or a clean annulus parallel to an almost clean annulus in ∂M . Hence, S'_1 and S'_2 must be isotopic to surfaces that are horizontal or vertical. Suppose that neither S'_1 nor S'_2 has any horizontal components. Then each component is vertical. Hence, the same is true of each component of S_1 and S_2 . But then S_1 and S_2 would intersect each fibre an even number of times, whereas S does not, which is a contradiction. So, one of S'_1 and S'_2 has a horizontal component, and so the same is true of a component of S_1 or S_2 . This is isotopic to a section, since it intersects each fibre in C at most once. Now, S_2 is disjoint from P and so cannot be a section. Therefore, S_1 is isotopic to a section. Moreover, its boundary is equal to ∂S away from annular components of $\partial M \setminus P$. So there is a homeomorphism of M preserving each vertex, edge and circle of P , taking S_1 to S . It has smaller prioritised weight than S . So, by Lemma 2.22, there is a homeomorphism of M preserving each vertex, edge and circle of P and each component of A , taking S_1 to S . But this contradicts our assumption that S had smallest prioritised weight among such surfaces.

Suppose now that $S = S_1 + S_2 + S_3$, where S_1 is non-empty and S_2 and S_3 are clean Möbius bands. The only orientable circle bundles over a compact surface that have a section and that contain a properly embedded Möbius band are the solid torus and the orientable circle bundle over the Möbius band. Suppose first M is a solid torus. In that case, both S_2 and S_3 represent the non-trivial element of $H_2(M, \partial M; \mathbb{Z}_2)$. So $S_2 + S_3$ is homologically trivial and therefore has no Möbius band component. (Note that the solid torus does not contain two disjoint properly embedded Möbius bands.) Now, $S_2 + S_3$ has zero Euler characteristic and therefore some component is a clean disc or clean annulus. So, S can be written as the sum of a non-empty normal surface and a clean disc or a clean annulus. We have already ruled that out above. Suppose now that M is the orientable circle bundle over the Möbius band. Then S_2 and S_3 are isotopic to sections of the bundle. However, by assumption, each component of ∂M contains a component of C , which lies in P , and hence M does not admit a clean section. This contradiction completes the proof. \square

As above, let M be a circle bundle, and suppose that C is a union of finitely many fibres with $C \cap \partial M \subset P$, and that A is a collection of annuli with boundary in C . Suppose that these divide M into solid tori, and that each component of ∂M contains at least one component of C . A *broken section* for M is obtained by picking distinct angles in $[0, 2\pi)$, one for each solid torus, and isotoping the intersection between S and each solid torus, so that for any fibre in that solid torus, its intersection with S has been moved through the relevant angle.

Proposition 2.24. *Let M , P , C , A , E_1 and E_2 be as in Proposition 2.23. Let S be a properly embedded surface that is isotopic, via an isotopy preserving A , to a section. Suppose that, among such surfaces, S intersects P as few times as possible. Then M has a broken section that is $(3c, k)$ -nearly exponentially controlled, where $c = -\chi(S) + 2|S \cap P| + 2|S|$ and $k = 2^{15}$.*

Proof. By Proposition 2.23, there is a homeomorphism of M preserving each vertex, edge and circle of P and each component of A , taking S to a weakly fundamental normal surface. By Proposition 2.6,

$$w(S) \leq t^2 2^{7t+6} (-\chi(S) + 2|S \cap P| + 2|S|).$$

We now perturb it to a broken section. By choosing the perturbation angles to be sufficiently close to zero, the resulting surface is still normal and it has at most three times the weight. (The weight on the edges of C will increase by a factor of three, since the three adjacent meridian discs have been perturbed to be disjoint.) Hence, it is $(3c, k)$ -nearly exponentially controlled. \square

3. HIERARCHIES

In this section, we recall some of the main concepts from the theory of hierarchies, as developed by Haken [9] and elucidated by Matveev [18].

3.1. Hierarchies and boundary patterns. Let M be a compact orientable 3-manifold with a boundary pattern P . A *partial hierarchy* for M is a sequence of surfaces S_1, \dots, S_n and a sequence of 3-manifolds $M = M_1, \dots, M_{n+1}$, where each S_i is a properly embedded orientable incompressible surface in M_i , and each M_{i+1} is obtained from M_i by cutting along S_i . We write this as

$$M = M_1 \xrightarrow{S_1} M_2 \xrightarrow{S_2} \dots \xrightarrow{S_n} M_{n+1}.$$

A partial hierarchy is a *hierarchy* if M_{n+1} is a collection of 3-balls.

In the partial hierarchies that we consider, the boundaries of the surfaces will be in general position. Hence, ∂S_i intersects $P \cup \partial S_1 \cup \dots \cup \partial S_{i-1}$ in finitely many points. Also, ∂S_i is non-embedded at only finitely many points. This latter phenomenon occurs when S_i runs over both sides of an earlier surface. We will ensure that this latter situation does not arise with the hierarchies that we consider.

The topological space $S_1 \cup \dots \cup S_i$ is homeomorphic to a 2-complex, but this construction cannot be made in a canonical way, in general. We therefore introduce the following structure.

We say that a space is a *generalised 2-complex* if it is obtained as follows. Let the *1-skeleton* K^1 be the disjoint union of finitely many circles and a finite 1-complex. Then attach a compact surface F via an attaching map $f: \partial F \rightarrow K^1$. We say that the image of each component of F is a *generalised 2-cell*.

Note that $S_1 \cup \dots \cup S_i$ has the structure of a generalised 2-complex, as follows. The union $P \cup \partial S_1 \cup \dots \cup \partial S_i$ is the 1-skeleton. The 0-cells are the points where some ∂S_i meets P , some other ∂S_j or itself.

Let Γ_i be $P \cup \partial S_1 \cup \dots \cup \partial S_i$. The inverse image of Γ_i in M_{i+1} is a union of simple closed curves and graphs P_{i+1} embedded within ∂M_{i+1} . This is the boundary pattern that M_{i+1} inherits from the partial hierarchy.

Note that, when visualising this boundary pattern, it is convenient to regard each surface S_i as ‘transparent’, in the following sense. When a later surface S_j has boundary that runs over a part of ∂M_j associated with S_i , then boundary pattern is created not only at this part of ∂M_{j+1} but also in the part of ∂M_{j+1} lying on the other side of S_i . See Figure 9, for example.

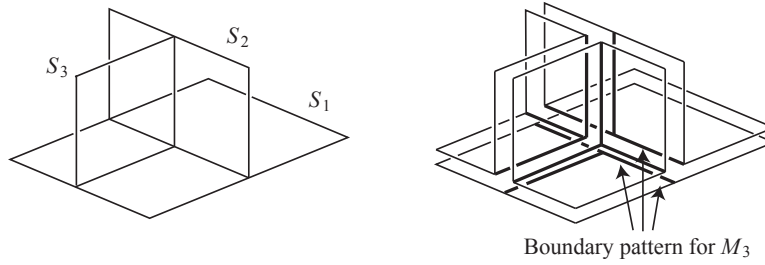


FIGURE 9. A partial hierarchy and associated boundary pattern

The following simple fact is well known. See, for instance, the discussion following Definition 6.3.6 in [18].

Lemma 3.1. *Let (M, P) be a compact orientable irreducible boundary-irreducible 3-manifold with boundary pattern. Let*

$$M = M_1 \xrightarrow{S_1} M_2 \xrightarrow{S_2} \dots \xrightarrow{S_n} M_{n+1}$$

be a partial hierarchy, where no component of any S_i is a sphere or clean disc. For each i , let P_i be the boundary pattern that M_i inherits. Then (M_i, P_i) is irreducible and boundary-irreducible.

In our proof of Theorem 1.1, we will reduce to the case where K is a non-split link other than the unknot. In the hierarchies we consider, the initial manifold M will be the exterior of K . By Lemma 3.1, the manifolds with boundary pattern in such a hierarchy will all be irreducible and boundary-irreducible.

3.2. Exponentially controlled hierarchies. Let M be a compact orientable 3-manifold with boundary pattern P . Let

$$M = M_1 \xrightarrow{S_1} M_2 \xrightarrow{S_2} \dots \xrightarrow{S_n} M_{n+1}$$

be a hierarchy. We say that it is (c, k) -*exponentially controlled* if, for each $i \leq n$ and each triangulation of M with t tetrahedra, in which P, S_1, \dots, S_{i-1} are simplicial, there is a homeomorphism of M that is isotopic to the identity on ∂M and that preserves P and S_j , for each $j < i$, and that takes S_i to a surface with weight at most ck^t .

Remark 3.2. Clearly, if each surface in the hierarchy is exponentially controlled, then the entire hierarchy is also exponentially controlled. However, there is one circumstance where we cannot ensure that the surfaces are exponentially controlled. This is when the manifold M_i is a union of solid tori and 3-balls, and where the solid tori patch together to form a circle bundle embedded in M . Any two such solid tori intersect in a collection of clean annuli. The next surface in the hierarchy will be a broken section. It might not be the case that any two sections of the circle bundle are strongly equivalent. However, applying the techniques of Section 2.15, and using specific properties of the partial hierarchy that we have constructed, we are able to find a homeomorphism of M that is isotopic to the identity on ∂M and preserving all the surfaces in the partial hierarchy, taking one section to another.

The following result will be crucial.

Theorem 3.3. *Let M be the exterior of a non-split link in the 3-sphere other than the unknot. For each component of ∂M that is incident to a Seifert fibred JSJ piece other than $T^2 \times [0, 1]$, suppose that $P \cap \partial M$ is a single curve with slope 0 or ± 1 , other than the fibre slope. Suppose also that P is disjoint from the other components of ∂M . When M is a copy of $T^2 \times [0, 1]$, then suppose that $\partial M \cap P$ consists of a longitude in each component of ∂M . Then (M, P) has an exponentially controlled hierarchy.*

The outline of the proof is as follows. Suppose that we have formed a partial hierarchy that is exponentially controlled, and let M_i be the final manifold. Suppose that it is not a disjoint union of 3-balls. We would like to find an exponentially controlled surface in M_i . Since M_i is not a disjoint union of 3-balls, it contains a properly embedded (non-separating) orientable incompressible boundary-incompressible surface S . In the case that M_i contains no properly embedded essential clean annuli or essential tori, then Theorem 2.8 can probably be used to make S weakly fundamental. On the other hand, if M_i contains an essential clean annulus or essential torus, then it probably contains a JSJ torus or annulus. In this case, Theorem 2.14 can be used to make the JSJ annuli and tori fundamental.

There are, of course, several complications. One is that Theorem 2.8 also has a hypothesis about the way that the boundary pattern intersects the incompressible boundary tori of M_i , and this leads to some technical issues.

The second complication is that it is not clear that the above procedure terminates. Fortunately, one can guarantee that the decomposition along S reduces some notion of complexity, and hence that the hierarchy eventually terminates, as long as one is careful about the choice of surface S . A significant complication is that the JSJ annuli in M_i may in fact be boundary parallel. Cutting along such a surface produces a solid torus together with a copy of the original manifold, but with different boundary pattern. In order not to get stuck in a repeating loop, considerable care is required here.

The third main complication is that it is possible for a manifold to contain an essential clean annulus or essential torus, and yet have no JSJ torus or annulus. This is clearly problematic, because neither Theorem 2.8 nor Theorem 2.14 can be used. But according to Theorem 2.9, this only occurs in two very restricted situations: when a component of M_i is Seifert fibred with pattern consisting of fibres, and when a component of M_i is an I -bundle over a surface and the pattern P_i is a collection of simple closed curves in the vertical boundary of the bundle.

This second case arises, for example, when the first surface in the hierarchy is a fibre in a fibration over the circle. When Haken first developed the theory of hierarchies [9], this was a situation that he could not handle. Fortunately, in the case of a non-split link in the 3-sphere other than the unknot, it is now known that one can avoid this situation, as follows.

Let S be a surface properly embedded in a compact orientable 3-manifold M . Then S is a *fibre surface* if it is a fibre in a fibration of M over the circle. This is equivalent to $M \setminus S$ having the structure of an I -bundle with horizontal boundary equal to the copies of S .

The surface S is a *semi-fibre* if it is connected and separating and each component of $M \setminus S$ is homeomorphic to an I -bundle over some non-orientable surface, with S forming the horizontal boundary of the bundle. Semi-fibres do not arise as properly embedded surfaces when M is the exterior of a link in the 3-sphere. This is because the zero-section of each I -bundle is a non-orientable surface, and one may join these two surfaces by annuli in the boundary to form a closed non-orientable surface properly embedded in M . However, the 3-sphere does not support any such surface.

The following is due to Culler and Shalen [7].

Theorem 3.4. *Let M be the exterior of a hyperbolic link. Then M contains a properly embedded orientable incompressible, boundary-incompressible surface that is neither a fibre nor a semi-fibre, and with no sphere or disc components.*

We now give a procedure for finding an exponentially controlled hierarchy for the exterior M of a non-split link other than the unknot. We give M the following boundary pattern. For each component of ∂M that is incident to a Seifert fibred JSJ piece of M other than $T^2 \times [0, 1]$, we give it boundary pattern that equals a single curve with slope 0 or ± 1 , chosen so that its slope does not agree with the fibre slope. The remaining components of ∂M are given empty boundary pattern. In the case where M is a copy of $T^2 \times [0, 1]$ (which is exactly when the link is the Hopf link), then we pick boundary pattern that is a longitude on each boundary component.

- (1) If M has any JSJ tori, decompose along two parallel copies of each torus. Let X be the union of the copies of $T^2 \times I$ between these tori. We will not touch any component of X until both of its boundary components have inherited non-empty boundary pattern.
- (2) For each Seifert fibred JSJ piece of $M_2 - X$, pick a properly embedded annulus as in Section 2.13. Insert one or two copies of it, depending on whether the annulus is separating or non-separating. Repeat this process until each Seifert fibred component of $M_2 - X$ has been decomposed into solid tori.
- (3) In each hyperbolic complementary region, find a properly embedded orientable incompressible boundary-incompressible surface that is not a union of fibre surfaces and that has no sphere or disc components. Cut along it or two parallel copies of it, depending on whether it is separating or non-separating.
- (4) We now embark upon a sequence of the following decompositions. The process continues until one of two situations arises. If every component of the 3-manifold is a ball, then the procedure terminates with the required hierarchy. If every component of the 3-manifold is a 3-ball or a solid torus

with a clean longitude, and there is at least one component of the latter type, then pass to step (5). The decompositions are as follows:

- (i) If a component of the 3-manifold contains a clean essential annulus that is parallel to an annulus in the boundary, then cut along this annulus.
 - (ii) If a component of the 3-manifold does not contain an annulus as in (i), but has some JSJ annuli, then cut along them all.
 - (iii) Suppose that a component of the 3-manifold is a copy of $T^2 \times [0, 1]$ with non-empty pattern consisting of essential parallel curves. If it is a component of X , suppose also that both components of $T^2 \times \{0, 1\}$ have non-empty intersection with the pattern. Pick an essential clean annulus, joining the two boundary components, and cut along two copies of it.
 - (iv) Suppose that a component of the 3-manifold is a copy of $T^2 \times [0, 1]$ and is simple (with respect to the boundary pattern). If it is a component of X , suppose also that both components of $T^2 \times \{0, 1\}$ have non-empty intersection with the pattern. Pick an incompressible boundary-incompressible annulus joining the two boundary components and cut along two copies of it.
 - (v) If a component of the 3-manifold is simple (with respect to the boundary pattern), and its boundary has a compression disc (that intersects the boundary pattern), but the manifold is not a solid torus with a clean longitude, then pick a disjoint collection of such discs that are boundary-incompressible (with respect to the boundary pattern), taking two parallel copies of a disc if it is non-separating, and cut along this collection.
 - (vi) If a component of the 3-manifold is simple and its boundary is incompressible, but it is not a copy of $T^2 \times [0, 1]$, then, in this component, pick a properly embedded orientable incompressible surface with non-empty boundary, with no sphere or disc components, that is boundary-incompressible with respect to the inherited pattern and with respect to the empty pattern. Cut along one or two parallel copies of it, depending on whether it is separating or non-separating.
- (5) By reaching this stage, every component of the 3-manifold is a 3-ball or a solid torus with a clean longitude. We now perform the following decompositions until no more such decompositions are possible.
- (i) If a solid torus contains a necklace annulus in its boundary, then push this annulus into the manifold so that it becomes properly embedded, and then cut along it. This creates a new solid torus. Cut along two copies of a meridian disc for this solid torus that intersects the pattern as few times as possible.
 - (ii) If two solid tori have a component of intersection that is a disc, then cut along a parallel copy of this disc lying in one of the solid tori.
- (6) Upon reaching this stage, it is still the case that every component of the 3-manifold is a 3-ball or a solid torus with a clean longitude. Consider the union of these solid tori. This is a circle bundle over a surface. The interior of the circle bundle intersects the union of the previous surfaces in the partial hierarchy in a union of annuli A . Let S be a properly embedded surface

that is isotopic, via an isotopy preserving A , to a section. Suppose that, among such surfaces, S intersects the boundary pattern P_n as few times as possible. Then M has a broken section that is $(3c, k)$ -nearly exponentially controlled, where $c = -2\chi(S) + |S \cap P_n| + 2|S|$ and $k = 2^{15}$. This therefore decomposes the solid tori to 3-balls.

Remark 3.5. In the above procedure, each surface S_i lies in a single component of M_i . However, that restriction can be weakened, as follows. We can instead require that no two components of M_i that contain components of S_i are incident across an earlier surface of the hierarchy.

Theorem 3.6. *This procedure produces a $(\mu, 2^{15})$ -exponentially controlled hierarchy, where μ is the maximal value of $-6\chi(S_i) + 3|S_i \cap P_i| + 6|S_i|$ over all surfaces S_i in the hierarchy, and where P_i is the boundary pattern of the i th manifold.*

Proof. The fact that each of these surfaces is $(\mu, 2^{15})$ -exponentially controlled follows from the results in Section 2, apart from the broken section in (6) which we will discuss below. Specifically, the fact that each JSJ torus is fundamental is the content of Theorem 2.14. Propositions 2.16, 2.17 and 2.18 give that the surfaces in (2) are weakly fundamental. Theorem 2.8 implies that the surfaces in (3) are weakly fundamental. Similarly, the surfaces in (4) and (5) are exponentially controlled, either by Theorem 2.8, Theorem 2.14, Proposition 2.17, Proposition 2.18 or Lemma 2.19. However, Theorem 2.8 has some hypotheses about toral boundary components of ∂M_i that need to be checked.

We claim that from step (3) onwards M_i has no essential tori. Suppose that on the contrary there is a properly embedded essential torus T in M_i . Note that T is disjoint from the surfaces S_1, \dots, S_{i-1} because it lies in the interior of M_i . Since T is essential in M_i , it is π_1 -injective in M_i . As at each stage of the partial hierarchy, we have decomposed along a 2-sided incompressible surface, M_i π_1 -injects into M . Hence, T is incompressible in M . It is disjoint from the JSJ surfaces in M , since these are the first decomposition in (1). It does not lie in a Seifert fibred piece, since these are decomposed into solid tori in (2). Hence T is boundary parallel in M or parallel to a JSJ torus of M . So, there is a copy of $T^2 \times [0, 1]$ between this torus and a boundary component of M_2 , where $T^2 \times \{1\} = T$. Now each surface in the partial hierarchy is incompressible, and the only connected incompressible surfaces properly embedded in $T^2 \times [0, 1]$ disjoint from $T^2 \times \{1\}$ are parallel to a subsurface of $T^2 \times \{0\}$. Thus, we deduce that in M_i , there is a copy of $T^2 \times [0, 1]$ with $T^2 \times \{0\}$ being a component of ∂M_i and $T^2 \times \{1\} = T$. In other words, T is boundary parallel in M_i and hence not essential, as claimed.

We also claim that from step (3) onwards, each incompressible toral boundary component of M_i either is clean or lies in a component of M_i that is a copy of $T^2 \times [0, 1]$. Consider an incompressible toral boundary component T of some M_i . As argued above, T is boundary parallel in M or parallel to a JSJ torus of M . Suppose that the component of M_i containing T is not a copy of $T^2 \times [0, 1]$. Hence, T lies in $M_2 - X$. If T is a boundary component of M_2 , then it must lie in a hyperbolic piece of M_2 since the Seifert fibred pieces are decomposed into solid tori in (2). But then T is clean, because any component of X on the other side of T is not decomposed until the copy of T in X inherits non-empty pattern, by which stage T is no longer a component of ∂M_i . If T intersects a component of ∂M_2 but does not lie entirely in ∂M_2 , then consider the first surface S_j that intersects

the region between T and ∂M_2 . This is parallel to a clean subsurface of ∂M_j and hence is boundary-compressible, contrary to our construction. Finally, suppose that T is disjoint from ∂M_2 . Then it is parallel to a boundary component T' of M_2 . Between T and T' is a copy of $T^2 \times [0, 1]$, and as argued above, this implies that the component of M_i containing T is a copy of $T^2 \times [0, 1]$, as claimed.

When a component of M_i that is a copy of $T^2 \times [0, 1]$ is decomposed, we ensure in (4)(iii) and (4)(iv) that we cut along an annulus joining the two boundary components. In (4)(iii), we use Proposition 2.18. In (4)(iv), we use Proposition 2.17. Hence, when we apply Theorem 2.8, its hypotheses hold. In particular, when we use it in steps (3) and (4), any incompressible toral boundary components in the relevant component of M_i are clean.

We need to show that the process terminates, with the final manifold being a collection of balls. To do so, we use the machinery developed by Matveev in [18] that was based on work of Haken [9].

The procedure described here should be compared with the procedure given by Matveev in [18]. The initial part of Matveev's procedure, described in [18, Section 6.5.1] is to perform the following moves whenever they are possible:

- (E1) If (M_i, P_i) contains an essential torus, then add one or two parallel copies of such a torus depending on whether the torus is separating or non-separating.
- (E2) If (M_i, P_i) contains a clean essential longitudinal annulus that is not parallel to an annulus in ∂M_i , then add one or two parallel copies of such an annulus, depending on whether the annulus is separating or non-separating.
- (E3) If (M_i, P_i) contains a clean essential annulus that is parallel to an annulus in the boundary, then cut along this annulus.
- (E4) If a component of (M_i, P_i) is simple, and its boundary has a compression disc, but the manifold is not a solid torus with a clean longitude, then pick a compression disc that intersects P_i as few times as possible, and cut along one or two parallel copies of this, depending on whether it is separating or non-separating.
- (E5) If a component of (M_i, P_i) is simple and its boundary is incompressible, then, in this component, pick a properly embedded connected orientable incompressible boundary-incompressible surface F with non-empty boundary, other than a sphere or disc, and for which $-\chi(F) + |F \cap P_i|$ is as small as possible. Then cut along one or two parallel copies of it, depending on whether it is separating or non-separating.

Matveev shows in Corollary 6.5.7 in [18] that the above procedure terminates, although the final manifold is not necessarily a 3-ball. Matveev then uses further 'extension moves' (E6)-(E10) and 'auxiliary moves' (E3'), (E4') and (E4'').

We do not follow Matveev's approach precisely for several reasons. First of all, we require the hierarchy to be exponentially controlled, whereas Matveev requires less of his hierarchies. Secondly, we are in a more particular situation where the initial manifold is the exterior of a link, whereas he has to work with more general Haken 3-manifolds.

We claim that by the time we have completed steps (1)-(4) above, none of Matveev's moves (E1)-(E5) can be applied. Hence, the manifold with boundary pattern that we obtain after completing moves (1)-(4) has all the properties that

Matveev requires before he embarks on (E6)-(E10) and (E3'), (E4') and (E4''). Let (M_i, P_i) be the manifold with pattern obtained after steps (1)-(4).

Suppose now that some component of $M_i - X$ contains a clean essential longitudinal annulus, but that none of 4(i), 4(ii) and 4(iii) can be applied. After a pattern-isotopy, this annulus lies in a JSJ piece for M_i , which by Theorem 2.9 is Seifert fibred with pattern consisting of fibres. This is not a solid torus, because such a manifold does not support an essential longitudinal annulus. The Seifert fibre pieces of M were decomposed in (2). Hence, the component of $M_i - X$ containing the annulus is a copy of $T^2 \times [0, 1]$. It might be a regular neighbourhood of a JSJ torus for M or it might be parallel to component of ∂M . Its pattern must be non-empty, since the annulus is longitudinal. Hence, 4(iii) can be applied to it, which is contrary to hypothesis.

So we have shown that neither (E1) nor (E2) can be applied to $M_i - X$. Note that (E3) is exactly 4(i). Note further that if (E4) can be applied, then so too can 4(v). Similarly, if (E5) can be applied to a component of $M_i - X$, then so too can 4(vi) or (iv). Since (4) has been completed, we deduce that none of (E3) - (E5) can be applied to $M_i - X$.

Hence, by Lemma 6.5.10 in [18], each component of $(M_i - X, P_i - X)$ is either a 3-ball, a solid torus having a clean longitude or an I -bundle chamber. Matveev defines an *I-bundle chamber* to be an I -bundle over some compact surface with negative Euler characteristic, with boundary pattern consisting a non-empty collection of simple closed curves in each vertical boundary component. He also requires that in each vertical boundary component of the I -bundle, its two boundary curves are part of the pattern.

We will show that in fact there are no I -bundle chambers. As explained by Matveev in the proof of Proposition 6.5.14 in [18], each I -bundle chamber has horizontal boundary equal to the horizontal boundary of another I -bundle chamber. Hence, by [18, Proposition 6.5.14], the union of the I -bundle chambers, minus a small regular neighbourhood of the vertical boundary, consists of the following four types of components:

- (1) a copy of $F \times I$ where F is a compact orientable surface, with $F \times \partial I \subset \partial N(K)$;
- (2) a twisted I -bundle $F \tilde{\times} I$, where F is a compact non-orientable surface, with the horizontal boundary of the I -bundle in $\partial N(K)$;
- (3) a fibration over the circle with fibre a compact orientable surface;
- (4) a union of two twisted I -bundles glued along their horizontal boundary, called a *quasi-Stallings manifold*.

We say that a component of the union of the I -bundle chambers, minus a small regular neighbourhood of their vertical boundary, is a *chamber compendium*. Its *vertical boundary* is the union the vertical boundaries of its constituent I -bundles. It is a collection of annuli and tori properly embedded in M , the exterior of K .

We claim that the vertical boundary of each chamber compendium is incompressible in M . Suppose not, and hence that the vertical boundary is compressible. Hence, there is a vertical boundary component T that admits a compression disc with interior disjoint from the vertical boundary. Any compressible torus in the irreducible 3-manifold M either lies in a 3-ball or bounds a solid torus. Any compressible annulus lies in a 3-ball. The surface T is incident to at least one horizontal boundary component F of an I -bundle chamber. Each component of ∂F represents

a non-trivial element of the fundamental group of the chamber, and hence is a non-trivial element of $\pi_1(M)$. Hence, T cannot lie in a 3-ball. It is therefore a torus bounding a solid torus V . Consider the first surface S_i in the hierarchy to intersect V . Then $S_i \cap V$ is incompressible in V , for the following reason. The boundary of a compression disc for $S_i \cap V$ would have to bound a disc in S_i , since S_i is π_1 -injective. This disc does not lie wholly in V . But each surface in the hierarchy incident to V but not lying wholly in V must contain a horizontal boundary component F of an I -bundle chamber incident to V . Hence, we deduce that a boundary component of F is homotopically trivial in M , which is ruled out above. Since $S_i \cap V$ is incompressible and orientable in the solid torus V , it is a collection of incompressible annuli and discs. In fact, it cannot contain any discs, since this would again imply that a horizontal boundary component of a chamber did not π_1 -inject into M . There is therefore some annulus A of $S_i \cap V$ that is outermost in V . This is incident to horizontal boundary components F and F' of I -bundle chambers. Lying between F and F' is a union of I -bundle chambers, which is again an I -bundle. It is possible that $F = F'$ if they bound a twisted I -bundle. Enlarge this I -bundle so that it contains the region in V between A and ∂V . This is again an I -bundle, but A is now a vertical boundary component. Now the base of this I -bundle has negative Euler characteristic and so contains an essential properly embedded arc with both endpoints on the boundary component corresponding to ∂A . Over this arc is a disc E consisting of I -fibres, with $E \cap S_i = \partial E$. By the incompressibility of S_i , ∂E bounds a disc E' in S_i . The intersection of E' with F is a planar surface component of $F \setminus \partial E$. Hence, we deduce that some component of ∂F bounds a disc in S_i , which we have shown to be impossible.

Thus, we have shown that the vertical boundary of each chamber compendium is incompressible.

The vertical boundary of chamber compendium cannot contain any annulus component, since (1) or (2) would then apply, and hence the horizontal boundary of the bundle would lie in ∂M . It would then be a disc, annulus or torus, all of which are ruled out by the assumption that the base of the bundle has negative Euler characteristic.

So, the vertical boundary of the chamber compendiums consists of incompressible tori. Each torus is therefore boundary parallel or JSJ or lying in the interior of a Seifert fibre piece of the JSJ decomposition of M . The latter situation can be ruled out as follows. Suppose some vertical boundary component lies in a Seifert fibre piece but is not JSJ or boundary parallel. This Seifert fibre space has been decomposed into solid tori in step (2). Hence, we deduce that some I -bundle chamber in the compendium is a 3-ball or solid torus, but the I -bundle chambers cannot be 3-balls or solid tori.

Hence, each chamber compendium is either a component of the JSJ decomposition of M or a copy of $T^2 \times I$ lying in a regular neighbourhood of a JSJ torus or a component of ∂M . The latter cannot happen because each chamber compendium is a surface bundle or Stallings manifold with fibre having negative Euler characteristic. Thus, we deduce that each chamber compendium is a component of the JSJ decomposition of M . The intersection between each surface of the hierarchy and the chamber compendium is horizontal in the chamber compendium. However, in step (3), we decomposed the hyperbolic JSJ pieces along a surface that was neither a fibre nor a semi-fibre, and we decomposed the Seifert fibred pieces into solid tori.

Thus, we have shown that there are no I -bundle chambers.

So, once Step (4) has been completed, each component of $M_i - X$ is a 3-ball or solid torus. We claim in fact that each component of M_i is a 3-ball or solid torus. Consider any component of X . The two components of $M_i - X$ that are incident to it have been decomposed into solid tori and 3-balls. Hence, both its boundary components have inherited non-empty boundary pattern. Therefore the restriction in 4(iii) and (iv) on decomposing such a component of X no longer applies. Hence, this component of X has indeed been decomposed into solid tori in 4(iii) or 4(iv). Any further decompositions to these components only give solid tori and 3-balls. This proves the claim.

We now apply the moves in (5). These are exactly moves (E3'), (E4') of Matveev [18]. Note that Matveev's move (E4'') is not possible in our situation, since it only applies when the initial manifold M_1 is closed. After these moves are performed, the union of the solid torus chambers are what Matveev terms *faithful* (as in Definition 6.5.52 in [18]). He proves in Lemma 6.5.53 that when this holds, the union of the solid toral chambers is a circle bundle over a surface. Moreover, each solid torus chamber is a union of circle fibres, and each component of intersection between them is an annulus that is a union of circle fibres. Thus, in step (6), this circle bundle is decomposed along a broken section. In particular, each remaining solid torus is decomposed along a meridian disc. Thus this procedure does terminate with a hierarchy.

We need to show that this hierarchy is $(\mu, 2^{15})$ -exponentially controlled. This was explained at the beginning of the proof for all surfaces in the hierarchy except the broken section, which is only $(\mu, 2^{15})$ -nearly exponentially controlled, by Proposition 2.24. (See Section 2.15 for the precise definition of being nearly exponentially controlled.) In our case, the surface that we are considering is a broken section, that is obtained from a section of a circle bundle Y lying in M . Let P be the boundary pattern for Y . Then $\partial Y \setminus P$ is a collection of discs and annuli. The annular components lie in ∂M and are vertical in the fibration. Any homeomorphism of Y that preserves each vertex, edge and circle of P is, up to pattern-isotopy, a composition of Dehn twists about clean vertical annuli and tori. A Dehn twist about a vertical torus restricts to the identity on ∂M and hence determines a strong equivalence of M . Consider a Dehn twist about a vertical annulus R . We claim that this Dehn twist is isotopic to the identity on M .

Now, we assigned boundary pattern to components of ∂M that were incident to Seifert fibred JSJ pieces, and this pattern consisted of a single essential curve not having a fibre slope. Hence, when the Seifert fibred pieces are decomposed in step (2), these components of ∂M incident to JSJ pieces inherit pattern that decomposes these tori into discs. We therefore deduce that the annular components of $\partial Y \setminus P$ are disjoint from those components of ∂M . The annulus R has clean boundary in ∂M , and hence misses the Seifert fibred pieces of M . It is therefore inessential in M . Hence, a Dehn twist about R gives a homeomorphism of M that is isotopic to the identity, as claimed.

So suppose that we are given a triangulation of M with t tetrahedra in which the pattern and all the surfaces before the broken section are simplicial. The broken section is $(\mu, 2^{15})$ -nearly exponentially controlled, and hence, as argued above, there is a homeomorphism of M , isotopic to the identity on ∂M , preserving the earlier

surfaces in the hierarchy, taking the broken section to one that has weight at most $\mu \cdot 2^{15t}$, as required. \square

3.3. Adequately separating hierarchies. Let

$$M = M_1 \xrightarrow{S_1} M_2 \xrightarrow{S_2} \cdots \xrightarrow{S_n} M_{n+1}$$

be a hierarchy. This is said to be *adequately separating* if the following two conditions hold:

- (1) for each $i < n$, there is no curve in M_i that intersects S_i transversely once, and
- (2) for each $i < n$, S_i is contained within a single component of M_i .

Furthermore, either these two conditions hold for $i = n$ or S_n is broken section of a circle bundle.

The first condition implies that for each component Y of $M_i \setminus S_i$, the restriction of $M_i \setminus S_i \rightarrow M_i$ to Y is an injection. Hence, with both these conditions, we deduce that no surface S_j ($j < n$) runs over two parts of ∂M_j that are identified in M . Note that this remains true for S_n even when it is a broken section of a circle bundle. This is technically useful, since ∂S_j creates boundary pattern, and if two parts of ∂S_j were to be identified in M , then this would create possibly unexpected vertices in the boundary pattern.

We note the following, particularly in reference to Remark 3.5.

Remark 3.7. The hierarchy constructed in Theorem 3.6 is adequately separating.

4. HANDLE STRUCTURES

We will deal with handle structures on 3-manifolds. These will always have the following properties:

Convention 4.1. Let \mathcal{H} be a handle structure on a 3-manifold. Then, for $0 \leq i \leq 3$, the union of the i -handles is denoted \mathcal{H}^i . The *index* of an i -handle is i . We will always require the handle structure \mathcal{H} to satisfy the following conditions.

- (1) For each i -handle $D^i \times D^{3-i}$, its intersection with the handles of lower index is $\partial D^i \times D^{3-i}$.
- (2) For each 2-handle $H_2 = D^2 \times D^1$ and 1-handle $H_1 = D^1 \times D^2$, the intersection $H_1 \cap H_2$ must respect the product structure of each, in the sense that $H_1 \cap H_2$ is equal to $\alpha \times D^1$ in H_2 , where α is a collection of arcs in ∂D^2 , and is equal to $D^1 \times \beta$ in H_1 , for a collection of arcs β in ∂D^2 .
- (3) Each 2-handle runs over at least one 1-handle.

A triangulation of a 3-manifold M determines a handle structure for M , where each i -simplex not lying in ∂M gives rise to a $(3-i)$ -handle.

When analysing handle structures of 3-manifolds, the following surface is of particular relevance. We let \mathcal{F} be $\mathcal{H}^0 \cap (\mathcal{H}^1 \cup \mathcal{H}^2)$. This is a subsurface of the spheres $\partial \mathcal{H}^0$. It has a handle structure where the 0-handles, denoted \mathcal{F}^0 , are $\mathcal{H}^0 \cap \mathcal{H}^1$, and the 1-handles, denoted \mathcal{F}^1 , are $\mathcal{H}^0 \cap \mathcal{H}^2$. Note that \mathcal{F}^1 is indeed a collection of 1-handles, because each 2-handle of \mathcal{H} runs over at least 1-handle of \mathcal{H} , and so \mathcal{F}^1 has no annular components.

4.1. Normal surfaces in handle structures. Let \mathcal{H} be a handle structure for a 3-manifold M . Then a surface properly embedded in M is said to be in *standard form* with respect to \mathcal{H} if

- (1) it is disjoint from the 3-handles;
- (2) it intersects each 2-handle $D^2 \times D^1$ in discs of the form $D^2 \times F$, for some finite set of points F in the interior of D^1 ;
- (3) it intersects each 1-handle $D^1 \times D^2$ in discs of the form $D^1 \times \alpha$, for a collection of properly embedded arcs α in D^2 ;
- (4) it intersects each 0-handle in a collection of properly embedded discs.

A standard surface S is *normal* if, for each component D of $S \cap \mathcal{H}^0$, the following hold:

- (1) D intersects each component of \mathcal{F}^1 in at most one arc;
- (2) D intersects each component of $\partial\mathcal{H}^0 \setminus \mathcal{F}$ in at most one arc and no simple closed curves that bound discs in $\partial\mathcal{H}^0 \setminus \mathcal{F}$;
- (3) D does not intersect a component of $\partial\mathcal{F}^0 \setminus \mathcal{F}^1$ more than once;
- (4) D does not intersect components of $\partial\mathcal{H}^0 \setminus \mathcal{F}$ and \mathcal{F}^1 that are adjacent.

These components D of $S \cap \mathcal{H}^0$ are called *elementary normal discs*.

4.2. Curves and graphs in the boundary. Let \mathcal{H} be a handle structure for a 3-manifold M . We say that a union of disjoint simple closed curves C in ∂M is *standard* if

- (1) C is disjoint from the 2-handles of \mathcal{H} ;
- (2) C intersects each 1-handle $D^1 \cap D^2$ in $D^1 \times F$ for some finite subset F of ∂D^2 ;
- (3) C intersects each 0-handle in a collection of arcs.

A graph P in the boundary of M is in *transverse form* if

- (1) P is disjoint from the 0-handles of \mathcal{H} ;
- (2) P intersects each 1-handle $D^1 \cap D^2$ in a subset of $\{*\} \times \partial D^2$ for some point $*$ in the interior of D^1 ;
- (3) each vertex of P lies in a 2-handle of \mathcal{H} ;
- (4) for any 2-handle $D^2 \times D^1$ of \mathcal{H} , the intersection between P and a component of $D^2 \times \partial D^1$ is at most one vertex of P plus arcs that each run from the vertex to the boundary of $D^2 \times \partial D^1$.

4.3. The weight and extended weight of a standard surface. The *weight* of a standard surface S in a handle structure \mathcal{H} is the number of components of $S \cap \mathcal{H}^2$. Note that, when the handle structure arises from a triangulation \mathcal{T} of a closed 3-manifold, then there is a bijection between the components of $S \cap \mathcal{H}^2$ and the points of intersection between S and the 1-skeleton of \mathcal{T} . Thus, in this situation, the two measures of weight coincide.

We also introduce another measure of complexity. The *extended weight* of a standard surface S is defined to be

$$|S \cap \mathcal{H}^2| + |S \cap \mathcal{H}^1|.$$

The rationale for this definition is that weight is a somewhat imperfect measure of complexity. For example, there can be surfaces that avoid the 2-handles completely but that run through the 0-handles and 1-handles many times. These have zero weight but large extended weight. Furthermore, we have the following lemma.

Lemma 4.2. *Let \mathcal{H} be a handle structure for a compact 3-manifold M . Let P be a boundary pattern that is in transverse form. Let S be a standard incompressible boundary-incompressible surface with no component that is a sphere or disc intersecting P at most twice. Then S is pattern-isotopic to a normal surface S' such that the extended weight of S' is at most the extended weight of S . Furthermore, for each 2-handle of \mathcal{H} , S' intersects the 2-handle at most as many times as S did.*

This is essentially proved in [18, Theorem 3.4.7], where normalisation moves are given. It is straightforward to check that they do not increase the extended weight.

4.4. Distinguished subsurfaces of the boundary. When we analyse a handle structure of a 3-manifold N , its boundary will sometimes contain a distinguished subsurface X . In addition, ∂N will contain some arcs α properly embedded in $\partial N \setminus X$. These will satisfy the condition that ∂X is a collection of standard curves and that each component of α lies in a 0-handle. We then say that \mathcal{H} is a handle structure for (N, α, X) .

When dealing with such extra structure on the boundary, we will slightly vary the definition of a normal surface S . We now say that a properly embedded surface S is *normal* if it is standard, its boundary is disjoint from X , and for each component D of $S \cap \mathcal{H}^0$, the following hold:

- (1) D intersects each component of \mathcal{F}^1 in at most one arc;
- (2) D intersects each component of $\partial \mathcal{H}^0 \setminus (\mathcal{F} \cup \alpha \cup X)$ in at most one arc and no simple closed curves that bound discs in $\partial \mathcal{H}^0 \setminus (\mathcal{F} \cup \alpha \cup X)$;
- (3) D does not intersect a component of $\partial \mathcal{F}^0 \setminus (\mathcal{F}^1 \cup \alpha \cup X)$ more than once;
- (4) D does not intersect components of $\partial \mathcal{H}^0 \setminus (\mathcal{F} \cup \alpha \cup X)$ and \mathcal{F}^1 that are incident.

We also say that S is *crudely normal* if it is standard, its boundary is disjoint from X , and each component D of $S \cap \mathcal{H}^0$ satisfies (1) above, as well as:

- (3') no component of $D \cap \mathcal{F}^0$ is an arc with endpoints on the same component of $\partial \mathcal{F}^0 \setminus (\mathcal{F}^1 \cup \alpha \cup X)$;
- (4') no component of $D \cap \mathcal{F}^0$ is an arc with endpoints on components of $\partial \mathcal{F}^0 \setminus (\mathcal{F}^1 \cup \alpha \cup X)$ and $\partial \mathcal{F}^0 \cap \mathcal{F}^1$ that are incident.

4.5. Parallelism. Let \mathcal{H} be a handle structure for (N, P, X) , and let S be a normal surface. We say that two disc components D_0 and D_1 of $S \cap \mathcal{H}^i$ are *parallel* if there is an isotopy of N , preserving all the handles of \mathcal{H} and preserving X and P , which takes D_0 onto D_1 , and such that the restriction of this isotopy to D_0 is an embedding of $D_0 \times [0, 1]$ into N . We then say that D_0 and D_1 are of the same *type*.

More generally, we say that two subsurfaces S_0 and S_1 of S are said to be *normally parallel* if there are subsurfaces S'_0 and S'_1 of S , each of which is a union of elementary normal discs, and satisfying $S'_0 \supseteq S_0$ and $S'_1 \supseteq S_1$, and an isotopy $H: N \times [0, 1] \rightarrow N$ preserving all the handles of \mathcal{H} , and preserving X and P , such that the following hold:

- (1) The restriction of H to $S'_0 \times [0, 1]$ is an embedding.
- (2) For each elementary normal disc D of S'_0 and each $t \in [0, 1]$, $H(D, t)$ is an elementary normal disc.
- (3) $H(S'_0 \times \{i\}) = S'_i$ for $i = 0$ and 1 .
- (4) $H(S_0 \times \{i\}) = S_i$ for $i = 0$ and 1 .

5. THE BASIC THEORY OF ARC PRESENTATIONS

In this section, we give an introduction to rectangular diagrams and arc presentations, as developed mostly by Cromwell [6] and Dynnikov [8].

5.1. Rectangular diagrams. A *rectangular diagram* is a diagram in the plane $\mathbb{R} \times \mathbb{R}$, consisting of arcs, each of which is horizontal (of the form $[s_1, s_2] \times \{t\}$ for $s_1, s_2, t \in \mathbb{R}$) or vertical (of the form $\{s\} \times [t_1, t_2]$ for $s, t_1, t_2 \in \mathbb{R}$). Each crossing arises when a vertical arc and a horizontal arc cross; the vertical arc is required to be the over-arc. Moreover, no two arcs may be collinear. (See Figure 10.)

There are the same number of horizontal and vertical arcs. This is the *arc index* of the rectangular diagram.

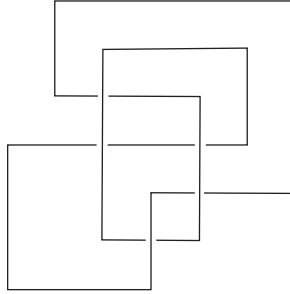


FIGURE 10. A rectangular diagram

It is an elementary matter to convert any link diagram into a rectangular diagram by an isotopy of the plane. In [17, Lemma 2.1], the following was proved, which gives a quantified version of this observation.

Lemma 5.1. *Let D be a diagram of a link with c crossings. Then D is isotopic to a rectangular diagram with arc index at most $(81/20)c$.*

5.2. Arc presentations. Closely related is the concept of an arc presentation for a link.

One fixes a description of the 3-sphere as the join $S^1 * S^1$ of two circles. Thus, a point in S^3 is specified by three coordinates (ϕ, τ, θ) , where $\phi, \theta \in \mathbb{R}/(2\pi\mathbb{Z})$ and $0 \leq \tau \leq 1$. The points $(\phi, 0, \theta)$ and $(\phi, 0, \theta')$ are identified for all θ and θ' . The resulting circle $\{\tau = 0\}$ is denoted S^1_ϕ and is called the *binding circle*. Similarly, $(\phi, 1, \theta)$ and $(\phi', 1, \theta)$ are identified for all ϕ and ϕ' , and the resulting circle $\{\tau = 1\}$ is written S^1_θ . For $t \in \mathbb{R}/(2\pi\mathbb{Z})$, the open disc $\{\theta = t, \tau > 0\}$ is termed a *page* and is denoted \mathcal{D}_t .

A link K is in an *arc presentation* if

- (1) it intersects the binding circle in finitely many points, called *vertices*;
- (2) its intersection with each page is either empty or a single arc joining distinct vertices.

The number of vertices is equal to the number of arcs, and is called the *arc index* of the arc presentation.

5.3. The relationship between arc presentations and rectangular diagrams. An arc presentation determines a rectangular diagram as follows. Let the arc index of the arc presentation be n . For each arc in a page \mathcal{D}_t joining vertices s_1 and s_2 in $[0, 2\pi)$, insert the horizontal arc $[s_1, s_2] \times \{t\}$ in the plane. For each vertex $s \in S_\phi^1$, let t_1, t_2 be the θ values in $[0, 2\pi)$ of the arcs incident to it, and insert a vertical arc $\{s\} \times [t_1, t_2]$ into the plane. These vertical and horizontal arcs in the plane specify a rectangular diagram, once we declare that at each crossing, the over-arc is the vertical one.

It is clear that this procedure may be reversed: given a rectangular diagram, one may form the corresponding arc presentation.

It was shown by Cromwell [6] that the arc presentation and the rectangular diagram specify the same link.

5.4. Modifications to a rectangular diagram. In [6], Cromwell introduced a collection of modifications that one can make to a rectangular diagram:

- (1) *Cyclic permutation*: Here one cyclically permutes the horizontal (or vertical arcs).

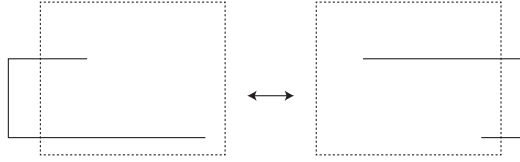


FIGURE 11. Cyclic permutation of the vertical arcs

- (2) *Stabilisation/destabilisation*: These are local modifications to the diagram, as shown in Figure 12, which increase or decrease the arc index.

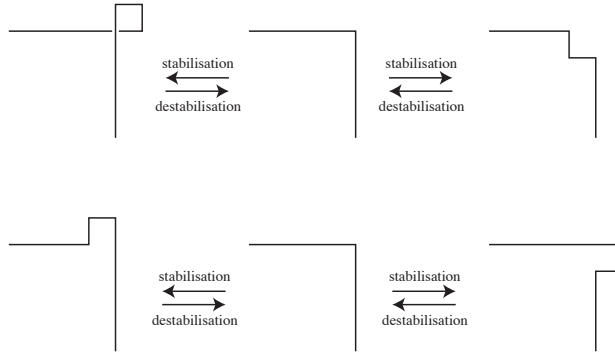


FIGURE 12. Stabilisations and destabilisations

- (3) *Exchange move*: This moves two horizontal (or two vertical) arcs of the diagram past each other. The arcs are required to have adjacent vertical heights, which means that no other arc has vertical height lying between them. Also, the horizontal endpoints of these arcs must not be interleaved.

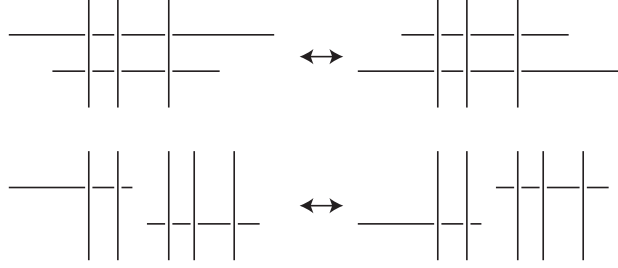


FIGURE 13. Exchange moves

Cromwell showed in [6] that any two diagrams of a link differ by a sequence of these moves. In the case where the link K is unknot, Dynnikov [8] proved that the striking result that stabilisations (which are the only moves that increase the arc index) are not required to reduce a rectangular diagram to the diagram with arc index 2.

5.5. Upper bounds on Reidemeister moves. It is clear that the above modifications to a rectangular diagram can be achieved using Reidemeister moves, and it is unsurprising that one can obtain explicit upper bounds. A stabilisation or a destabilisation can obviously be achieved using at most one Reidemeister move. Exchange moves and cyclic permutations were examined in [17], where the following bounds were established.

Lemma 5.2. *Let n be the arc index of a rectangular diagram D . Suppose that an exchange move is performed and let D' be the resulting rectangular diagram. Then D' and D differ by a sequence of at most n Reidemeister moves.*

Lemma 5.3. *Let n be the arc index of a rectangular diagram D . Suppose that a cyclic permutation is performed on the vertical (or horizontal) arcs giving a rectangular diagram D' . Then D and D' differ by a sequence of at most $(n - 1)^2$ Reidemeister moves.*

5.6. Generalised exchange moves. These are defined as follows.

Let $0 < s_1 < s_2 < s_3 < 2\pi$ be values of ϕ disjoint from the vertices of the link K . Let $0 \leq t_1 < t_2 < 2\pi$ be values of θ disjoint from the arcs of K . Suppose that, for each horizontal arc $[s, s'] \times \{t\}$ of the rectangular diagram, the following hold:

- (1) If $t \in (t_1, t_2)$, then $\{s, s'\}$ is not interleaved with $\{s_2, s_3\}$.
- (2) If $t \in S_\phi^1 - (t_1, t_2)$, then $\{s, s'\}$ is not interleaved with $\{s_1, s_2\}$.

Then one can change the rectangular diagram by shifting the ϕ value of all vertices between s_1 and s_2 so that they lie between s_2 and s_3 in the same order, and by shifting the ϕ value of all vertices between s_2 and s_3 so that they lie between s_1 and s_2 in the same order. This is a *generalised exchange move*.

The effect on the rectangular diagram is shown in the case where $t_1 = 0$ in Figure 14.

The following was Lemma 2.4 in [17].

Lemma 5.4. *Let n be the arc index of an arc presentation of K . A generalised exchange move on this arc presentation is a composition of at most $(3/2)n^3$ Reidemeister moves. It is also a composition of at most n cyclic permutations and at most $(3/4)n^2$ exchange moves.*

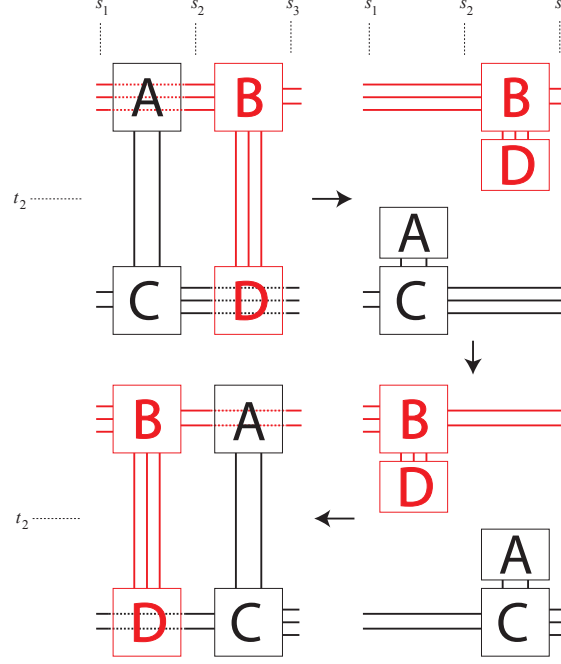


FIGURE 14. A generalised exchange move

5.7. Nearly admissible form. We fix S^3 as the join $S_\phi^1 * S_\theta^1$. Let S be a surface embedded in S^3 . Then $S - S_\phi^1$ inherits a singular foliation F with tangent space at each point given by the kernel of the 1-form $d\theta$.

We say that S is in *nearly admissible form* if it satisfies the following:

- (1) The boundary ∂S is in an arc presentation.
- (2) The surface S is smoothly embedded, except at $\partial S \cap S_\phi^1$.
- (3) $S - \partial S$ intersects S_ϕ^1 transversely at finitely many points.
- (4) The foliation F has only finitely many singularities, each of which is a point of tangency with some page.
- (5) Each singularity is of Morse type, in other words, a local maximum, a local minimum, an interior saddle or a boundary saddle. (See Figure 15). Note that local minima and local maxima do not occur on ∂S , since ∂S is in an arc presentation.
- (6) The foliation F is radial near each vertex of S .
- (7) Each page contains at most one arc of ∂S or one singularity of $F|_{S-\partial S}$ but not both.
- (8) Each arc of ∂S contains at most one singularity of F in its interior (which is a boundary saddle as shown in Figure 15 (e)).

Near a singular point of F or vertex of S , the foliation has one of the following local pictures shown in Figure 15.

In (a), the interior of S intersects the binding circle transversely. This is an *interior vertex*. In (b), a local maximum or minimum is shown, which is known as a *pole*. In (c), there is an *interior saddle*. A *boundary vertex* is shown in (d),

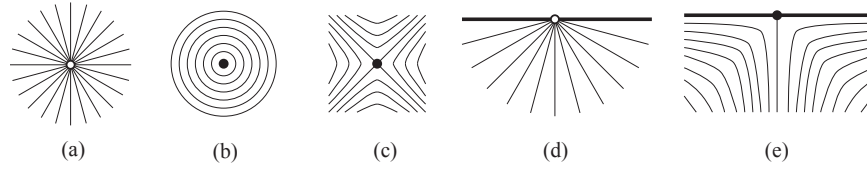


FIGURE 15. Singularities of the foliation

which is where the boundary of S intersects the binding circle. A *boundary saddle* is given in (e). When a singular point is an interior saddle or a boundary saddle, we call it a *saddle*.

An elementary general position argument gives that any embedded surface, with boundary that is in an arc presentation, may be isotoped into nearly admissible form, keeping its boundary fixed.

There are various arrangements for the surface S that are closely related to nearly admissible form. These were introduced by Dynnikov [8] and the author [17]. We briefly recall these below.

5.8. Admissible form and winding vertices. Let v be a boundary vertex of S . Pick a properly embedded arc α in S that lies within small regular neighbourhood of v and with endpoints on each side of v . Then the *winding angle* of v is $|\int_{\alpha} d\theta|$. We say that v is a *winding vertex* if its winding angle is more than 2π .

Admissible form was defined by Dynnikov [8]. The surface is required to be nearly admissible, but also to satisfy some extra conditions which constrain the behaviour near ∂S . Specifically, there can be at most one winding vertex. There can also be at most one boundary saddle. Moreover, if there is both a winding vertex and a boundary saddle, then these are required to lie in the same arc of ∂S .

Dynnikov shows in [8] that if S is an embedded surface and ∂S is connected and in an arc presentation, then there is an ambient isotopy keeping ∂S fixed taking S into admissible form.

5.9. Alternative admissible form. A surface S is in *alternative admissible form* if it is nearly admissible and it contains no winding vertices.

This was defined by the author in [17]. It has some advantages over admissible form and some disadvantages. The main advantage is the absence of winding vertices, which can be problematic. The main disadvantage is that it need not be the case that a surface can be isotoped into alternative admissible form keeping its boundary fixed. This is because the constraints that there are no winding vertices and that there is at most one boundary saddle per arc of ∂S impose restrictions on the framing of ∂S .

5.10. Generalised admissible form. It will be convenient to consider a further variation on the above theme, where we permit the surface to have singularities that are more general than Morse singularities.

We consider \mathbb{R}^3 with height function z . We say that a surface embedded in \mathbb{R}^3 has a *generalised interior saddle* at the origin if when using cylindrical polar coordinates near the origin, it has the following form:

$$\{(r, \alpha, z) : z = r^2 \sin(k\alpha)\}$$

for some integer $k \geq 2$. The $k = 2$ is the case of a Morse saddle.

We also consider surfaces embedded in the half space $\{0 \leq \alpha \leq \pi\}$ (which in Cartesian coordinates is just $\{y \geq 0\}$). We say that the surface has a *generalised boundary saddle* at the origin if it has the form

$$\{(r, \alpha, z) : z = r^2 \sin(k\alpha)\} \cap \{0 \leq \alpha \leq \pi\}$$

for some integer $k \geq 2$. Thus, at a generalised boundary saddle, the boundary of the surface is level with respect to the height function z . A *generalised saddle* is a generalised interior saddle or a generalised boundary saddle.

We say that a surface is in *generalised admissible form* if it satisfies (1)-(7) in the definition of nearly admissible form, except that generalised interior saddles and generalised boundary saddles are permitted, and it has no winding vertices.

The singular foliation near a generalised interior saddle and a generalised boundary saddle is shown in Figure 16.

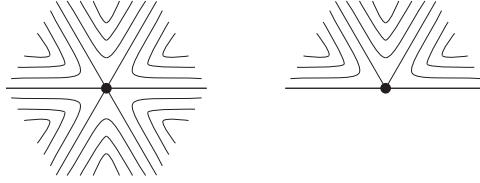


FIGURE 16. Generalised interior saddle and generalised boundary saddle (in the case $k = 3$)

5.11. Nearly embedded surfaces. In this paper, it will be necessary to consider surfaces that are not quite embedded. We say that a surface S in a 3-manifold M is *nearly embedded* if $S \cap \partial M = \partial S$ and $S - \partial S$ is embedded.

Thus, the boundary of a nearly embedded surface is required to lie in ∂M but it may not be embedded there. However, if one were to remove a very thin regular neighbourhood of ∂M , the intersection between S and the boundary of the resulting 3-manifold would be a collection of embedded curves. Hence, although ∂S is not necessarily embedded in ∂M , it is possible to perform a small homotopy to ∂S in ∂M to make it embedded.

We will consider nearly embedded surfaces that are in alternative admissible form. In this situation, we will require that ∂S is a union of fibres, and so it can be viewed as being in an arc presentation. However, because ∂S is not necessarily embedded, its image will not necessarily be a union of disjoint simple closed curves.

5.12. The binding weight of surfaces. For an embedded surface S in nearly admissible form, its *binding weight* $w_\beta(S)$ is $|S \cap S_\phi^1|$. However, when S is nearly embedded, we use a different definition. A nearly embedded surface is given by a map $f: S \rightarrow S^3$. The *binding weight* $w_\beta(S)$ of S is $|f^{-1}(S_\phi^1)|$. Thus, when ∂S runs over a vertex several times, this vertex is counted with multiplicity.

5.13. Separatrices and tiles. All the above configurations for a surface share various common features, which we now describe.

If one removes the vertices and generalised saddles, the result is a foliation of a subsurface of S .

A *separatrix* is a component of a leaf of this foliation that is incident to a generalised saddle. Since every generalised saddle has a different value of θ , the separatrices do not join distinct generalised saddles. They therefore run from a generalised saddle to a vertex or from a generalised saddle back to itself.

If one removes the singularities, the separatrices and ∂S from S , each component of the resulting surface is called a *tile*.

Each tile has a foliation, and since it is non-singular and defined by a global function θ , this foliation induces a product structure on the tile. Therefore, each tile is an open annulus or open disc. The separatrices in the boundary of a disc tile run from a generalised saddle to a vertex. Hence, disc tiles come in three types (see Figure 17):

- (1) *square* tiles, which have two vertices in their boundary (which may be identified) and two generalised saddles;
- (2) *half* tiles, which have two vertices in their boundary (which may again be identified) and one generalised saddle; the arc in the boundary of the half tile joining the vertices lies in ∂S ;
- (3) *bigon* tiles, which have two vertices in their boundary and no generalised saddles. In this case, the bigon tile is a component of S .

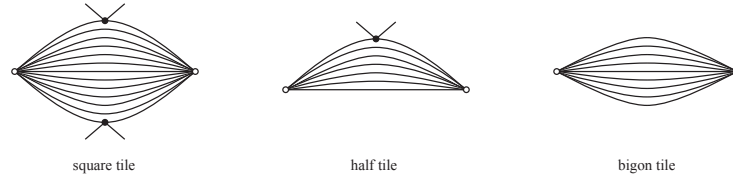


FIGURE 17. Disc tiles

It is frequently the case that annular tiles can be removed without increasing the binding weight of the surface, or can be otherwise avoided. Thus, we will frequently make the hypothesis that a given nearly admissible or generalised admissible surface contains no annular tiles.

5.14. Relating the number of saddles and vertices.

Lemma 5.5. *Let S be a generalised admissible surface with no annular tiles. Let $v^i(S)$ denote the number of interior vertices, let $v^b(S)$ be the number of boundary vertices, let $x^i(S)$ be the number of generalised interior saddles, and let $x^b(S)$ denote the number of generalised boundary saddles. Then*

$$2v^i(S) + v^b(S) - 2x^i(S) - x^b(S) \geq 2\chi(S).$$

Proof. We consider first the case where S is closed. The singular foliation can be used to create a cell structure on S , as follows. The 0-cells are the vertices. Within each tile, we pick a non-singular leaf, and declare it to be a 1-cell. The complement of this 1-skeleton is a collection of discs, each of which forms an open regular neighbourhood of a generalised saddle. Here, we are using the assumption that there are no annular tiles. Let V , E and F be the number of 0-cells, 1-cells and 2-cells of this cell structure. Then $4F \leq 2E$, because each 2-cell is at least 4-valent. Hence,

$$\chi(S) = V - E + F \leq V - F = v^i(S) - x^i(S),$$

proving the lemma in this case.

Note that when a component of S contains a bigon tile, then that tile is all of the component. The required formula is readily verified in this situation. So we may assume that S contains no bigon tiles.

When S has non-empty boundary, we form its double \hat{S} . The singular foliation on S induces one on \hat{S} . Since ∂S is in an arc presentation and S has no bigon tiles, no annular tiles on \hat{S} are formed. Thus, we may create a cell structure on \hat{S} , exactly as in the closed case. Let V , E and F be its number of 0-cells, 1-cells and 2-cells. Again each 2-cell is at least 4-valent, because it arises either from an interior generalised saddle of S , or from two generalised boundary saddles fused together. Hence,

$$2\chi(S) = \chi(\hat{S}) = V - E + F \leq V - F = 2v^i(S) + v^b(S) - 2x^i(S) - x^b(S).$$

□

6. ADMISSIBLE PARTIAL HIERARCHIES

6.1. Definition. We are given a link K in an arc presentation. This arc presentation corresponds to a rectangular diagram for K .

We give the exterior of K a boundary pattern as described in Theorem 3.3. On each component of $\partial N(K)$, this pattern is either empty or single curve with slope 0 or ± 1 . We stabilise the arc presentation enough times so that for each component of $\partial N(K)$ with non-empty boundary pattern, the slope of this boundary pattern is equal to the writhe of the relevant component of K in its rectangular diagram.

Let S_1, \dots, S_n be a partial hierarchy for the exterior K . We set S_0 to be $\partial N(K)$. Thus, although S_0 is not properly embedded in the ambient manifold $S^3 \setminus N(K)$, we can still think of it as, in some sense, the first surface of the hierarchy. Because of the above stabilisations, the initial boundary pattern can be realised as a union of fibres in the singular foliation on $\partial N(K)$.

We say that S_0, \dots, S_n is in *admissible form* if each S_i is an embedded nearly admissible surface with no annular tiles and no winding vertices, and ∂S_i avoids all the saddles in the earlier surfaces. Note that we require each S_i to be embedded, rather than just nearly embedded. The *binding weight* of this partial hierarchy is $\sum_{i=0}^n w_\beta(S_i)$.

Let $M = M_{n+1}$ be the final manifold in the partial hierarchy arising from S_1, \dots, S_n . It inherits a boundary pattern $P = P_{n+1}$ which is in an arc presentation. Each component of $\partial M_{n+1} \setminus P_{n+1}$ lies in some S_i and so is in nearly admissible form.

6.2. A polyhedral decomposition for an admissible partial hierarchy. Let H be the partial hierarchy in admissible form for the exterior of the link K . We now explain how to construct a polyhedral decomposition Δ of the exterior of K that contains H as a subcomplex. Note that the boundary of M is comprised of parts of H and so inherits the singular foliation of H .

The polyhedral decomposition will be an expression of the exterior of K as a union of polyhedra with some faces identified in pairs. The polyhedra that we will use will be *expanded tetrahedra*, *expanded turnovers*, and *rugby balls*, which are defined as follows.

A *bigon* is a cell complex, consisting of two 0-cells, two 1-cells, and a 2-cell. The 1-cells are each attached to both 0-cells, so that the 1-skeleton is a circle. The boundary of the 2-cell is attached homeomorphically onto the 1-skeleton.

An *expanded tetrahedron* is a 3-ball that is obtained from a solid tetrahedron by replacing some of its edges by bigons.

A *turnover* is a cell complex, consisting of three 0-cells, three 1-cells, two 2-cells and a 3-cell. The 1-cells are attached to the 0-cells so that the 1-skeleton is a circle. The boundary of each 2-cell is attached homeomorphically onto this circle. The 2-skeleton is then a 2-sphere, to which the boundary of the 3-cell is attached homeomorphically. An *expanded turnover* is a 3-ball that is obtained from a turnover by replacing some 1-cells by bigons.

A *rugby ball of order n* is a cell complex consisting of two 0-cells, n 1-cells, n 2-cells and a 3-cell. Each 1-cell joins the two 0-cells. Each 2-cell is a bigon. The 2-skeleton forms a 2-sphere to which the boundary of the 3-cell is attached homeomorphically.

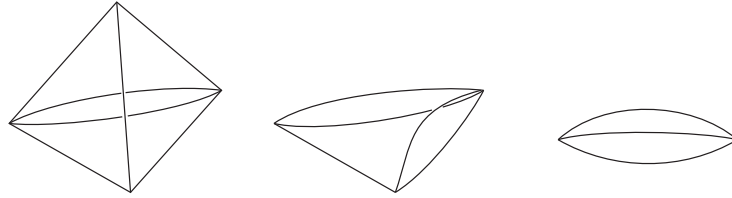


FIGURE 18. The polyhedra used in the polyhedral decomposition. Left: expanded tetrahedron. Middle: expanded turnover. Right: rugby ball of order 3.

The binding circle intersects $H \cup N(K)$ in a finite collection of points and arcs. Cutting the binding circle at these points and arcs gives a union of closed arcs. We declare that each of these is a 1-cell of Δ .

The complement of the binding circle is foliated by pages \mathcal{D}_t , where $t \in \mathbb{R}/(2\pi\mathbb{Z})$. We say that a page \mathcal{D}_t contains an *event* if it has non-empty intersection with the boundary of some surface of the partial hierarchy, or some Morse singularity of some surface in the partial hierarchy. There are only finitely many events, and in each component of the complement of these events, the intersections $\mathcal{D}_t \cap H$ are isotopic as t varies. Between the event pages, we will also pick a finite collection of pages and these, together with the event pages, will be called *specified pages*.

Consider an event page that contains an interior saddle of a surface in H . Then just below this point, two arcs of the surface come together, forming the interior saddle, and then just above the interior saddle, they separate into two different arcs. We pick a page just below and a page just above the interior saddle, and we declare that these are specified pages. Each of these pages intersects the exterior of H in a collection of discs. We triangulate each of these discs by coning from a point on the boundary, unless a disc is a bigon, in which case we leave it untouched. We ensure that the arcs involved in the interior saddle move are edges of the polyhedral decomposition. These pages are now decomposed into triangles and isolated bigons. Despite the possible existence of bigons, we call this a ‘triangulation’.

Now suppose that the event page contains an arc α in the boundary of one of the surfaces S_j in H . This runs over a previous surface S_i in the hierarchy. By

assumption, this arc in S_i misses its interior saddles and boundary saddles. Hence, near α , one part of $S_i \setminus \alpha$ lies just above α and the other part lies just below α . There are two possible arrangements for S_j : either α contains a boundary saddle of S_j or it does not.

Let us first consider the case where α does not contain a boundary saddle of S_j . Then near α , S_j lies entirely above α or entirely below it. We have three sheets of H emanating from this arc. Say that one sheet lies below the arc and two sheets lie above. We declare that a page just above and a page just below the event page are specified pages. Between the higher page and the event page, we insert a rugby ball of order 3. The event page intersects the exterior of H in discs. Pick a triangulation of these discs by coning from points on the boundary. This specifies triangulations in the two nearby specified pages. Between these triangulations, we insert expanded turnovers and expanded bigons. We will consider later the case where α contains a boundary saddle.

When two event pages are adjacent, the discs in the specified page just below the higher event page and in the specified page just above the lower event page have been triangulated. These triangulations need not agree. We can change one triangulation into the other using 2-2 Pachner moves. Pick such a sequence of 2-2 Pachner moves that is as short as possible. We insert new specified pages, the number being one less than the number of Pachner moves. We triangulate these as indicated by the Pachner moves. Each of these moves can be realised by the addition of an expanded tetrahedron. Two of the triangular faces of the expanded tetrahedron lie in a specified page, as do the other two triangular faces. Between these, the expanded tetrahedron has four bigons. For each of the triangles not involved in the 2-2 Pachner move, we insert an expanded turnover.

We now need to explain how to triangulate the space between the pages just below an interior saddle and just above. The triangulations on these pages mostly agree. Where they do agree, we insert expanded turnovers. The remainder lies near the separatrices emanating from the interior saddle. We declare that the interior saddle point is a vertex of Δ , and that the four separatrices emanating from it are edges of Δ . The interior saddle is contained in some surface of the hierarchy H , and near this interior saddle, there is a square in the surface bounded by edges of Δ lying in the adjacent pages. This square is divided into 4 triangles by the interior saddle and its incident separatrices. We declare that each of these triangles is a face of Δ . The 4 separatrices lie in a page, and this has also been decomposed by other edges of Δ . In the complement of these edges, there are 4 triangles that are incident to the separatrices. We also declare that these 4 triangles are faces of Δ . We attach polyhedra of Δ onto these 4 faces. Four of these polyhedra are expanded tetrahedra, and four are expanded turnovers. See Figure 19.

When there is a boundary saddle in some S_j , Δ is defined similarly. The boundary saddle is a vertex of Δ . The three separatrices in S_j emanating from it are edges of Δ . On each side of S_j , we insert an expanded tetrahedron and an expanded turnover.

Note that by construction, the boundary pattern P for the exterior of H is a subcomplex of the 1-skeleton of the polyhedral decomposition. This is because the boundary pattern is a union of fibres, each of which is part of the boundary of a surface in the hierarchy. This therefore lies in an event page and becomes an edge of Δ .

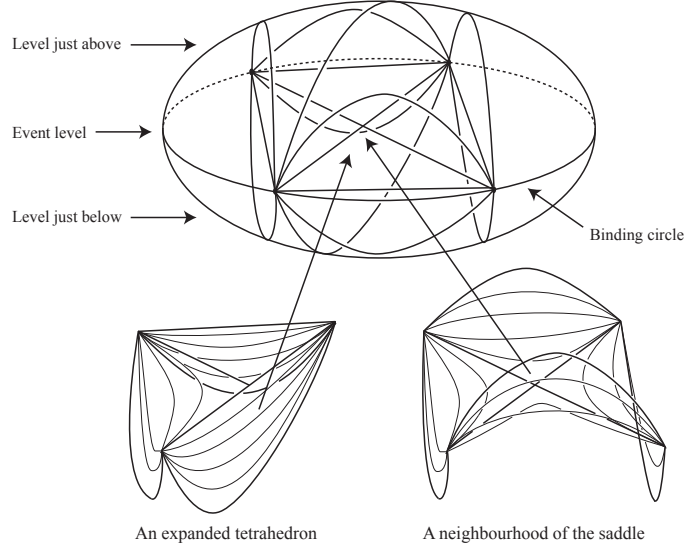


FIGURE 19. An expanded tetrahedron lying near an interior saddle

The hierarchy H is a subcomplex of Δ , and hence the exterior of H inherits a polyhedral decomposition Δ_M .

6.3. A triangulation for an admissible partial hierarchy. In the previous subsection, we defined a polyhedral decomposition Δ for the exterior of the link, with the admissible partial hierarchy H as a subcomplex. In this section, we show that this can be collapsed to a triangulation \mathcal{T} . We collapse each rugby ball to a single edge, in such a way that each edge of the rugby ball maps homeomorphically onto the new edge. We collapse each expanded turnover to a single triangular face. Again, each edge of the expanded turnover is mapped homeomorphically to an edge of the triangle. We collapse each expanded tetrahedron to a tetrahedron. This gives a triangulation \mathcal{T} for the link exterior.

Note that, in principle, this sequence of collapses might create a topological space that is not a 3-manifold. For example, suppose that a sequence of expanded turnovers were glued together in a circular fashion, so that they are indexed by the integers modulo n , and so that the expander turnover with index i is glued the one with index $i-1$ and the one with index $i+1$ along their triangular faces. Then when these expanded turnovers are collapsed, they are collapsed to a single triangle, and hence points in this triangle are not 3-manifold points. However, it is easy to check that in our situation, we do obtain a triangulation for the link exterior.

By performing the same collapses to Δ_M , we obtain a triangulation \mathcal{T}_M for M .

6.4. Bounding the number of tetrahedra.

Lemma 6.1. *Let $H = \{S_0, S_1, \dots, S_n\}$ be an admissible partial hierarchy for the exterior of K . Let w be its binding weight. Then the total number of event pages is at most $2(w - \sum_{i=1}^n \chi(S_i))$.*

Proof. An event page occurs when some surface S_i has an interior saddle or when its boundary has an arc in a page. The number of interior saddles of the surface

S_i is $x^i(S_i)$, which is at most $2x^i(S_i) + x^b(S_i)$. By Lemma 5.5, this is at most $2v^i(S_i) + v^b(S_i) - 2\chi(S_i)$. So the sum of the number of interior saddles and boundary vertices is at most $2(v^i(S_i) + v^b(S_i) - \chi(S_i))$. Summing this over i gives the required upper bound on the number of event pages. \square

Lemma 6.2. *The total number of specified pages is at most $4w(w - \sum_{i=1}^n \chi(S_i))$.*

Proof. Consider a specified page just below or just above an event page. This intersects the exterior of H in discs. Some of these are bigons which are not triangulated, and some of these are triangles. The remainder are polygons with at least four edges in their boundary. Each edge of each polygon is either an arc in the binding circle or an arc in the interior of the page. There are at most w of the former arcs. There are at most $w/2$ arcs in the interior of the page, and each contributes at most two edges to the boundaries of the polygons. So the polygons have a total of at most $2w$ sides. Each of these discs has been triangulated by coning from a vertex in its boundary. Now, the number of 2-2 Pachner moves relating two such triangulations of a polygon with n sides is at most $n - 2$. Hence, the number of 2-2 Pachner moves needed to relate the two triangulated specified pages is at most $2w - 2$. So, the number of specified pages strictly between two adjacent event pages is at most $2w - 1$. Adding the number of event pages gives the required bound. \square

Lemma 6.3. *The number of tetrahedra in the triangulation \mathcal{T} of the link exterior is at most $16w(w - \sum_{i=1}^n \chi(S_i))$. This is also an upper bound for the number of tetrahedra in \mathcal{T}_M .*

Proof. The tetrahedra in the triangulations \mathcal{T} and \mathcal{T}_M come from expanded tetrahedra of the polyhedral decompositions Δ and Δ_M . We observe that between two specified pages, neither of which are event pages, there is at most one expanded tetrahedron. Between any event page and an adjacent specified page, there are at most four expanded tetrahedra. The result now follows from Lemma 6.2. \square

Lemma 6.4. *The number of expanded turnovers in the polyhedral decomposition Δ is at most $8w^2(w - \sum_{i=1}^n \chi(S_i))$.*

Proof. Between adjacent specified pages, there are at most $2w$ expanded turnovers. Now apply Lemma 6.2. \square

Lemma 6.5. *Each edge of \mathcal{T} has pre-image in Δ consisting of at most $8w(w - \sum_{i=1}^n \chi(S_i))$ edges. Similarly, each edge in \mathcal{T}_M has pre-image in Δ_M consisting of at most $8w(w - \sum_{i=1}^n \chi(S_i))$ edges.*

Proof. The edges of \mathcal{T} are obtained by collapsing bigons of Δ . Within each page, each edge is incident to at most one bigon. In adjacent specified pages, edges of Δ may be joined by bigons. So the number of edges of Δ collapsed to a single edge of \mathcal{T} is at most twice the number of specified pages. Now apply Lemma 6.2. The same argument also applies to the edges in \mathcal{T}_M . \square

Lemma 6.6. *For each point of intersection between ∂M and the binding circle, the number of edges of Δ_M that are incident to that vertex is at most $4w^2(w - \sum_{i=1}^n \chi(S_i))$. This is also an upper bound for the number of edges of \mathcal{T}_M that are incident to a vertex in ∂M .*

Proof. We see from the construction of Δ in Section 6.2 that each edge lies in a specified page. It runs from a vertex of H either to another vertex or to a saddle. Therefore, in any page, the number of edges running from a given vertex is at most w , the binding weight of the partial hierarchy. So, the total number of edges of Δ_M emanating from a given vertex of H is at most w times the number of specified pages, hence at most $4w^2(w - \sum_{i=1}^n \chi(S_i))$ by Lemma 6.2. When Δ_M is collapsed to form \mathcal{T}_M , several edges may be combined to a single edge, but the number of edges emanating from a given vertex in ∂M does not go up. \square

6.5. A handle structure for the exterior of the partial hierarchy. Let M be the exterior of the partial hierarchy H . We can attach a partial collar onto ∂M , as follows. For each edge of Δ_M in ∂M , attach a bigon along one of the bigon's edges. For each triangle of Δ_M in ∂M , attach on an expanded turnover with three bigons in its boundary. Identify each of these bigons with the bigon incident to the relevant edge in ∂M .

Let $\hat{\Delta}_M$ be this new polyhedral decomposition of M , and let \mathcal{K} be the dual handle structure. Recall that this has an i -handle for each $(3 - i)$ -cell of $\hat{\Delta}_M$ that does not lie wholly in ∂M .

For example, dual to the expanded tetrahedron shown in Figure 19, there is 0-handle of \mathcal{K} . Its intersection with the 1-handles and 2-handles is shown in Figure 21.

Suppose that S is a surface properly embedded in M that is in general position with respect to \mathcal{T}_M . The inverse image of S is a surface in Δ_M , which we will call S' . This also gives a surface \hat{S} that is standard in the handle structure \mathcal{K} .

Lemma 6.7. *Let S be a surface properly embedded in M that is in general position with respect to \mathcal{T}_M , and let \hat{S} be the resulting surface in \mathcal{K} . Then the extended weight of \hat{S} is at most $40w(w - \sum_i \chi(S_i))w(S)$, where w is the binding weight of H and $w(S)$ is the weight of S in \mathcal{T}_M .*

Proof. The polyhedral decomposition Δ_M collapses to \mathcal{T}_M . By Lemma 6.5, each edge of \mathcal{T}_M has pre-image in Δ_M consisting of at most $8w(w - \sum_{i=1}^n \chi(S_i))$ edges. So, the inverse image of S in Δ_M has weight at most $8w(w - \sum_{i=1}^n \chi(S_i))w(S)$. There is a 1-1 correspondence between the edges of Δ_M not lying in ∂M and the 2-handles of \mathcal{K} . Hence, $8w(w - \sum_{i=1}^n \chi(S_i))w(S)$ is an upper bound for the weight of \hat{S} in \mathcal{K} .

Each arc of intersection between S and a face of \mathcal{T}_M has both its endpoints on edges of \mathcal{T}_M . The same is true of the inverse image of S in Δ_M . Hence, each component of intersection between \hat{S} and a 1-handle of \mathcal{K} is incident to a 2-handle of \mathcal{K} . Each 2-handle of \mathcal{K} runs over at most four 1-handles of \mathcal{K} . Hence, the number of intersections between \hat{S} and the 1-handles of \mathcal{K} is at most 4 times the weight of \hat{S} . Therefore, the extended weight of \hat{S} in \mathcal{K} is at most $40w(w - \sum_{i=1}^n \chi(S_i))w(S)$, as required. \square

6.6. Making a surface transverse to P . We are considering a partial hierarchy H in admissible form. We also suppose that it is *well-spaced*, which means that it satisfies all the following conditions:

- (1) the closure of the union of tiles in H adjacent to P is a regular neighbourhood $N(P)$ of P ;

- (2) for each vertex s of H that lies on P and for each annular component A' of $N(P) \setminus P$ containing s , there is a tile in A' incident to s such that the other vertex of the tile lies on $\partial A' - P$;
- (3) the above vertices include all the vertices lying in A' .

Section 7.2 describes how to make a hierarchy well-spaced, and Figures 25 and 26 show examples of the singular foliations in a well-spaced hierarchy.

We have defined a polyhedral decomposition Δ and a triangulation \mathcal{T} for the link exterior. These restrict to a polyhedral decomposition Δ_M and a triangulation \mathcal{T}_M for the exterior M of H . Suppose that S is a surface properly embedded in M that is normal with respect to \mathcal{T}_M . In this subsection, we will explain how to pattern-isotope ∂S so that afterwards it intersects the annuli $N(P) \setminus P$ in a collection of essential arcs, each of which intersects P transversely at a single point away from the vertices of P .

Let A' be a component of $N(P) \setminus P$, which is an annulus. By assumption, for each vertex s of H that lies on $A' \cap P$, there is a tile in A' incident to s such that the other vertex of the tile lies on $\partial A' - P$. This tile is a union of faces of \mathcal{T} . Each such face has corners that are vertices of H or a saddle, and there is at most one corner of the latter type. Hence, there is some edge of \mathcal{T} running across the tile from s to the other vertex. Pick one such edge for each vertex of H on $P \cap A'$. We may choose a product structure $S^1 \times [0, 1]$ on A' so that each of these edges is of the form $\{*\} \times [0, 1]$ for some point $\{*\}$ on S^1 . Pick a very thin regular neighbourhood A'' of $A' \cap P$ in A' that also intersects these edges vertically. There is an isotopy of ∂M that expands A'' to all of A' , and that respects the product structures on A'' and A' . Now consider the curves ∂S . We may assume that ∂S intersects A'' in a collection of vertical arcs. Hence, when we expand A'' to fill A' , afterwards the intersection $\partial S \cap A'$ is a union of vertical arcs in A' , as required. At certain times during the isotopy, the image of ∂S passes through a vertex on $\partial A' - P$. The number of times this happens is equal to the number of intersections between the initial position of ∂S and the edges running from the vertices on $\partial A' - P$ to P . Hence, the number of times this happens is at most $w(\partial S)$. We can realise the isotopy of ∂S as an isotopy of S supported in a collar neighbourhood of ∂M . At each moment in time that ∂S passes through a vertex on ∂M , we obtain new intersection points between S and the edges of \mathcal{T}_M . By Lemma 6.6, the total number of these intersection points is at most $4w^2(w - \sum_{i=1}^n \chi(S_i))w(\partial S)$.

6.7. Homotoping S into alternative admissible form. As in the previous subsection, suppose that the partial hierarchy H is in admissible form and is well-spaced. Again suppose that S is a surface properly embedded in M that is normal with respect to \mathcal{T}_M . We also suppose that S is incompressible and boundary-incompressible with no component that is a sphere or disc intersecting P at most twice.

We pattern-isotoped S so ∂S intersects each component of $N(P) \setminus P$ in a collection of essential arcs, each of which intersects P transversely at a single point away from the vertices of P . This isotopy need not have kept S normal in \mathcal{T}_M .

The boundary of S need not be in an arc presentation, in other words, a concatenation of properly embedded arcs in pages. We will homotope ∂S to achieve this. The homotopy runs from time 0 to time 1, say. For all times in the interval $[0, 1)$, the curves remain embedded. But at time 1, this need not be the case. This

homotopy therefore specifies a modification to S near the boundary, taking S to a nearly embedded surface. In fact, S will be embedded away from the binding circle.

Recall that associated with the surface S in \mathcal{T}_M , there are surfaces S' and \hat{S} in Δ_M and \mathcal{K} . The homotopy that we will perform to S will also induce a homotopy to S' and \hat{S} . We start by describing the homotopy to S .

Since the hierarchy H is in admissible form, ∂M inherits the structure of a union of tiles. The boundary pattern P is a union of fibres, each running between vertices. We now homotope ∂S , via a homotopy supported away from P , as follows. In each triangle of \mathcal{T}_M in ∂M , we homotope ∂S near the edges incident to a saddle. We homotope its points of intersection along that edge away from the saddle until they run through the vertex at the other end of the edge. (See Figure 20.)

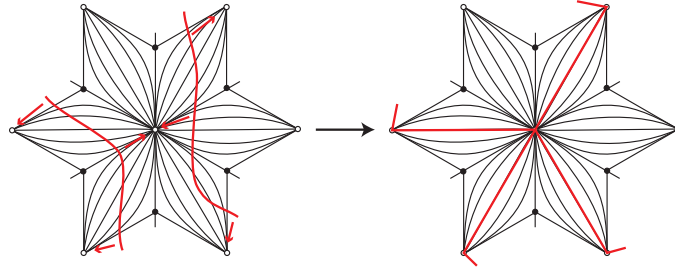


FIGURE 20. Homotoping the boundary curves of a surface

The reason that we do not require ∂S to be embedded is as follows. The surface S is constructed using normal surface theory, and all we have is an exponential upper bound on its weight. Hence, ∂S may run over the tiles in $S_0 \cup \dots \cup S_n$ many times. So it might not be possible to isotope ∂S to become a union of fibres and preserve the property that it is embedded, without somehow introducing many more vertices to $S_0 \cup \dots \cup S_n$. This runs counter to our main goal, which is to control the binding weight of the hierarchy.

This homotopy creates the surfaces S' in $\hat{\Delta}_M$ and \hat{S} in \mathcal{K} . The latter can also be viewed as a properly embedded surface in the handle structure \mathcal{K} . Moreover, the homotopy that we have applied to \hat{S} is actually an isotopy. This is because in \mathcal{K} , we have removed a regular neighbourhood of the vertices on the binding circle, and away from the binding circle, the surface really is embedded. An example of this isotopy is shown in Figure 21. Note that after this isotopy, the surface is disjoint from the 2-handles that are dual to the edges incident to saddles.

This surface \hat{S} can easily be made standard with respect to \mathcal{K} , but it is not necessarily normal. However, a further pattern-isotopy makes \hat{S} normal with respect to \mathcal{K} , by Lemma 4.2. This does not increase the extended weight of the surface.

Note also that this isotopy does not increase the number of components of intersection with any 2-handle. Hence, the resulting surface is also disjoint from the 2-handles that are dual to all the edges incident to saddles. This imposes the following restrictions on how $\partial \hat{S}$ runs over ∂M . Consider any face F of a polyhedron of Δ_M that lies in ∂M . Dual to this polyhedron is a 0-handle of \mathcal{K} and dual to the face is a 1-handle incident to the 0-handle, since we enlarged Δ_M to $\hat{\Delta}_M$. In the case where F is triangular, it has three vertices, one of which is a saddle and the other two being on the binding circle. The 1-handle of \mathcal{K} dual to F has three

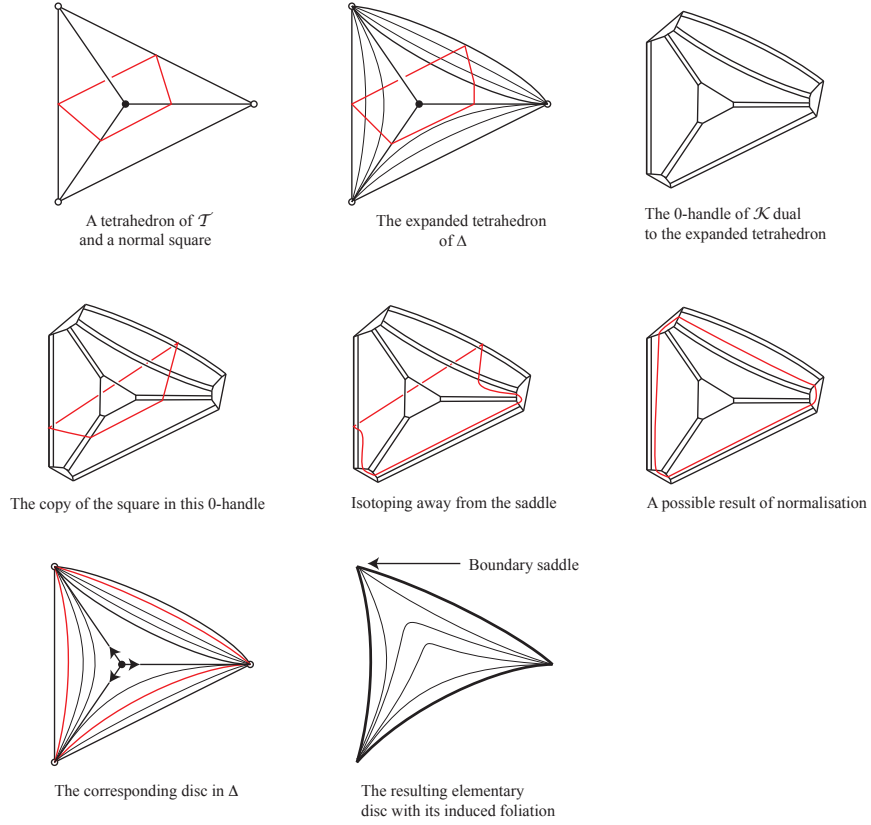


FIGURE 21. Isotoping the boundary curves of a surface

2-handles incident to it. Two of these are dual to edges of Δ_M incident to the saddle, and so \hat{S} misses these. So, up to handle-preserving isotopy, there is only one possible configuration for each component of intersection between S and the 1-handle. This corresponds to an arc in F joining the two vertices on the binding circle. In the case where F is a bigon, both of its vertices are on the binding circle. Again, there is only one possible configuration for each component of intersection between the dual 1-handle and \hat{S} .

We now summarise the result of the procedure so far.

Lemma 6.8. *Let S be a compact orientable incompressible boundary-incompressible surface properly embedded in M with no component that is a sphere or disc intersecting P at most twice. Suppose that S is normal with respect to \mathcal{T}_M . Then there is a pattern-isotopy taking S to a surface \hat{S} that is normal with respect to \mathcal{K} and with extended weight at most $200w^3(w - \sum_{i=1}^n \chi(S_i))^2w(S)$. This surface \hat{S} misses the 2-handles of \mathcal{K} that are dual to the edges incident to saddles. This also induces a surface (also called \hat{S}) in \mathcal{T}_M that is nearly embedded, that has boundary in admissible form, and that has binding weight at most $200w^3(w - \sum_{i=1}^n \chi(S_i))^2w(S)$. It misses the saddles in ∂M as well as the interior of any edge that is incident to a saddle.*

Proof. We start with the surface S that is normal with respect to \mathcal{T}_M . As described in Section 6.6, we then pattern-isotope S so afterwards it intersects each component of $N(P) \setminus P$ in a collection of essential arcs. Its weight in \mathcal{T}_M is then at most

$$w(S) + 4w^2(w - \sum_{i=1}^n \chi(S_i))w(\partial S) \leq 5w^2(w - \sum_{i=1}^n \chi(S_i))w(S).$$

By Lemma 6.7, the corresponding surface \hat{S} in \mathcal{K} has extended weight at most $200w^3(w - \sum_{i=1}^n \chi(S_i))^2w(S)$. By Lemma 4.2, we can then isotope \hat{S} to a surface that is normal with respect to \mathcal{K} without increasing its extended weight and without introducing any new intersections with 2-handles. The bound on the extended weight in \mathcal{K} then gives a bound on the binding weight of the surface in \mathcal{T}_M . \square

Thus, we have now specified, for each polyhedron of Δ_M , the intersection between its boundary and \hat{S} . In particular, the boundary of \hat{S} is in an arc presentation. We now arrange the intersection between \hat{S} and the interior of the polyhedra, using the following lemmas.

We start by examining the elementary discs of \hat{S} in \mathcal{K} . The following straightforward result classifies elementary normal discs in a rugby ball.

Lemma 6.9. *Let D be an elementary normal disc in a rugby ball with an open regular neighbourhood of its vertices removed. Then D satisfies one of the following:*

- (1) *The boundary of D is a concatenation of two arcs, each of which lies in a bigon and runs between the vertices of the rugby ball.*
- (2) *The boundary of D runs over each bigon in exactly one arc, which runs between the two edges of the bigon.*

Proof. We refer to the definition of normality in Section 4.1. By (1) in the definition, ∂D runs over the interior of each edge of the rugby ball at most once. By (2), it runs over each vertex of the rugby ball at most once. By (4), it cannot run over both a vertex and the interior of an incident edge. By (3), it cannot be comprised of an arc in a face with endpoints in the same vertex. Hence, the only possibilities are as described in the lemma. \square

Similarly, we have the following classification of normal discs in an expanded turnover.

Lemma 6.10. *Let D be an elementary disc in an expanded turnover with an open regular neighbourhood of its vertices removed. Then D satisfies one of the following:*

- (1) *The boundary of D is a concatenation of two arcs, one in a bigon and one in triangular face, each one running between distinct vertices.*
- (2) *The boundary of D runs over each bigon in at most one arc, and it runs over each triangular face in exactly one arc.*
- (3) *The boundary of D is a concatenation of three arcs, each of which runs between distinct vertices.*

Examples of these possible discs are shown in Figure 22. In (1) and (2) of each of these lemmas, the disc D can be arranged so that its singular foliation is a product foliation with no singularities. But in (3) of Lemma 6.10, it has a singular foliation with one boundary saddle (which may be placed at a vertex of the expanded turnover).

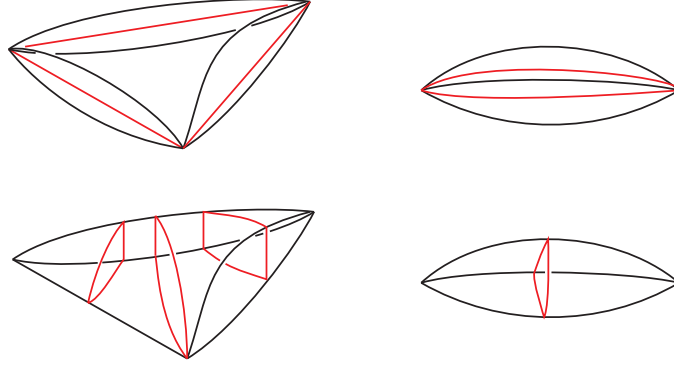


FIGURE 22. Some normal discs in expanded turnovers and in a rugby ball

Lemma 6.11. *Let B be an expanded tetrahedron of Δ_M , with an open neighbourhood of its vertices removed, that is incident to a saddle of the partial hierarchy H . Let D be an elementary normal disc in B that misses the 2-handles dual to edges incident to the saddle. Then ∂D is a concatenation of two or three arcs, each of which runs between distinct vertices of the expanded turnover. Furthermore, D can be arranged so that its singular foliation either is a product or has a single boundary saddle, at a vertex of the expanded turnover.*

Proof. The expanded tetrahedron has four triangular faces. Three of these are incident to the saddle. As shown above, when ∂D intersects one of these three faces, it does so in an arc joining the two vertices in its boundary that are not the saddle. Hence, if D intersects the fourth triangular face, it must also do so in an arc running between two distinct vertices of B . For otherwise, it runs over the interior of an edge of that triangular face, but this then forces ∂D to intersect the interior of one of the edges of the other three triangular faces, and we have ruled this out. So, ∂D is a concatenation of arcs, each of which runs between distinct vertices of B . By normality, it cannot intersect a vertex of B more than once. It also misses the vertex that is the saddle. Hence, it is a concatenation of two or three arcs. In each case, we can arrange its singular foliation to be of the specified form. \square

See Figure 21 for an example of such a normal disc. A similar case-by-case analysis gives the following result for an expanded tetrahedron of Δ_M that is not incident to a saddle of the partial hierarchy H . Recall from Section 6.2 that each triangular face of the expanded tetrahedron lies in some page, with two triangular faces lying one page, and two lying in another.

Lemma 6.12. *Let B be an expanded tetrahedron of Δ_M , with an open neighbourhood of its vertices removed, that is not incident to a saddle of the partial hierarchy H . Let D be a normal disc in B . Then D may be isotoped, keeping its boundary fixed, so that it satisfies one of the following possibilities:*

- (1) *It inherits a product foliation. Its boundary consists of an arc in the union of the top two triangular faces and an arc in the union of the bottom two triangular faces, possibly joined by arcs in the bigons.*
- (2) *It inherits a product foliation. Its boundary consists of an arc in a triangular face and an arc in bigon face, both joining distinct vertices.*

- (3) *It has a single boundary saddle. Its boundary consists of three, four or five arcs. If it has four arcs, then two of these lie in the top two triangles or the bottom two triangles and are incident. If it has five arcs, then two of these lie in the top two triangles and are incident, and two of the arcs lie in the bottom two triangles and are incident. The remaining arcs run between distinct vertices and lie in a bigon or triangular face.*
- (4) *Its boundary runs through the four bigons, with each arc joining one vertex of the bigon to the other. It inherits a singular foliation with two boundary saddles, on distinct vertices.*
- (5) *Its boundary consists of six arcs. Two of these lie in the top two triangles and are incident, and two of the arcs lie in the bottom two triangles and are incident. Each of the remaining two arcs runs between distinct vertices and lies in a bigon. This disc has two boundary saddles.*
- (6) *Its boundary runs through each of the four triangular faces and across each of the bigons. It is disjoint from the vertices and comes from a square in the collapsed triangulation. It inherits a singular foliation, with a single interior saddle and with separatrices running to each of the four bigons.*

See Figure 23 for examples of these discs.

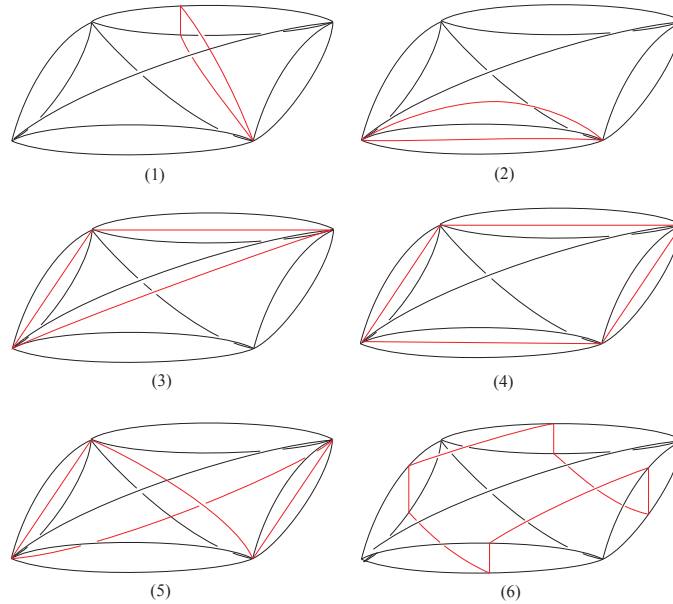


FIGURE 23. The five possibilities for elementary normal discs in an expanded tetrahedron, as described in Lemma 6.12

Lemma 6.13. *Each elementary normal disc of \hat{S} intersects at most 4 separatrices and most 4 tiles.*

Proof. Each separatrix starts at an interior saddle or boundary saddle. Consider the case of a boundary saddle first. This occurs in an elementary disc as in (3), (4) or (5) of Lemma 6.12 or as in (3) of Lemma 6.10 or as in Lemma 6.11. Therefore,

the separatrices starting at a boundary saddle lie wholly within a single elementary normal disc.

Now consider the case of an interior saddle. This lies in a disc as in (6) of Lemma 6.12, which we will call a *square*. Just above and below the event page containing the interior saddle there are two specified pages. These two pages have triangulations with no vertices of the triangulation in the interior of these pages. Note that S intersects of these discs in normal arcs. As the height of the pages increase, some of these arcs come together in interior saddles lying in squares parallel to the given square. The separatrices emanating from an interior saddle lie in a square. They then run to bigons in the boundary of the expanded tetrahedron. After that, they run through discs as in (2) of Lemma 6.10 and Lemma 6.9 until they reach a vertex on the binding circle. These latter discs do not contain any boundary saddles. So they each intersect at most one separatrix, and hence intersect at most 2 tiles. We also deduce that the only separatrices lying in squares are the ones emanating from the interior saddle. So they intersect 4 separatrices and 4 tiles. \square

7. EXTENDING THE PARTIAL HIERARCHY

7.1. Overview. Theorem 3.3 provides an exponentially controlled hierarchy for the exterior of the link K . Let S_1, \dots, S_n be the first n surfaces in this hierarchy. Let S_0 be $\partial N(K)$. Suppose that this partial hierarchy is in admissible form, and let w be its binding weight.

We now explain, in overview, how to extend this partial hierarchy.

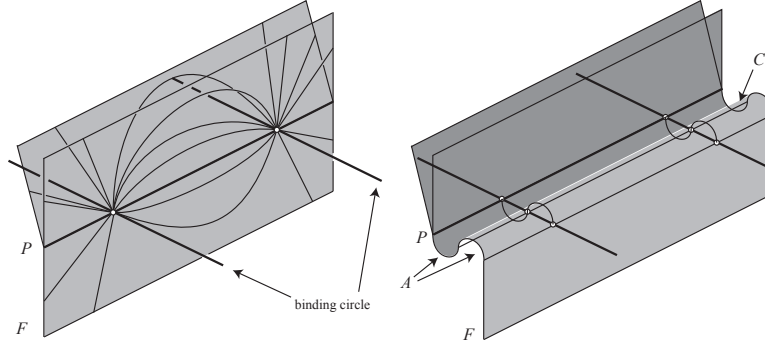
- (1) Modify the hierarchy so that it is well-spaced, as defined in Section 6.7.
- (2) Build a triangulation \mathcal{T} and a polyhedral decomposition Δ for the exterior of K , with $S_0 \cup \dots \cup S_n$ as a subcomplex, as described in Section 6.2. Let \mathcal{T}_M be the restriction of \mathcal{T} to M , the exterior of $S_0 \cup \dots \cup S_n$.
- (3) Let $S = S_{n+1}$ be the next surface in the hierarchy. Since S is exponentially controlled, we may assume that it has exponentially bounded weight in \mathcal{T}_M .
- (4) Homotope S to a nearly embedded alternative admissible surface \hat{S} that is embedded away from the binding circle, that is normal with respect to the handle structure \mathcal{K} , and with binding weight at most $200w^3(w - \sum_i \chi(S_i))^2 w(S)$, as described in Section 6.7.
- (5) This bound on the binding weight is too large for our purposes, and so we modify \hat{S} to reduce its binding weight significantly. This may require exchange moves and also a modification to the earlier surfaces in the hierarchy.
- (6) Modify S_0, \dots, S_n, \hat{S} further so that \hat{S} is embedded, rather than just nearly embedded.
- (7) The surface \hat{S} is in generalised admissible form, but we require it to be alternative admissible. A small perturbation to the surface achieves this.

Step (1) is described in Section 7.2. Step (2) is explained in Sections 6.2, 6.3 and 6.5. Step (4) is described in Section 6.7. Step (5) will require considerable work and will be explained in Sections 9, 10, 11, 12 and 13. Step (6) is described in Section 7.3. Step (7) is given in Section 7.4.

7.2. Making the hierarchy well-spaced. We suppose that we have an admissible partial hierarchy S_0, \dots, S_n . We now show how to make it well-spaced. The boundary pattern P forms an embedded graph in an arc presentation. Let F be

a component of $(S_0 \cup \dots \cup S_n) \setminus P$. We will show how to modify F so that it is still nearly admissible, with the same boundary, and so that the closure in F of the union of the tiles incident to ∂F is a regular neighbourhood of ∂F . Performing this operation for each component of $(S_0 \cup \dots \cup S_n) \setminus P$ will make the hierarchy well-spaced.

Pick a transverse orientation on F . We cut the hierarchy along ∂F and isotope F a little in this transverse direction. Thus, ∂F no longer becomes attached to P . We then insert between ∂F and P an annulus A as shown in Figure 24. Specifically, between each fibre of ∂F and the adjacent fibre of P , we insert another fibre in the same page. The union of these new fibres will form a core curve C for the annulus A , which will divide A into two smaller annuli. One of these annuli is attached to ∂F , extends beyond F and then rapidly returns to C . The other annulus emanates from P initially in the same way that the original copy of F did, but it then rapidly returns to C . (See Figure 24.)

FIGURE 24. Making F well-spaced

A small isotopy places A into nearly admissible form keeping its boundary fixed. In the case where F has no boundary saddles, the induced foliation on F is then as shown in Figure 25.

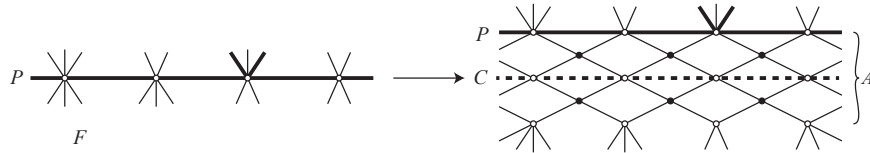


FIGURE 25. The effect on the foliation

When F has a boundary saddle incident to an arc in ∂F , then one possibility for the new singular foliation near that arc is shown in Figure 26.

Note that each vertex of ∂F has been replaced by 3 vertices. Since in each surface S_i of the partial hierarchy, the graph $(\partial S_0 \cup \dots \cup \partial S_n) \cap S_i$ has vertices with degree at most 3, we deduce that this gives rise to at most 6 new vertices. Thus, the binding weight of the hierarchy increases by at most a factor of 7.

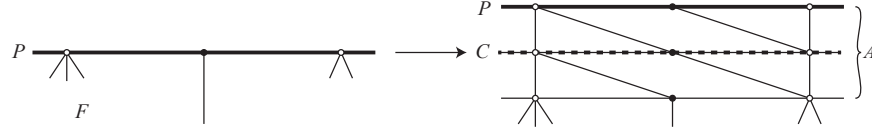
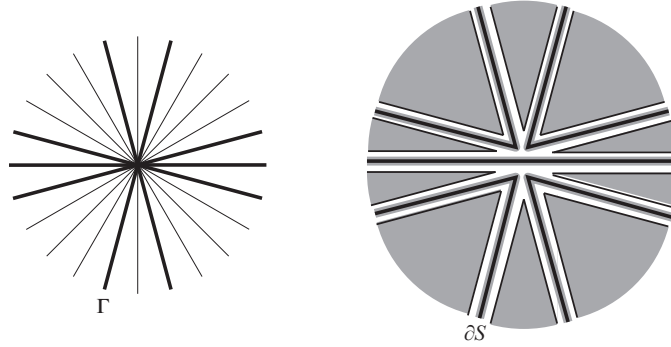


FIGURE 26. The effect on the foliation near a boundary saddle

7.3. Making the surface embedded. In this subsection, we suppose that we have a hierarchy S_0, \dots, S_n in admissible form. In particular, each surface is embedded. We also have the next surface \hat{S} in alternative admissible form, but it might just be nearly embedded. However, it is embedded away from the binding circle. We will show how to modify the hierarchy so that all the surfaces end up embedded.

The exterior of $S_0 \cup \dots \cup S_n$ is a 3-manifold M . Its boundary has a singular foliation. The boundary of S maps to a graph Γ in ∂M that is a union of fibres. We will modify the surfaces S_0, \dots, S_n so that Γ becomes a union of disjoint simple closed curves. We can then make ∂S run along these curves so that it becomes embedded.

The first stage is to cut $S_0 \cup \dots \cup S_n$ along Γ . Then isotope the resulting components of $(S_0 \cup \dots \cup S_n) \setminus \Gamma$ away from Γ but keeping them admissible. The second stage is to realise ∂S as a collection of simple closed curves in their arc presentation. The third stage is to thicken ∂S to a union A of admissible annuli in which ∂S forms the core curves. The final stage is to isotope ∂A so that it runs along $(S_0 \cup \dots \cup S_n) \setminus \Gamma$. In this way, we replace $(S_0 \cup \dots \cup S_n)$ by a new 2-complex $((S_0 \cup \dots \cup S_n) \setminus \Gamma) \cup A$. The binding weight of this 2-complex has gone up by at most $3w_\beta(\partial S)$, since each new vertex is a vertex of A , and A has three vertices for each vertex of ∂S . See Figure 27.

FIGURE 27. Making S embedded. Left: a neighbourhood of a vertex in S . The image of ∂S is a graph Γ . Right: Cutting S along Γ , and thickening ∂S to annuli.

Since the binding weight of $S_0 \cup \dots \cup S_n$ increases by at least the binding weight of ∂S , we only perform this procedure when the latter quantity is relatively small.

7.4. From generalised admissible form to alternative admissible form.

Given a generalised admissible surface, we may perturb it to an alternative admissible surface, as follows. Any generalised interior saddle with n singular leaves emanating from it may be perturbed to $(n-2)/2$ saddles. An example where $n = 6$ is shown in Figure 28. Similarly, any generalised boundary saddle with n singular leaves may be perturbed to $(n-2)/2$ saddles (if n is even) or $(n-3)/2$ saddles plus a boundary saddle (if n is odd). This does not affect the binding weight of the surface.

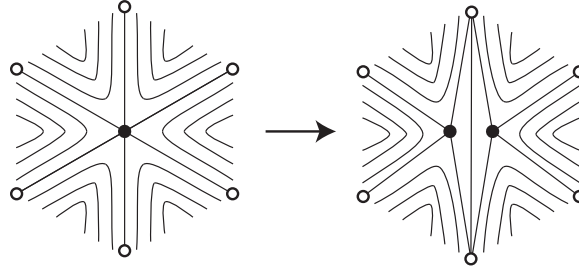


FIGURE 28. A small isotopy replaces a generalised interior saddle with several interior saddles

8. BRANCHED SURFACES

8.1. Definitions. A *generalised branched surface* is a compact generalised 2-complex B smoothly embedded in a 3-manifold M , with the following extra structure:

- (1) At each point x in B , there is a specified tangent plane in $T_x(M)$ that is denoted by $T_x(B)$. All the 1-cells and generalised 2-cells containing x are required to have tangent spaces that lie in $T_x(B)$.
- (2) As a consequence, for each point x in the interior of a 1-cell, the tangent plane $T_x(B)$ is divided into two half-planes by the tangent space of the 1-cell. These are the two *sides* of the 1-cell. At each such point x , we require that there are generalised 2-cells on both sides of x or just one side. The closure of the union of points of the former type is the *branching locus* of B . The closure of the union of points of the latter type is ∂B , the *boundary* of B .

The generalised 2-cells of B are known as *patches*. We permit a 1-cell in the branch locus to have a single patch on each side. A generalised branched surface is known as a *branched surface* if each 1-cell either has patches on both sides or it is incident to a single patch. The union of points of the latter type is the *boundary* of B and is denoted ∂B .

We let $N(B)$ be a thickening of B in M . This thickening is almost a regular neighbourhood, except that ∂B is required to lie in $\partial N(B)$. (This would hold true for a regular neighbourhood if $\partial B \subset \partial M$, but we will not be making this requirement.) This thickening has a decomposition as a union of *fibres*, each of which is a closed interval. There is a map $\pi: N(B) \rightarrow B$ that collapses each fibre to a point. Away from a small regular neighbourhood of the 1-skeleton of B , π is the projection map for an I -bundle. We choose this decomposition of $N(B)$ into fibres so that the following conditions hold:

- (1) For each $x \in B$, the fibre through x is required to have tangent space that does not lie in $T_x(B)$.
- (2) Each fibre intersects $\partial N(B)$ in its endpoints, together possibly with a collection of closed intervals.
- (3) When the fibre does not intersect ∂B , these closed intervals in (2) are required to lie in the interior of the fibre.

The boundary of $N(B)$ decomposes into two subsurfaces:

- (1) the horizontal boundary $\partial_h B$, which is the closure of the union of the endpoints of the fibres;
- (2) the vertical boundary $\partial_v B$, which is $\partial N(B) \setminus \partial_h B$.

Each component of $\partial_v B \setminus \pi^{-1}(\partial B)$ is a *cusps*. Cusps arise near the branching locus of B . (See Figure 29.)

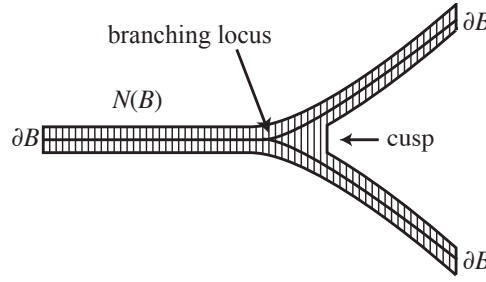


FIGURE 29. The fibred neighbourhood of a branched surface B

Note that the branched surfaces that we consider are far from generic. In particular, more than three 2-cells may be incident to a 1-cell of B . This means the cusps of $\partial N(B)$ can be quite complicated. However, the following result was established in Lemma 8.1 of [17]. When $B \cap \partial M = \partial B$ and $\pi^{-1}(\partial B) = N(B) \cap \partial M$ and at each $x \in \partial B$, the tangent plane $T_x(B)$ does not equal $T_x(\partial M)$, then each cusp is either an annulus or a disc D such that $D \cap \partial M$ is two arcs in ∂D .

8.2. Natural handle structure. Let B be a branched surface with a fibred neighbourhood $N(B)$. Suppose that B has only disc patches. Then $N(B)$ has a handle structure \mathcal{H} where each i -handle is a thickened i -cell of B . We term this the *natural handle structure* of $N(B)$.

As explained in Section 4, the surface $\mathcal{F} = \partial \mathcal{H}^0 \cap (\mathcal{H}^1 \cup \mathcal{H}^2)$ plays an important role when considering handle structures. In the case of the natural handle structure, we can arrange that this surface is a union of fibres. So one could view this as a fibred neighbourhood of a generalised train track (once this has been suitably defined). In fact, the branched surface structure imposes further conditions on \mathcal{F} , as follows.

Let v be a 0-cell of B and let H_0 be the associated 0-handle of \mathcal{H} . We require that H_0 is a copy of $D^2 \times [-1, 1]$, where the fibres are of the form $\{*\} \times [-1, 1]$. The vertex v has a well-defined tangent space $T_v B$, which we may view as intersecting ∂H_0 in $\partial D^2 \times \{0\}$. Because B is smoothly embedded, its intersection with ∂H_0 lies within $\partial D^2 \times [-1, 1]$, provided H_0 is chosen appropriately. Moreover, the tangent planes of B near v all are close to the tangent plane through the equator.

Similarly, we require that each 1-handle of \mathcal{H} is of the form $[-1, 1] \times [-1, 1] \times I$, where the fibres are of the form $\{*\} \times I$ for a point $*$ in $[-1, 1] \times [-1, 1]$.

8.3. Surfaces weakly carried by branched surfaces. A compact surface S is *weakly carried* by a generalised branched surface B if S is properly embedded in $N(B)$ transversely to the fibres of $N(B)$, and also $\partial S = S \cap \partial N(B)$ lies within a small regular neighbourhood of the 1-cells of B . Note that we permit ∂S to intersect the cusps and the horizontal boundary of $N(B)$.

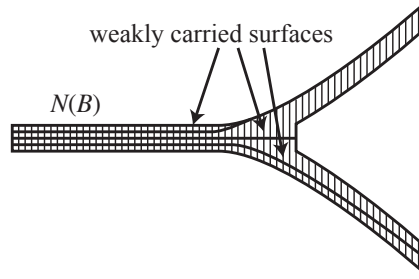


FIGURE 30. Weakly carried surfaces

When a surface S is weakly carried by B , each patch of B inherits a non-negative integer, which is the *weight* of S in that patch. This is equal to the number of intersections between S and any fibre away from a regular neighbourhood of the branching locus.

The surfaces that we will consider will be standard in the natural handle structure on $N(B)$. Such surfaces satisfy one of the conditions of normality, as follows.

Lemma 8.1. *Let B be a generalised branched surface in which each patch is a disc. Let S be a surface weakly carried by B that is standard in the natural handle structure \mathcal{H} on $N(B)$. Then in each 0-handle H_0 of \mathcal{H} , each disc D of $S \cap H_0$ satisfies (1) in the definition of normality. That is, ∂D runs over each 1-handle of \mathcal{F} at most once.*

Proof. As discussed in Section 8.2, the surface $\mathcal{F} \cap \partial H_0$ lies within $\partial D^2 \times [-1, 1] \subset \partial H_0$, where $\partial D^2 \times \{0\}$ is the equator of ∂H_0 corresponding to the tangent plane of the branched surface. Each fibre of $N(B) \cap \partial H_0$ lies within $\{*\} \times [-1, 1]$ for some point $*$ in ∂D^2 . Each 1-handle of \mathcal{F} has a product foliation by fibres, each parallel to the co-core of the 1-handle. The boundary of the disc D lies within $\partial D^2 \times [-1, 1]$ and intersects $\{*\} \times [-1, 1]$ at most once for each $*$ in ∂D^2 . It therefore runs over each 1-handle of \mathcal{F} at most once. \square

8.4. Surfaces carried by branched surfaces. We say that S is *carried* by a branched surface B if it is weakly carried and, in addition, $S \cap \pi^{-1}(\partial B) = \partial S$. Thus, in Figure 30, only the lower surface satisfies this condition.

A surface S carried by B is determined, up to fibre preserving ambient isotopy, by its weights in the patches. The collection of these weights is known as the *vector* for S , and is denoted $[S]$.

When surfaces S , S_1 and S_2 are carried by B , we say that S is the *sum* of S_1 and S_2 if $[S] = [S_1] + [S_2]$. Then S_1 and S_2 are *summands*. There is an intrinsic characterisation of one carried surface S_1 being a summand of another carried

surface S : this occurs if and only if the weight of S in each patch is at least that of S_1 .

8.5. A branched surface carrying an alternative admissible surface. Suppose that we have a hierarchy H in admissible form. We give the exterior of H the handle structure \mathcal{K} described in Section 6.5. Let S be a nearly embedded surface in the exterior of H in alternative admissible form with no annular tiles. We assume that S is embedded away from the binding circle. Hence, in \mathcal{K} , the surface is actually embedded. We also make the assumption that it is normal with respect to \mathcal{K} . We now construct a branched surface B that carries S .

Since S is nearly admissible, it is a union of tiles. We say that two tiles of S have the same *type* if they are normally parallel with respect to the handle structure \mathcal{K} . For each type of tile of S , take one representative of this tile. These tiles will form the patches of B . When two tiles share a separatrix, the corresponding patches are glued along a 1-cell of B . This may lead to a sequence of identifications. For example, if a patch P_1 shares a separatrix with P_2 , and P_2 shares the same separatrix with a patch P_3 , then P_1 and P_3 are, of course, glued along this separatrix.

The result is the branched surface B . Note that each patch of B is a disc and so B is a 2-complex, rather than a generalised 2-complex. Let $N(B)$ be a fibred regular neighbourhood of B . We give $N(B)$ its natural handle structure \mathcal{H} . For each i -cell of B , its regular neighbourhood forms an i -handle of \mathcal{H} .

We will measure parallelism using the 3-manifold $N(B)$ and its handle structure \mathcal{H} . More specifically, we say that two vertices of S have \mathcal{H} -parallel stars if their stars are normally parallel with respect to \mathcal{H} .

Note that S is indeed carried by B , for the following reason. Form a regular neighbourhood $N(S)$ of S , so that $N(S)$ intersects each handle of \mathcal{K} in a union of elementary normal discs. When two tiles of S are \mathcal{K} -parallel, then attach the space between them to $N(S)$. The result is the 3-manifold $N(B)$. By definition of parallelism, the region between two normally parallel tiles has the structure as $D \times [0, 1]$ for a disc D , where $D \times \{0, 1\}$ is the two tiles. Thus, $D \times [0, 1]$ has the structure of a union of intervals. These combine to form the fibres of $N(B)$. If we collapse each of these fibres to a point, we obtain the branched surface B . By construction, S is a subset of $N(B)$ transverse to the fibres and so it is weakly carried by B . In fact, $S \cap \pi^{-1}(\partial B) = \partial S$, for the following reason. Note that $\partial S = S \cap \partial N(S)$ lies in the interior of $\partial_v N(S)$. The fibred neighbourhood $N(B)$ is obtained from $N(S)$ by adding to it. But none of these additions affects the interior of $\partial_v N(S)$, and furthermore $\pi^{-1}(\partial B)$ contains $\partial_v N(S)$. Hence, $S \cap \pi^{-1}(\partial B) = \partial S$, and so S is carried by B .

9. MODIFICATIONS TO AN ADMISSIBLE SURFACE

When an admissible surface has a vertex of small valence, there is a modification that can be made to the surface and the arc presentation, which reduces the complexity of the surface. These modifications were key to Dynnikov's analysis of the unknot and its spanning disc. Similar moves had also been utilised in the context of braids by Birman and Menasco [3] and by Bennequin [2].

9.1. Two-valent interior vertex. Let S be a generalised admissible surface. The star of a two-valent interior vertex s in S is shown in Figure 31.

The arrangement of this star in 3-space is shown in Figure 32.

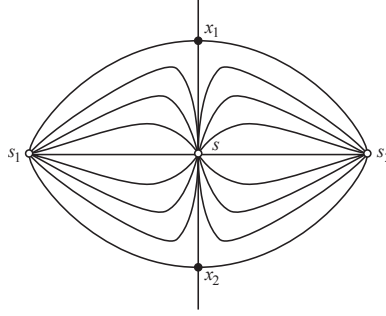


FIGURE 31. A 2-valent interior vertex and its star

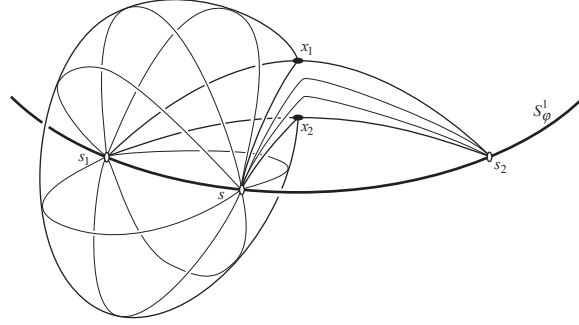


FIGURE 32. The arrangement of the surface with respect to the arc presentation

There are two separatrices emanating from s . Let t_1 and t_2 be their θ values, where $0 \leq t_1 < t_2 < 2\pi$. For each $t \in (t_1, t_2)$, there is a leaf of the foliation in \mathcal{D}_t running from s to a vertex s_1 in the star of s . For every $t \in S_\theta^1 - (t_1, t_2)$, there is a leaf of the foliation in \mathcal{D}_t running from s to the other vertex s_2 in the star of s . We may suppose, after relabelling if necessary, that the vertices s_1, s, s_2 occur in that order around S_ϕ^1 .

Let $i: S_\phi^1 - \{s_1, s, s_2\} \rightarrow S_\phi^1 - \{s_1, s, s_2\}$ be the following involution. It sends the interval (s_1, s) to the interval (s, s_2) linearly, preserving this ordering. Similarly, it sends (s, s_2) to the interval (s_1, s) also linearly, preserving this ordering. It fixes the interval between s_2 and s_1 .

We may now apply the following procedure, which will modify S , and which may also move the link and indeed any other surfaces in S^3 , including any surfaces in the partial hierarchy that have already been constructed.

- (1) Remove the star of s from S .
- (2) For any leaf of the foliation or arc of the link with an endpoint in $\phi \in S_\phi^1 \setminus \{s_1, s_2\}$, replace it with a leaf ending at $i(\phi)$.
- (3) For the leaves of the foliation that ended at s_1 or s_2 , but which were not part of the star of s , replace them with a leaf ending at s .

This procedure can be achieved via an isotopy of the link and the surfaces. The effect of this on the rectangular diagram is to perform a generalised exchange move (see [8]). The effect on S is to remove the star of s from S , and to identify the leaves running from s_1 in the boundary of this star with the leaves running from

s_2 . Therefore this decreases the binding weight of S by 2. It has no effect on the singular foliation in the remainder of S , and no effect on the singular foliation of the other surfaces.

9.2. Three-valent interior vertex. Let s be a 3-valent interior vertex of S . Let s_2, s_3, s_4 be the three vertices in its star. After relabelling, we may suppose that these vertices occur in the order s, s_2, s_3, s_4 around S_ϕ^1 . Then the singular foliation near s is as shown in Figure 33. (In the figure, only saddles are shown for simplicity, but they may be generalised saddles.) Each generalised saddle has an *orientation*. Here, one picks a transverse orientation on S near s . If the transverse orientation on S at a generalised saddle points in the direction of increasing θ , we say that it is *positive*. Otherwise it is *negative*. There are three generalised saddles in the boundary of the star of s . Therefore, there are two, x_1 and x_2 say, that have the same orientation. By relabelling if necessary, we may assume that x_1 is the generalised saddle involving s, s_2 and s_3 , and that x_2 is the generalised saddle involving s, s_3 and s_4 . Let $t_1 < t_2$ be the θ values of these saddles. Note that $t_1 \neq t_2$ because the pages \mathcal{D}_{t_1} and \mathcal{D}_{t_2} are distinct at s . So x_1 and x_2 are definitely distinct generalised saddles.

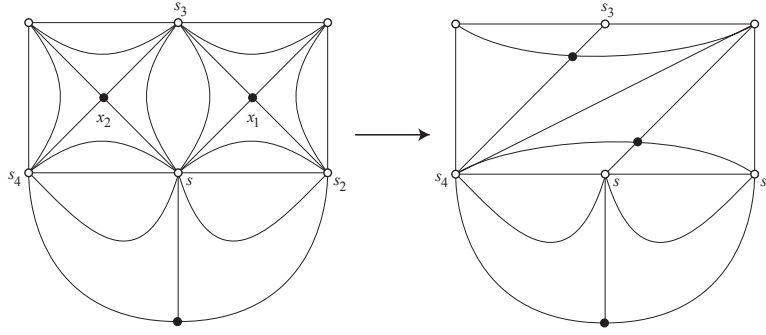


FIGURE 33. A 3-valent interior vertex

In previous work on this problem [8], the following procedure was used. The first stage of the modification was to remove all events from the interval (t_1, t_2) . This was achieved by moving all saddles in the interval (t_1, t_2) with separatrices ending in (s, s_3) into the future, and all saddles with separatrices ending in (s_3, s) into the past. Similarly, all boundary arcs in the interval (t_1, t_2) with endpoints in (s, s_3) move into the future, and those with endpoints in (s_3, s) move into the past. Then one increases the θ -value of x_1 until it equals that of x_2 , and then is slightly greater than it. The new singular foliation is shown in Figure 33. The effect of this process on the singular foliation is to turn s into a 2-valent interior vertex. One can then perform the procedure described in Section 9.1 to remove this vertex.

In this paper, we will do something that is a little different. Again, we remove all events from the interval (t_1, t_2) , by moving some into the future and some into the past. And again, we will increase the θ -value of x_1 , but we will stop when x_1 and x_2 have the same θ -values. These can then be combined to form a single generalised saddle. The effect on the singular foliation of S is shown in Figure 34. Once again, the vertex s has become 2-valent, and it can be removed using the method in Section 9.1.

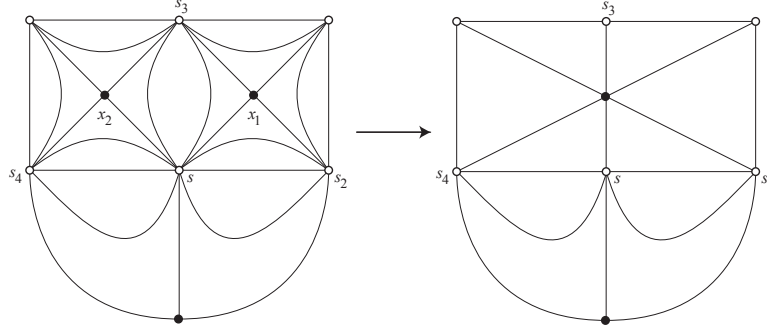


FIGURE 34. The modification used in this paper

9.3. Low-valence boundary vertices. For a vertex s on the boundary of a generalised admissible surface S , there are two types of separatrix that are incident to it: those lying in ∂S , and those lying in the interior of S . We say that the *boundary valence* $d_b(s)$ is equal to the number of separatrices in ∂S that are incident to s . The *interior valence* $d_i(s)$ is the number of interior separatrices incident to s . We say that a vertex s is a *low-valence boundary vertex* if it lies on ∂S and $2d_i(s) + d_b(s) < 4$. Thus, these vertices fall into one of five possible types:

- (1) $d_b(s) = 0$ and $d_i(s) = 0$.
- (2) $d_b(s) = 0$ and $d_i(s) = 1$.
- (3) $d_b(s) = 1$ and $d_i(s) = 0$.
- (4) $d_b(s) = 1$ and $d_i(s) = 1$.
- (5) $d_b(s) = 2$ and $d_i(s) = 0$.

Case 1 cannot occur, because a vertex cannot have zero valence unless the component of S containing it is a single bigon tile. The remaining cases are shown in Figure 35.

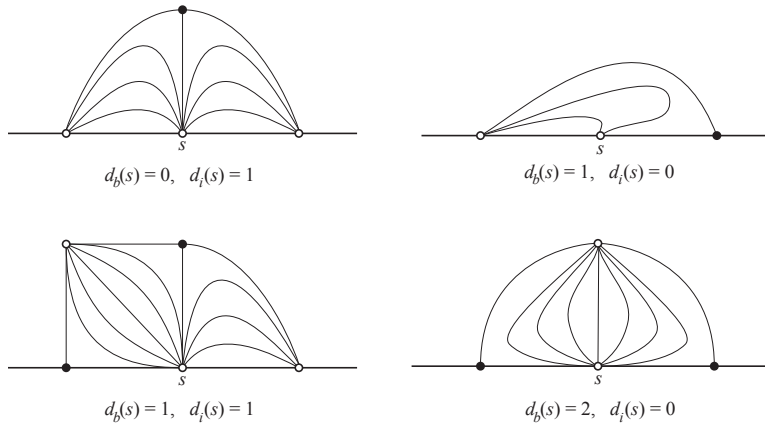


FIGURE 35. Low-valence boundary vertices

In previous papers [8, 17], such an arrangement usually specifies an ambient isotopy of ∂S . For example, when considering arc presentations of the unknot,

which bounds a spanning disc S , Case (2) above leads to an isotopy of the unknot, which results in a sequence of cyclic permutations and exchange moves, followed by a destabilisation. We want to do something similar here, but simply moving ∂S would appear to be problematic. The boundary of S lies in earlier surfaces of the hierarchy, and this needs to remain the case. We therefore introduce a completely new type of modification, called a *wedge insertion*.

9.4. Wedge insertion. Let M be the exterior of an admissible partial hierarchy. Let P be its boundary pattern. Let D be a nearly admissible disc embedded in M disjoint from P , with $D \cap \partial M$ a single arc in ∂D , and where the remainder of ∂D is an arc in some page. Then the associated *wedge* W is a regular neighbourhood of D in M . The 3-manifold $M' = M \setminus W$ is obtained *inserting the wedge* W into M . (See Figure 36.)

It is clear that, after a small isotopy, $\partial M' \setminus P$ may be placed in nearly admissible form. Moreover, the singular foliation on $\partial M'$ can be determined from that of M and D . We will only be concerned with the binding weight of $\partial M'$, which one can arrange to be at most the binding weight of ∂M , plus twice the binding weight of D . When D is the closure of the star of a low-valence boundary vertex, its binding weight is at most 3 and so the binding weight of ∂M increases by at most 6.

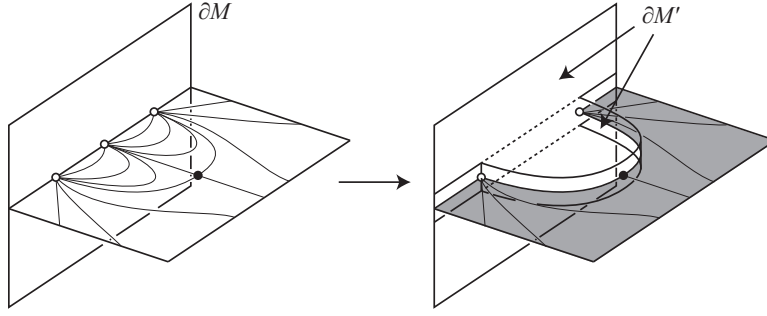


FIGURE 36. Inserting a wedge into M

We will apply this when there is a low-valence boundary vertex. We take D to be the star of this vertex. The effect of this is to remove this vertex, and thereby reduce the binding weight of S . However, it also makes ∂M more complicated, and therefore this move needs to be applied carefully and sparingly.

In our situation, M is the exterior of an admissible partial hierarchy for K . We are assuming that D is disjoint from the boundary pattern P of M , and so it intersects a single surface S' in this partial hierarchy. We can view the insertion of the wedge W as being achieved by an isotopy of S' . This isotopy slides S' across W , replacing $S' \cap W$ with $\partial W \setminus S'$.

9.5. Parallelism. The above moves do all simplify the surface S , in the sense that they reduce its binding weight. However, they do so at a cost. In the case of the 2-valent and 3-valent interior vertex, they require a generalised exchange move, and this is achieved using a sequence of Reidemeister moves. In the case of a low-valence boundary vertex, the link does not need to be moved, but this modification does make the earlier surfaces in the partial hierarchy more complicated. We therefore need to ensure that we use as few of these moves as possible. As in [17], the key is

to perform a variation of these moves, each of which has the same cost in terms of Reidemeister moves or the effect on the earlier parts of the hierarchy, but which lead to a much more substantial decrease in the binding weight of S . We will describe these moves in the next section.

10. PARALLELISM AND ENVELOPES

In the previous section, we described several modifications that one can make to a generalised admissible surface. As we have already mentioned, it will be important that when we perform these moves, we take advantage of parallelism. Initially, we will say that two parts of S are parallel if they are normally parallel in the handle structure \mathcal{H} . Recall from Section 8.5 that \mathcal{H} was a handle structure for the fibred regular neighbourhood N of a branched surface B that carries S . As we modify S , we will also need to modify B , N and \mathcal{H} . In this section, we will describe these modifications.

10.1. Distinguished subsurfaces. We will keep track of disjoint distinguished subsurfaces X and Y in ∂N . We initially set X to be $\partial_h N$ and set $Y = \emptyset$. As in Section 4.4, $\partial X \cup \partial Y$ is a union of standard curves in \mathcal{H} .

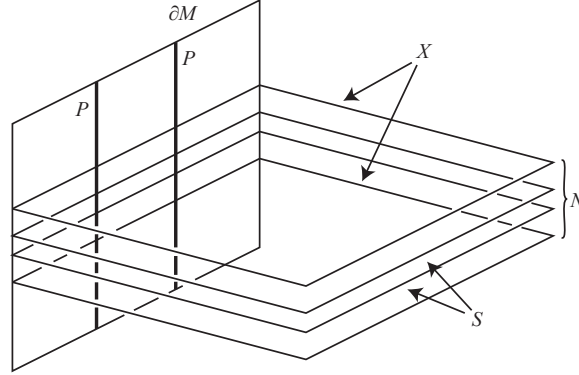


FIGURE 37. The arrangement of N , S and X

10.2. Admissible envelope. We formalise the properties of S , B , N and \mathcal{H} in the following definition. It will also be useful to keep track of the two disjoint distinguished subsurfaces X and Y , and the boundary pattern P .

We say that $(N, X, Y, P \cap N, \mathcal{H}, S)$ is an *admissible envelope* if

- (1) N is a fibred regular neighbourhood of a generalised branched surface with only disc patches.
- (2) \mathcal{H} is the natural handle structure for N .
- (3) X and Y are disjoint subsurfaces of ∂N with $\partial X \cup \partial Y$ standard in \mathcal{H} .
- (4) Each component of $\partial_h N$ lies in X or Y or is disjoint from both.
- (5) $P \cap N$ is a collection of disjoint arcs properly embedded in $\partial_v N$ with endpoints in X and is otherwise disjoint from $X \cup Y$. Each such arc lies in a 0-handle of \mathcal{H} , and at most one arc lies in each 0-handle.
- (6) S is a surface weakly carried by N that is standard with respect to \mathcal{H} .
- (7) S is disjoint from X .

- (8) S is in generalised admissible form with no annular tiles.
- (9) Each 2-handle of \mathcal{H} is a regular neighbourhood of a tile of S , and all the components of intersection between S and this 2-handle inherit corresponding singular foliations. In particular, each 2-handle of \mathcal{H} contains at least one tile of S .
- (10) For each 1-handle H_1 of \mathcal{H} , each component of $H_1 \cap S$ is a regular neighbourhood either of a separatrix or of an arc in ∂S joining two vertices and not containing a generalised boundary saddle.
- (11) For each 0-handle H_0 of \mathcal{H} , $H_0 \cap S$ consists either of a collection of regular neighbourhoods of vertices of S , or of a collection of regular neighbourhoods of generalised saddles of S (but not a mixture of the two). We refer to 0-handles of these types as *vertex* or *saddle* 0-handles.

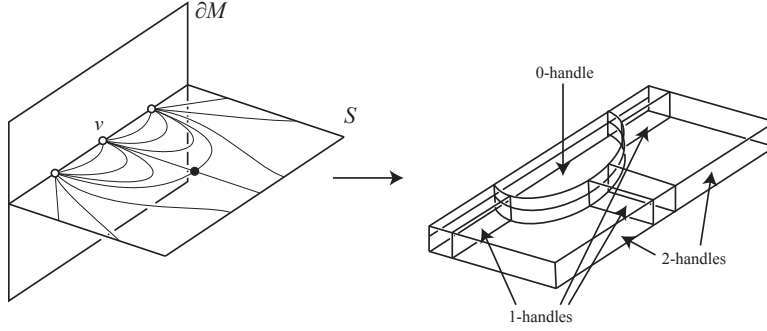
10.3. The role of the distinguished subsurface X . The reason for introducing the distinguished subsurface X is as follows. Consider a vertex v of S on the boundary of M , for example as in Figure 38. This lies in a 0-handle H_0 of \mathcal{H} . As explained in Section 8.2, this 0-handle is identified with $D^2 \times [-1, 1]$, where the fibres of N are of the form $\{*\} \times [-1, 1]$. Hence, $X \cap (D^2 \times [-1, 1]) = D^2 \times \{-1, 1\}$. Attached to this 0-handle are 1-handles and 2-handles, and these also have an I -bundle structure, although the I -fibres of these handles intersect $D^2 \times [-1, 1]$ in closed intervals in $\{*\} \times [-1, 1]$ rather than necessarily the whole of $\{*\} \times [-1, 1]$. Since H_0 is incident to ∂M , there is an arc α in ∂D^2 such that $H_0 \cap \partial M = \alpha \times [-1, 1]$. In particular, no handles are attached to $(\alpha - \partial\alpha) \times [-1, 1]$.

As discussed in Section 4, the surface $\mathcal{F} = \mathcal{H}^0 \cap (\mathcal{H}^1 \cup \mathcal{H}^2)$ plays an important role. This has a handle structure with $\mathcal{F}^0 = \mathcal{H}^0 \cap \mathcal{H}^1$ being the 0-handles and $\mathcal{F}^1 = \mathcal{H}^0 \cap \mathcal{H}^2$ equalling the 1-handles. Consider the 0-handles of \mathcal{F}^0 incident to some component of $\partial\alpha \times [-1, 1]$. Arranged along this component, there may be several 0-handles of \mathcal{F}^0 . Consider one 0-handle of \mathcal{F} that is outermost in $\partial\alpha \times [-1, 1]$. The intersection between S and this 0-handle of \mathcal{F} consists of arcs. Note that there is at least one such arc, since every handle of \mathcal{H} has non-empty intersection with S . If we had not specified the distinguished subsurface X , these arcs would have violated condition (4) in the definition of normality in Section 4.1. However, in the presence of X , they do satisfy condition (4) in the definition of normality in Section 4.4. Since it is convenient for S to be normal in \mathcal{H} , this explains the reason for using X .

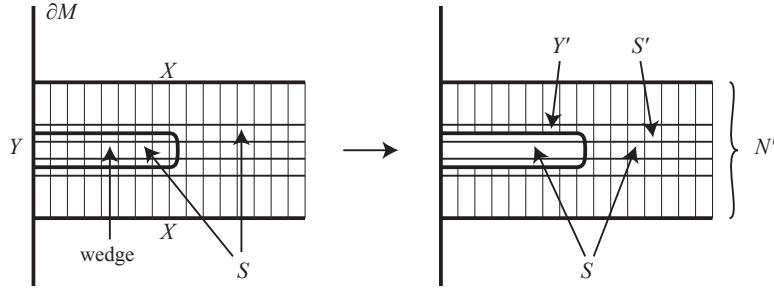
Note that P is disjoint from \mathcal{F} , for the following reason. Since $P \cap N$ lies in ∂N , the interior of $P \cap \partial\mathcal{H}^0$ is disjoint from \mathcal{F} . Moreover, as the endpoints of each arc of $P \cap N$ lie in X , these endpoints are disjoint from \mathcal{F} .

10.4. Another admissible envelope. As explained above, we are keeping track of a 3-manifold N with a handle structure \mathcal{H} . In fact, we will keep track of two 3-manifolds N and N' , with handle structures \mathcal{H} and \mathcal{H}' respectively. The 3-manifold N' will be a subset of N , and each handle of \mathcal{H}' will lie within a handle of \mathcal{H} .

Initially, $N' = N$ and $\mathcal{H}' = \mathcal{H}$. But when we insert the first wedge W , then these two manifolds and handle structures become distinct. With the insertion of W , N and \mathcal{H} remain unchanged, but we remove W from N' . When W is a regular neighbourhood of the star of a vertex of ∂S , then there is a natural way of defining \mathcal{H}' . We will also consider the surface $S' = S \cap N'$. Thus, when a wedge is inserted, we view S' as being shrunk but S as remaining unchanged.

FIGURE 38. A vertex v in ∂S and the 0-handle of \mathcal{H} that contains it

We say that two vertices of S' have \mathcal{H}' -parallel stars if their stars are normally parallel with respect to \mathcal{H}' .

FIGURE 39. The effect on N' when a wedge is inserted

10.5. Nested admissible envelopes. We formalise the above arrangements with the following definition. We say that two admissible envelopes $(N, X, Y, P \cap N, \mathcal{H}, S)$ and $(N', X', Y', P \cap N', \mathcal{H}', S')$ are *nested* if the following hold:

- (1) N' is a subset of N , and is obtained from N by removing a sequence of wedges, where each wedge is transverse to the fibres of N and is disjoint from $\partial_h N$.
- (2) Each fibre of N' is a subset of a fibre of N .
- (3) $X' = X = \partial_h N$ and $P' \cap N' = P \cap N$.
- (4) $Y = \emptyset$.
- (5) Y' is a surface properly embedded in N and that is parallel to a subsurface of $\partial_v N$; the product region between Y' and $\partial_v N$ is where the wedges are inserted.
- (6) $X' \cup Y'$ contains $\partial_h N'$.
- (7) $S \cap N' = S'$.
- (8) S' is obtained from S by wedge insertions and boundary-compressions along clean discs that intersect $\partial N'$ in arcs lying in Y' ;
- (9) S' is a union of tiles of S .
- (10) Each handle of \mathcal{H}' lies in a handle of \mathcal{H} with the same index and their product structures agree.

We explained above why we keep track of the surface $X = X'$. The reason for the designated surface Y' is that it encodes the location where the wedges are inserted when forming N' from N .

The following result helps to clarify the relationship between S and S' .

Lemma 10.1. *Let S be a compact incompressible boundary-incompressible surface properly embedded in (M, P) . Let $(N, X, Y, P \cap N, \mathcal{H}, S)$ and $(N', X', Y', P \cap N', \mathcal{H}', S')$ be nested admissible envelopes. Let M' be the 3-manifold $M \setminus (N \setminus N')$, which is obtained from M by removing wedges. Then S' is a copy of S in M , plus possibly some clean properly embedded discs, in the sense that there is an isotopy taking M to M' and taking S plus possibly some clean discs to S' . In particular, S' is incompressible and boundary-incompressible in M' , and S' is π_1 -injective in S .*

Proof. By assumption, S' is obtained from S by wedge insertions and boundary-compressions along clean discs that intersect $\partial N'$ in arcs lying in Y' . At each wedge insertion, M' and S' are obtained from what they were previously by an isotopy. When a boundary compression is performed along a clean disc that intersects $\partial N'$ in an arc lying in Y' , then this disc is a boundary compression disc for S' in M' . However, inductively, S' is a copy of S plus possibly some clean discs. Since S is boundary-incompressible, the arc of intersection between S' and a clean compression disc is boundary parallel in S' . Hence, after the boundary compression, S' remains a copy of S plus some clean discs. \square

10.6. The complexity of the handle structures. It will be important that, as we make the modifications to \mathcal{H} and \mathcal{H}' , they remain of bounded ‘complexity’. We now make this notion precise. It is defined using the vertex 0-handles.

We consider the intersection between the vertex 0-handles of \mathcal{H} with the union of the 1-handles and 2-handles of \mathcal{H} . This is a surface \mathcal{F}_v . We define \mathcal{F}'_v similarly, but using the handle structure \mathcal{H}' .

For a component F' of \mathcal{F}'_v , we define the following complexity:

$$c(F') = \max\{-\chi(F') + 2|\partial X \cap F'|, 0\}.$$

Here, X' is the distinguished subsurface of $\partial N'$. For a component F of \mathcal{F}_v , we define the following:

$$c_+(F) = -\chi(F) + 2|\partial X \cap F| + 1.$$

We define $c(\mathcal{F}')$ to be the sum of $c(F')$ over each component F' of \mathcal{F}' . We define $c_+(\mathcal{F})$ to be the sum of $c_+(F)$ over each component F of \mathcal{F} . We will measure the complexity of \mathcal{H} by $c_+(\mathcal{F})$, and will measure the complexity of \mathcal{H}' by $c(\mathcal{F}')$. We will ensure that, during any modification to \mathcal{H} and \mathcal{H}' , these complexities do not increase.

We will give some justification in Section 10.13 for this definition of $c(\mathcal{F}')$. The following also provides some explanation.

Lemma 10.2. *Let F be a component of \mathcal{F} or \mathcal{F}' . Then the maximal number of disjoint non-parallel properly embedded essential arcs in $F - X$ is at most $\max\{3c(F), 1\}$. This is less than $3c_+(F)$ unless F is a disc disjoint from X . Here, an arc in $F - X$ is said to be essential if it cannot be homotoped, keeping its endpoints fixed, to an arc in $\partial F - X$. Also, we say that two arcs α_1 and α_2 in F*

are parallel if they separate off a disc in F containing exactly one copy of α_1 and exactly one copy of α_2 and that is disjoint from X .

Proof. Note first that we may assume that F is not a disc disjoint from X , because such a disc cannot contain any essential arcs. Let α be a maximal set of arcs as in the lemma. For a component F_0 of $F \setminus \alpha$, define its index $I(F_0)$ to be

$$-\chi(F_0) + 2|F_0 \cap X| + |F_0 \cap \partial\alpha|/4.$$

Then the sum of $I(F_0)$ over all components F_0 of $F \setminus \alpha$ is equal to $c(F)$. Unless F is an annulus disjoint from X , each component F_0 has positive index, since otherwise a component of α is inessential or two components are parallel. In the case where F is an annulus disjoint from X , α consists of a single arc that cuts the annulus into a single region with zero index. We also note that if $I(F_0) > 0$, then in fact $I(F_0) \geq 1/2$. Hence, the number of components of $F \setminus \alpha$ is at most $2c(F)$ unless F is an annulus disjoint from X . Again using the additivity of index, we deduce that

$$|\alpha| = \sum_{F_0} |F_0 \cap \partial\alpha|/4 = c(F) + \sum_{F_0} \chi(F_0) - \sum_{F_0} 2|F_0 \cap X| \leq c(F) + |F \setminus \alpha| \leq 3c(F).$$

Again, this formula does not hold when F is an annulus disjoint from X , but in this case $|\alpha| = 1$ and $c(F) = 0$. Thus, we obtain that $|\alpha| \leq \max\{3c(F), 1\}$ in general. Note that this is clearly less than $3c_+(F)$ unless F is a disc disjoint from X . \square

10.7. The complexity of the natural handle structure. In this section, we compute an upper bound for $c(\mathcal{H})$ and $c_+(\mathcal{H})$ where \mathcal{H} is the natural handle structure arising from the initial alternative admissible surface S .

Lemma 10.3. *Let $H = \{S_1, \dots, S_k\}$ be a partial hierarchy with exterior M and boundary pattern P , and let w be its binding weight. Let \mathcal{K} be the handle structure of (M, P) described in Section 6.5. Let S be a nearly embedded alternative admissible surface that is in normal form with respect to \mathcal{K} . Then the number of tile types of S is at most $832w^2(w - \sum_i \chi(S_i))$.*

Proof. The possible normal disc types of S are described in Lemmas 6.9, 6.10, 6.11 and 6.12. Each tile must be incident to a saddle, and this must lie in an expanded tetrahedron or an expanded turnover. We see that from Lemmas 6.10 and 6.11 that there are at most three such elementary disc type in each expanded turnover and there are at most five such normal elementary disc types in each expanded tetrahedron. The outermost tiles of any given type must intersect some outermost disc in an expanded tetrahedron or expanded turnover. By Lemmas 6.3 and 6.4, there are at most $208w^2(w - \sum_{i=1}^n \chi(S_i))$ such normal discs. By Lemma 6.13, each such disc can intersect at most 4 tiles. So, the number of tile types is at most $832w^2(w - \sum_{i=1}^n \chi(S_i))$. \square

Lemma 10.4. *The complexity of the initial natural handle structure \mathcal{H} satisfies*

$$c(\mathcal{H}) \leq c_+(\mathcal{H}) \leq 11648w^2(w - \sum_i \chi(S_i)).$$

Proof. The number of tile types of S is at most $832w^2(w - \sum_i \chi(S_i))$ by Lemma 10.3. Each tile type gives rise to at most two 1-handles of \mathcal{F}_v (one at each vertex of the tile). So, $|\mathcal{F}_v^1| \leq 1664w^2(w - \sum_i \chi(S_i))$. Also, since each 0-handle of \mathcal{H} contains a component of \mathcal{F}^1 , this also forms an upper bound on the number of 0-handles of \mathcal{H} .

Each component of $X \cap \mathcal{F}_v$ intersects some 1-handle of \mathcal{F}_v and any given 1-handle intersects at most two such components of $X \cap \mathcal{F}_v$. Each component of $X \cap \mathcal{F}_v$ gives rise to at most two points of $\partial X \cap \mathcal{F}_v$. So, $|\partial X \cap \mathcal{F}_v| \leq 6656w^2(w - \sum_i \chi(S_i))$. So,

$$c_+(\mathcal{H}) = c_+(\mathcal{F}_v) \leq |\mathcal{F}_v^1| + 2|X \cap \mathcal{F}| \leq 11648w^2(w - \sum_i \chi(S_i)).$$

The inequality $c(\mathcal{H}) \leq c_+(\mathcal{H})$ is immediate. \square

10.8. Removing handles.

Lemma 10.5. *Let \mathcal{H}'_- be a handle structure obtained from \mathcal{H}' by removing some handles disjoint from P and X . For \mathcal{H}'_- to be a handle structure, we require that whenever a handle of \mathcal{H}' is removed, then so are all the incident handles with higher index. Then $c(\mathcal{H}'_-) \leq c(\mathcal{H}')$.*

Proof. Let F be a component of \mathcal{F}' . When a 1-handle of F disjoint from X is removed, this decreases $-\chi(F)$ by 1. If this 1-handle is non-separating in F , the resulting component therefore has smaller complexity than F . Suppose that the 1-handle is separating in F . Then the quantities $-\chi(F) + 1$ and $2|\partial X \cap F|$ are shared out between the two new components. So, the total complexity goes down by one, unless a component is created that is a disc disjoint from X . But in this case, the other component is a copy of F , and so complexity is unchanged.

Suppose now that a 0-handle of \mathcal{F}' is removed. The incident 1-handles have already been removed. Therefore, this procedure is simply the removal of a component of \mathcal{F}' , which clearly does not increase complexity. \square

10.9. Shrinking the handle structure to achieve crude normality. We now give another modification to N' . The end result will be that S' is crudely normal in (\mathcal{H}', X') , as defined in Section 4.4. The moves that we make are versions of the usual normalisation moves and hence are guaranteed to terminate. More precisely, each move decreases the extended weight of S' , as defined in Section 4.3.

We saw in Lemma 8.1 that since S' is weakly carried by N' , it satisfies (1) in the definition of crude normality.

Suppose that, contrary to (4') in the definition of crude normality, there is an arc α of intersection between S' and a 0-handle of \mathcal{F}' such one endpoint of α lies in a 1-handle of \mathcal{F}' and the other endpoint lies in $\partial N'$. Suppose that one of the arcs β in the boundary of the 0-handle of \mathcal{F}' joining $\partial\alpha$ is disjoint from X and P and from any other 1-handles of \mathcal{F}' . Let us assume that the interior of β is disjoint from S' . Then $\alpha \cup \beta$ bounds a disc in the 0-handle of \mathcal{F}' , and the product of this disc with the interval forms a ball B_1 in the incident 1-handle of \mathcal{H}' . One of the endpoints of α lies in a 2-handle of \mathcal{H}' . This is divided into balls by S' . Let B_2 be the ball that intersects the interior of β . (See Figure 40.)

The modification we make is to remove a regular neighbourhood of $B_1 \cup B_2$ from N' . The intersection between $B_1 \cup B_2$ and S' is a disc that intersects $\partial S'$ in a single arc. The modification removes this disc from S' . This leaves S' unchanged up to isotopy in S .

Alternatively, suppose that there is an arc α of intersection between S' and a 0-handle of \mathcal{F}' that separates off a disc from \mathcal{F}' that is disjoint from the 1-handles of \mathcal{F}' and from X , contrary to (3') in the definition of crude normality. Let us assume that the interior of this disc is disjoint from S' . The product of this disc with the interval is a ball B_1 in the incident handle of \mathcal{H}' . The modification we

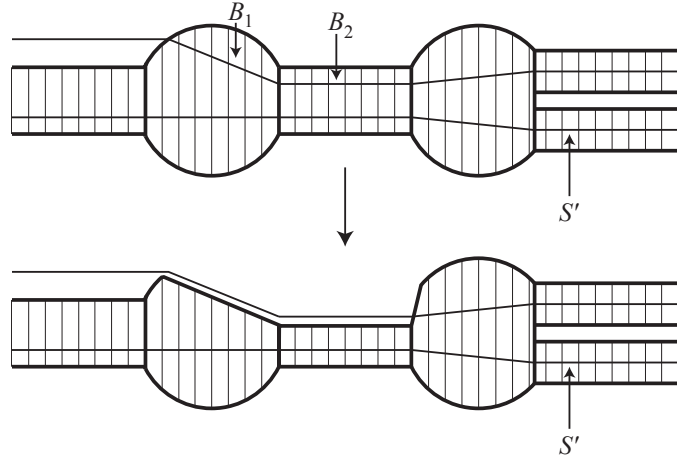


FIGURE 40. Shrinking the handle structure

can then perform is to remove a regular neighbourhood of this ball from N' . This has the effect of boundary-compressing S along a clean boundary-compression disc D . The intersection between D and $\partial_h N'$ misses X' and hence lies in Y' . The intersection between D and $\partial_v N'$ lies in those parts of $\partial_v N'$ intersecting $\partial S'$, which again lie in Y' . Hence, $D \cap \partial N'$ lies in Y' , as required in the definition of nested admissible envelopes.

These modifications may create handles of \mathcal{H}' that do not intersect S' . In this case, we remove these handles. By Lemma 10.5, this does not increase $c(\mathcal{H}')$.

Note that these are solely modifications to N' , \mathcal{H}' and S' and not N , \mathcal{H} and S .

Lemma 10.6. *Suppose that S' is crudely normal in \mathcal{H}' . Then, for each 0-handle of \mathcal{F}' lying in a 0-handle $D^2 \times [-1, 1]$ of \mathcal{H}' , one of the following must hold:*

- (1) *it intersects both $X \cap (D^2 \times \{1\})$ and $X \cap (D^2 \times \{-1\})$, or*
- (2) *it has 1-handles of \mathcal{F}' on both sides of it.*

Proof. Suppose that, on the contrary, there is a 0-handle D of \mathcal{F}' that is disjoint from $X \cap (D^2 \times \{1\})$ or $X \cap (D^2 \times \{-1\})$ and that has 1-handles emanating from at most one side of it. We are identifying D with $[-1, 1] \times I$ with fibres of the form $* \times I$. By assumption, the 1-handles of \mathcal{F}' are disjoint from some component of $\{-1, 1\} \times I$, say $\{-1\} \times I$. They are also disjoint from $[-1, 1] \times \partial I$. The intersection between S' and this 0-handle of \mathcal{F}' is a non-empty collection of arcs. Each arc of $S' \cap D$ intersects at most one 1-handle of \mathcal{F}' since it is transverse to the fibres. Consider an outermost such arc. It divides D into two discs, one of which, D' say, has interior disjoint from S' . We may choose the arc so that D' also misses X . The boundary of D' cannot intersect an entire component of $D \cap (\mathcal{F}')^1$, since S' runs over every 1-handle of \mathcal{F}' . Thus, we deduce that this arc violates (3') or (4') in the definition of crude normality. \square

10.10. Reduced handle structure. We say that \mathcal{H}' is *reduced* if both the following hold:

- (1) no 0-handle of \mathcal{F}' is disjoint from ∂X and intersects the 1-handles of \mathcal{F}' in at most one arc, and

(2) for each 0-handle H'_0 of \mathcal{H}' , $\partial H'_0 \cap (\mathcal{F}' \cup X \cup P)$ is connected.

We now show how to modify the handle structure \mathcal{H}' to make it reduced without increasing its complexity $c(\mathcal{F}')$.

Suppose that there is a 0-handle of \mathcal{F}' that is disjoint from X and intersects the 1-handles of \mathcal{F}' in at most one arc. Let H'_0 be the 0-handle of \mathcal{H}' containing this 0-handle. Then any component of $S' \cap H'_0$ violates (1), (3') or (4') in the definition of crude normality. Hence, S' must be disjoint from this 0-handle. But then the incident 1-handle of \mathcal{H}' , and possibly an incident 2-handle, would have been removed.

Thus we may assume that \mathcal{F}' satisfies (1) in the definition of reduced. Suppose that for some 0-handle H'_0 of \mathcal{H}' , $\partial H'_0 \cap (\mathcal{F}' \cup X \cup P)$ is disconnected. Then there is a disc D properly embedded in H'_0 that separates the components of $\partial H'_0 \cap (\mathcal{F}' \cup X \cup P)$. We may assume that D intersects each disc of $S' \cap H'_0$ in a (possibly empty) collection of arcs. The modification we perform is to cut H'_0 along D . This has no effect on the complexity $c(\mathcal{F}')$. But it does increase the number of 0-handles of \mathcal{H}' . Since each 0-handle of \mathcal{H}' contains a component of \mathcal{F}' , this process is guaranteed to terminate. The effect on S' is to boundary-compress it along clean boundary-compression discs. Hence, S' remains π_1 -injective in S , and $S' \setminus S'$ still lies in a collar on ∂S .

We need to justify why this modification maintains N' as a thickening of a generalised branched surface weakly carrying S' . Recall that H'_0 is identified with $D^2 \times [-1, 1]$. The discs $D^2 \times \{-1, 1\}$ lie in $\partial_h N'$, and so are disjoint from \mathcal{F}' . Each component of $D^2 \times \{-1, 1\}$ either lies in X or is disjoint from X . We may also assume that $\mathcal{F}' \cap (D^2 \times \{-1, 1\}) \subset X$. Since ∂D is disjoint from X , we may isotope ∂D so that it lies in $\partial D^2 \times [-1, 1]$. We claim that ∂D is an essential curve in $\partial D^2 \times [-1, 1]$. If not, some component of \mathcal{F}' lies within in a disc in $\partial D^2 \times [-1, 1]$. But this implies that some 0-handle of \mathcal{F}' has 1-handles of \mathcal{F}' emanating from only one side of it and is disjoint from X , contradicting Lemma 10.6. Hence, ∂D is essential in $\partial D^2 \times [-1, 1]$. We may then isotope it so that it is everywhere transverse to the I -fibres in $\partial D^2 \times [-1, 1]$. For the only obstruction to doing this would be a 0-handle of \mathcal{F}' with 1-handles emanating from only one side of it, and where this component of \mathcal{F}' prevented ∂D being pulled tight. So once ∂D is transverse to the I -fibres, we can make all of D transverse to the I -fibres. Hence, cutting the 0-handle along D creates two 0-handles that are still fibred with intervals and hence copies of $D^2 \times [-1, 1]$. Therefore, cutting N' and S' along D' still results in a 3-manifold that is a thickening of a generalised branched surface weakly carrying the new S' .

Lemma 10.7. *Suppose that \mathcal{H}' is reduced. Then \mathcal{F}' contains no components that are discs disjoint from X .*

Proof. Let F' be such a component. The handle structure on F' is a thickened graph. This graph is a tree, by the hypothesis that F' is a disc. Hence, it is either an isolated vertex or contains a leaf. This contradicts the hypothesis that \mathcal{H}' is reduced. \square

Note that the disc components F' of \mathcal{F}' disjoint from X were a special case in the definition of complexity $c(\mathcal{F}')$. The first part of the formula for $c(F')$ gives -1 , but this is upgraded to 0 by the fact that one takes a maximum. It is therefore convenient to be able to assume that such components do not arise.

We now give some consequences of crude normality.

Lemma 10.8. *Suppose that \mathcal{H}' is reduced and that S' is crudely normal. Then no component of $S' \cap \mathcal{F}'$ is a simple closed curve that bounds a disc in \mathcal{F}' or an arc parallel to an arc in $\partial\mathcal{F}' - X$.*

Proof. Consider first the case where a component of $S' \cap \mathcal{F}'$ is a simple closed curve that bounds a disc in \mathcal{F}' . An innermost such curve bounds a disc component of $\mathcal{F}' \setminus S'$. This inherits a handle structure that is a thickened tree. This is either a single 0-handle or it has a leaf. When it is a single 0-handle, then S' fails to be standard, since its intersection with the corresponding 1-handle of \mathcal{H}' has an annular component. When the disc component of $\mathcal{F}' \setminus S'$ has a leaf, this implies that the relevant component of $S' \cap (\mathcal{H}')^0$ runs over a 1-handle of \mathcal{F}' more than once, contradicting crude normality.

Now suppose that a component of $S' \cap \mathcal{F}'$ is an arc parallel to an arc in $\partial\mathcal{F}' - X$. An outermost such arc separates off a disc D in \mathcal{F}' . Again this inherits a handle structure that is a thickened tree. If this is a single 0-handle, then this contradicts (3') in the definition of crude normality. So suppose that this handle structure has a leaf. This is a 0-handle of D that is incident to a single 1-handle of D . This 0-handle intersects ∂D in an arc α . If α lies wholly in $\partial\mathcal{F}'$, then this implies that \mathcal{F}' has a leaf, contradicting the assumption that \mathcal{H}' is reduced. If both the endpoints of α lie in S' , then this contradicts (1) in the definition of crude normality. If one endpoint of α lies in S' and the other lies in $\partial\mathcal{F}'$, then this contradicts (4') in the definition of crude normality. The remaining possibility is that α contains $\partial D \cap S'$ in its interior. But then D has at least one other leaf, and the intersection between that leaf and ∂D lies wholly in $\partial\mathcal{F}'$, and again we have a contradiction. \square

Lemma 10.9. *Suppose that \mathcal{H}' is reduced and that S' is crudely normal. Suppose that two components of $S' \cap \mathcal{F}'$ are topologically parallel in $\mathcal{F}' - X$. Then they are normally parallel.*

Proof. Suppose that two components of $S' \cap \mathcal{F}'$ are topologically parallel in $\mathcal{F}' - \partial X$, and that these are arcs α_1 and α_2 . Then these two arcs separate off a disc D from \mathcal{F}' . This disc inherits a handle structure. One can define the index of a 0-handle of D to be half the number of intersection components with the 1-handles of D plus a quarter of the number of intersection points with $\partial\alpha_1 \cup \partial\alpha_2$ minus 1. Then the sum of the indices of the 0-handles of D is equal to 0. Each 0-handle of D has non-negative index, since otherwise crude normality is violated. So, every 0-handle of D has zero index. Thus, D consists of a line of 0-handles and 1-handles. Suppose that there is at least one 1-handle, for otherwise α_1 and α_2 are normally parallel. Then the end 0-handles each intersect $\partial\alpha_1 \cup \partial\alpha_2$ in two points. There are then two possibilities: either $\alpha_1 \cap D$ and $\alpha_2 \cap D$ each lie in a single 0-handle of D , or α_1 and α_2 are normally parallel. We must rule out the former possibility. Consider the 1-handle of D incident to the 0-handle containing α_1 . This must have non-empty intersection with S' . Hence, at least one arc of S' runs from this 1-handle of D into the 0-handle. But then crude normality is violated.

A similar but simpler argument applies when α_1 and α_2 are simple closed curves of $S' \cap \mathcal{F}'$ that are topologically parallel. They cobound an annulus A in \mathcal{F}' . Again each 0-handle of A has index zero. So A is formed of a circle of 0-handles and 1-handles. Therefore, α_1 and α_2 are normally parallel. \square

10.11. Low-valence boundary vertices (revisited). Suppose that \mathcal{H}' is reduced and that S' is crudely normal.

One can take advantage of parallelism in the case of wedge removal. Suppose that s^1, \dots, s^n are low-valence boundary vertices of S' . Let D_i be the star of s^i . Suppose that these stars are parallel in \mathcal{H}' . Then there is an embedding of $D \times [1, n]$ in N' , where $D_i = D \times \{i\}$, and where the remainder of S' is disjoint from $D \times [1, n]$. Then we make take a regular neighbourhood of $D \times [1, n]$ to be the wedge W that is removed from N' .

Thus, with a single wedge removal, the binding weight of S' is reduced by n .

We can see from Figure 35 that each of these stars contains at most one interior vertex of S' in its boundary. The effect on N' and \mathcal{H}' is as follows:

- (1) The component of \mathcal{F}' containing the vertices s^1, \dots, s^n is sliced along its intersection with W , which is a regular neighbourhood of a properly embedded arc γ in \mathcal{F}' . This increases $\chi(\mathcal{F}')$ by 1. We will show below that this does not create any disc components of \mathcal{F}' that are disjoint from X . Hence, $c(\mathcal{F}')$ decreases by 1.
- (2) If any vertices of the star other than s lie in $\partial S'$, then the effect on \mathcal{F}' at these points is cut along an arc, with precisely one boundary point in $\partial \mathcal{F}'$. Thus, the topology of \mathcal{F}' remains unchanged, and hence so does its complexity.
- (3) If there is an interior vertex in the star, then the component of \mathcal{F}' containing this vertex is cut open along an arc. This arc lies within the interior of \mathcal{F}' , and so the effect on $c(\mathcal{F}')$ is to increase it by 1.

So, the overall effect is leave $c(\mathcal{H}')$ unchanged. But we need to show that the procedure in (1), where \mathcal{F}' is sliced along an arc γ , does not create a disc component disjoint from X . The arc γ is a component of $S' \cap \mathcal{F}'$ and so is not parallel to a component of $\partial \mathcal{F}' - X$, by Lemma 10.8. Moreover, it is not an essential arc in an annular component of \mathcal{F}' because its endpoints are joined by an arc of $S' \cap (\partial(\mathcal{H}')^0 \setminus (\mathcal{F}' \cup X \cup P))$. This establishes the claim.

Note also that, when we insert a wedge, we do not change N and \mathcal{H} and so their complexity remains unchanged.

10.12. Cut vertices. The operation of inserting a wedge shrinks S' by a removing discs incident to $\partial S'$. In doing so, $\partial S'$ may, at some stage, bump into itself. At that point, a *cut vertex* may be created. This is a vertex in $\partial S'$ at which S' is not a surface with boundary. Instead, a neighbourhood of that point is a cone on at least two closed intervals.

We can cut S' at these points, giving a surface S'_{cut} . When we do this, we do not forget the remnants of the cut vertices in S'_{cut} . The reason is that if a cut-vertex of S'_{cut} is a low-valence vertex, then we cannot insert a wedge along it.

The *cut-adjusted Euler characteristic* of S'_{cut} , denoted $\chi_c(S'_{\text{cut}})$, is defined to be $\chi(S'_{\text{cut}})$ minus half its number of cut vertices. Since S'_{cut} is obtained from S' by cutting along some arcs, and because each such arc creates two cut vertices, we deduce that $\chi_c(S'_{\text{cut}}) = \chi_c(S')$.

10.13. Justification for the measure of complexity. We can now justify the measure of complexity $c(\mathcal{F}')$ that we chose for \mathcal{F}' . In some respects, it is an awkward quantity, because certain components of \mathcal{F}' have zero complexity. Hence, if one just has a bound on $c(\mathcal{F}')$, then one cannot control the number of these components.

These components come in two types: discs and annuli disjoint from X . The discs disjoint from X can be dealt with using Lemma 10.7, and we therefore may assume that \mathcal{H}' is reduced. However, annuli disjoint from X are harder to control. The following example explains why.

Suppose that, initially, we have two components of \mathcal{F}' , an annulus disjoint from $X \cup P$ and a pair of pants disjoint from $X \cup P$. (See Figure 41.) Suppose that a wedge is inserted, thereby slicing the pair of pants along an arc, into two annuli. Suppose also that the star contains an interior vertex and that the relevant component of \mathcal{F}' surrounding this interior vertex is the annulus disjoint from $X \cup P$. Then a slit is inserted into this annulus, forming a pair of pants. Thus, the wedge insertion has created a pair of pants and two annuli from a pair of pants and a single annulus. The conclusion that one must draw is that, whatever notion of complexity we have for \mathcal{F}' , annuli disjoint from $X \cup P$ must have zero complexity if wedge insertion is not going to increase complexity.

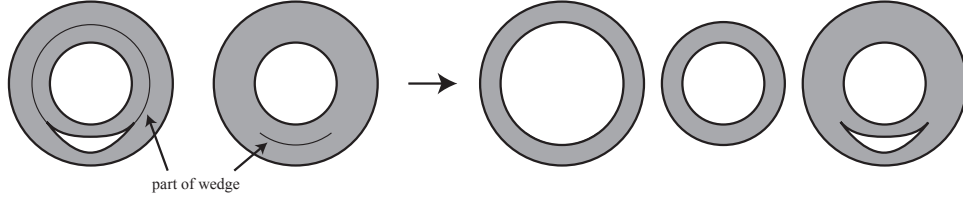


FIGURE 41. An example of the effect of wedge insertion on \mathcal{F}'

10.14. Approximate parallelism. Consider two normal simple closed curves α and α' in \mathcal{F}' . We say that they are *approximately parallel* if

- (1) they have the same \mathcal{H} -type; and
- (2) when \mathcal{F}' is decomposed along them forming a surface F'' , the resulting pieces of F'' lying in the annular region between them are just annuli in which every 0-handle of F'' has two arcs of intersection with the 1-handles of F'' .

We say that two disjoint properly embedded discs in some 0-handle of \mathcal{H}' are *approximately parallel* if their boundary curves lie in \mathcal{F}' and are approximately parallel. Similarly, we say that interior vertices of S' are *approximately parallel* if the discs of intersection between S' and the 0-handles of \mathcal{H}' that contain these vertices are approximately parallel.

When two curves, discs or vertices are approximately parallel, we also say that they have the same *approximate type*.

10.15. Interior 2-valent vertices (revisited). We now examine the ‘parallel’ version of the modification in Section 8.1. Suppose that \mathcal{H}' is reduced.

Consider again a 2-valent interior vertex s of S' . We will now describe a modification to S and S' which removes this vertex and all other vertices of S with approximately parallel stars.

Let s^1, \dots, s^n be these vertices, and for each i , let s_1^i and s_2^i be the two vertices in its star. Then these vertices occur along S_ϕ^1 in the order

$$\dots, s_1^1, \dots, s_1^n, s^n, \dots, s_2^1, s_2^n, \dots$$

Let $i: S_\phi^1 \rightarrow S_\phi^1$ map the interval (s_1^n, s^n) linearly onto the interval (s^1, s_2^1) , map the interval (s^1, s_2^1) linearly onto the interval (s_1^n, s^n) , and fix everything else.

- (1) Remove the stars of s^1, \dots, s^n from S .
- (2) For any leaf of the foliation or arc of the link with an endpoint in $\phi \in (s_1^n, s^n) \cup (s^1, s_2^1)$, replace it with a leaf ending at $i(\phi)$.
- (3) For the leaves of the foliation that ended at s_1^i or s_2^i , but which were not part of the star of s^i , replace them with a leaf ending at s^i .

Thus, a single generalised exchange move is performed, but the binding weight of S is reduced by $2n$. Furthermore, if m is the number of vertices among s^1, \dots, s^n that lie in S' , then the binding weight of S' is reduced by $2m$.

Adjacent to s are two tiles, T_1 and T_2 , say, each of which lies in a 2-handle of \mathcal{H} . Any other vertex with star that is \mathcal{H} -parallel to that of s contains parallel copies of T_1 and T_2 . Consider a sequence of these stars, moving away from s in some transverse direction. At some stage, we reach the last star parallel to s . Beyond this may be a parallel copy of T_1 or a parallel copy of T_2 , but not both. For the sake of being definite, suppose that there is a parallel copy of T_1 . Then the first stage of the procedure is to slice the 2-handle of \mathcal{H} containing T_1 into two handles, thereby separating this parallel copy of T_1 that is not part of a star parallel to that of s from the stars that are parallel to that of s . We also slice the two 1-handles that are incident to this 2-handle and that contain the separatrices incident to s . This procedure is shown in Figure 42.

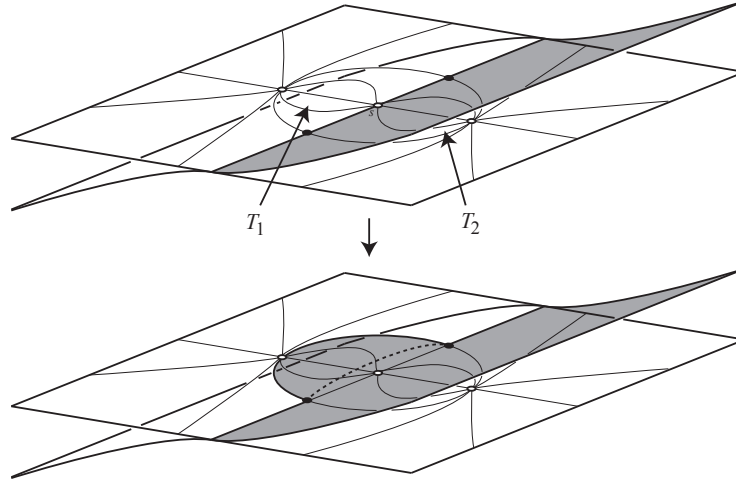
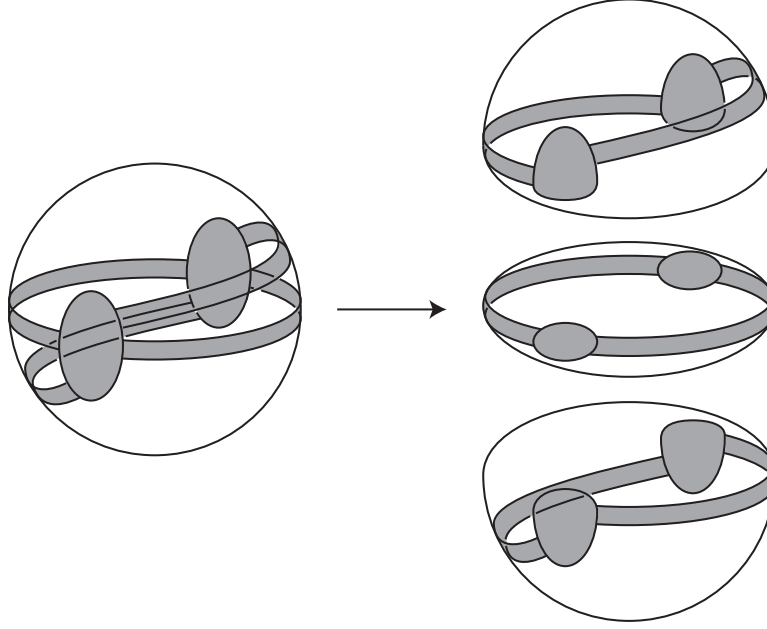


FIGURE 42. The 2-handle of \mathcal{H} containing T_1 is sliced into two 2-handles

We perform a similar procedure, but going away from s in the other transverse direction. Again for the sake of being definite, suppose that in that direction, there is a parallel copy of T_2 but not T_1 . The effect of these two procedures on the 0-handle containing s is to slice it into three 0-handles. The way that this affects \mathcal{F} in these 0-handles is shown in Figure 43.

Note that the original component of \mathcal{F}_v near s is sliced along two arcs, thereby forming three components. This decreases $-\chi(\mathcal{F}_v)$ by 2, but it increases the number of components by 2, and so the contribution to $c_+(\mathcal{F}_v)$ from $-\chi(\mathcal{F}_v)$ is unchanged.

FIGURE 43. The effect on \mathcal{F} near s

However, this is not the only effect on \mathcal{F}_v . Let s_1 and s_2 be the vertices in the star of s . These lie in vertex 0-handles which are also affected by this modification. Each 0-handle intersects the 2-handle containing T_1 or T_2 in a 1-handle of \mathcal{F}_v . This 1-handle of \mathcal{F}_v is replaced by two 1-handles. When this is performed in the two vertex 0-handles of \mathcal{H} , this increases $-\chi(\mathcal{F}_v)$ by 2, but it leaves the number of components of \mathcal{F}_v unchanged. So, the contribution to $c_+(\mathcal{F}_v)$ from $-\chi(\mathcal{F}_v)$ has increased by 2.

We must also analyse the contribution to $c_+(\mathcal{F}_v)$ arising from $\mathcal{F}_v \cap \partial X$. When the 2-handle of \mathcal{H} containing T_1 is sliced in two, this enlarges the horizontal boundary of N . If any components of $\partial_h N$ containing X are enlarged, then we enlarge X also. This moves ∂X but it does not increase the number of vertex 0-handles that it runs over. So the contribution to $c_+(\mathcal{F}_v)$ from $\mathcal{F}_v \cap \partial X$ remains unchanged.

We now perform the procedure described above, which removes the star of s and all parallel stars. This new surface is carried by a new handle structure. One removes the vertex handle containing s , and all the incident handles. Thus, the annular component of \mathcal{F} in this 0-handle is removed, thereby decreasing the complexity $c_+(\mathcal{F})$ by 1.

The vertex 0-handles containing v_1 and v_2 are amalgamated to form a single 0-handle. The way that this affects \mathcal{F} is as follows. The two components of \mathcal{F} in these 0-handles are incident to a 2-handle that was removed because it contained T_1 or T_2 . These two components of \mathcal{F} are then combined to form a single component. The effect here is leave $-\chi(\mathcal{F})$ unchanged but to reduce the number of components by 1. The number of points of $\partial X \cap \mathcal{F}_v$ remains unchanged. So the complexity goes down by 1.

We see that the net effect of this process is to leave the complexity $c_+(\mathcal{F})$ unchanged.

We also observe that the complexity of \mathcal{F}' is unchanged, as follows. The modifications to $\partial X \cap \mathcal{F}'_v$ are exactly as described above. Hence, the only thing that concerns us is the effect on Euler characteristic and whether the modifications create or remove any components of \mathcal{F}'_v that are discs disjoint from X .

The first stage is to increase $-\chi(\mathcal{F}'_v)$ by 2, but it then decreases $-\chi(\mathcal{F}'_v)$ by 2. At no stage is a disc component of \mathcal{F}'_v disjoint from X created. At the first stage, this is because each component of \mathcal{F}'_v is sliced along at most one arc that is parallel to a component of $S' \cap \mathcal{F}'_v$ with endpoints on the same component of $\partial \mathcal{F}'_v$. This does not create a disc component of \mathcal{F}'_v disjoint from X by Lemma 10.8. In the second stage, a non-separating 1-handle of \mathcal{F}'_v is removed from two components of \mathcal{F}'_v , and then these are combined into a single component of \mathcal{F}'_v , which again cannot be a disc disjoint from X .

In the above discussion, we assumed that, in one transverse direction from the tiles parallel to s , there is a copy of T_1 but not T_2 , and in the other transverse direction, there is a copy of T_2 but not T_1 . However, this is not the only possible arrangement. For example, it might be the case that in both transverse directions, there is a copy of T_1 but not T_2 . Alternatively, in one or both transverse directions, there might be neither a copy of T_1 nor a copy of T_2 . In the former situation, the 0-handle is still sliced into three 0-handles. The above argument still applies to give that the complexities of \mathcal{F}_v and \mathcal{F}'_v are unchanged. In the latter situation, where in one or both transverse directions, there is neither a copy of T_1 nor a copy of T_2 , then the 0-handle is sliced into two 0-handles or is not sliced at all. Again it is easy to deduce that the complexities of \mathcal{F}_v and \mathcal{F}'_v are unchanged.

10.16. Interior 3-valent vertices (revisited). Consider a 3-valent interior vertex s . It is contained within an elementary normal disc D in a 0-handle of \mathcal{H} . Its star contains three tiles, and hence D runs through three 1-handles and three 0-handles of \mathcal{F} . Coming out of these three 0-handles are three 1-handles of \mathcal{H} , which end at other vertices s_2, s_3, s_4 . As in Section 9.2, suppose these vertices are arranged around S_ϕ^1 in the order s, s_2, s_3, s_4 . Let x_1 and x_2 be the generalised saddles that are defined in Section 9.2 and that have the same orientation, and let t_1 and t_2 be their θ -values.

We wish to perform the procedure described in Section 9.2, not just to the star of s but to all the vertices with stars that are parallel (with respect to \mathcal{H}). The procedure in Section 9.2 can be viewed as taking place in three steps:

- (1) Moving events occurring in the interval (t_1, t_2) into the future or past.
- (2) Collapsing the tile with boundary s, x_1, s_3, x_2 so that x_1 and x_2 are amalgamated to form a single generalised saddle.
- (3) The process given in Section 9.1 for dealing with 2-valent interior vertices.

Step (1) can be achieved by an isotopy of \mathcal{H} and \mathcal{H}' , and so this does not affect the complexities of \mathcal{H} and \mathcal{H}' . We have already seen in Section 10.15 that the complexities of \mathcal{H} and \mathcal{H}' are unchanged when dealing with parallel copies of a 2-valent interior vertex. So, it suffices to show that the complexities of \mathcal{H} and \mathcal{H}' are unchanged by Step (2).

In Step (2), we collapse the 2-handle containing the tile. The 0-handles containing the generalised saddles x_1 and x_2 are amalgamated into a single 0-handle. These are not vertex 0-handles, and so do not contribute to the complexity of \mathcal{H} . The 1-handles of \mathcal{H} running from s to x_1 and from s to x_2 are amalgamated into a

single 1-handle. The 1-handles running out of s_3 are similarly amalgamated. The effect on the two relevant components of \mathcal{F}_v is to collapse, in each component, a 1-handle of \mathcal{F}_v . This leaves \mathcal{F}_v unchanged up to homeomorphism preserving $X \cap \mathcal{F}_v$, and so does not change its complexity $c_+(\mathcal{F}_v)$. It also leaves \mathcal{F}'_v unchanged up to homeomorphism preserving $X \cap \mathcal{F}'_v$, and so does not change its complexity $c(\mathcal{F}'_v)$.

10.17. The number of star types.

Proposition 10.10. *Suppose that \mathcal{H}' is reduced and that S' is crudely normal. Then the number of star types of non-cut vertices in $\partial S'$ is at most $3c(\mathcal{H}')$.*

Proof. Each vertex s of S' lies within a vertex 0-handle H_0 of \mathcal{H}' . The intersection between S' and H_0 is a collection of elementary normal discs. Let E be the elementary normal disc containing s . This disc E intersects \mathcal{F}' in a single arc or simple closed curve, because s is a non-cut vertex. For the purposes of this proof, we call such a disc *relevant*. We also call the intersection of the disc with \mathcal{F}' a *relevant arc*.

We now show that a relevant disc E in H_0 is determined (up to an isotopy of H_0 preserving X , P and the handles) by its relevant arc $E \cap \mathcal{F}'$. First observe that, because \mathcal{H}' is reduced, $H_0 \cap (\mathcal{F}' \cup X \cup P)$ is connected, and hence its complement is a collection of discs. The arc $E \setminus \mathcal{F}'$ intersects one or two such discs, depending on whether it intersects $P \cap H_0$. (Note that there is at most one arc of $P \cap H_0$.) Hence, the arc $\partial E \setminus \mathcal{F}'$ is determined by its endpoints, as required.

It therefore suffices to bound the number of such relevant arcs. By Lemma 10.8, $S' \cap \mathcal{F}'$ is essential in \mathcal{F}' , in the sense that no simple closed curve bounds a disc in \mathcal{F}' or no arc is parallel to an arc in $\partial \mathcal{F}' - X$. By Lemma 10.9, whenever two arcs are topologically parallel in $\mathcal{F}' - X$, they are normally parallel.

So, when we consider a collection of disjoint relevant discs in H_0 , no two of which are \mathcal{H}' -parallel, we deduce that their intersection with \mathcal{F}' forms a collection of properly embedded arcs, no two of which are topologically parallel. Each arc is essential in the complement of X . So, the maximal number of relevant non-parallel discs in H_0 is at most the number of disjoint non-parallel essential arcs in $(\mathcal{F}' \cap H_0) - X$. This is at most $3 \max\{c(\mathcal{F}' \cap H_0), 1\}$, by Lemma 10.2. However, in fact, the number of such arcs is bounded above by $3c(\mathcal{F}' \cap H_0)$, for the following reason. When the maximum is 1, then $\mathcal{F}' \cap H_0$ is a disc or annulus disjoint from X . We saw in Lemma 10.7 that when \mathcal{H}' is reduced, \mathcal{F}' has no disc components disjoint from X . Any annular component of \mathcal{F}' disjoint from X is composed of a string of 0-handles and 1-handles joined together in a circular fashion. Hence, any normal arc in such an annulus must join distinct components of $\partial H_0 \setminus \mathcal{F}'$. Therefore such an arc is not relevant, as it comes from a cut vertex. \square

Proposition 10.11. *Suppose that \mathcal{H}' is reduced and that S' is crudely normal. Then the number of approximate types of interior vertices in S' is at most $2c_+(\mathcal{H}) + c(\mathcal{H}')$.*

Proof. Consider a maximal collection of interior vertices of S' , no two of which have the same approximate type. Each gives rise to a simple closed curve in \mathcal{F}_v . None of these curves bounds a disc in \mathcal{F}_v , by Lemma 10.8. The maximal number of disjoint non-parallel simple closed curves in \mathcal{F}_v , none of which bounds a disc, is at most $2c_+(\mathcal{F}_v)$. To prove this, consider a component F of \mathcal{F}_v , and maximal collection of disjoint non-parallel curves in F , none of which bounds a disc in F . Since F is planar, this is $-2\chi(F) + 1 \leq 2c_+(F)$.

However, it may be the case that two closed curves in \mathcal{F} are topologically parallel but not of the same approximate type. Certainly, if they are topologically parallel in \mathcal{F} , then they have the same \mathcal{H} -type. However, condition (2) in the definition of approximate type may not hold. We may cut \mathcal{F}' along these curves and one of the resulting pieces of surface between them may not be an annulus. Note that we do not in fact obtain any components of this surface that are discs disjoint from ∂X , since these discs would then have been part of \mathcal{F}'_v , contradicting the assumption that \mathcal{H}' is reduced. Thus, when \mathcal{F}'_v is cut along this collection of vertices, the result is a collection \mathcal{F}'' of surfaces, none of which is a disc disjoint from ∂X . Since we are cutting along simple closed curves, $\chi(\mathcal{F}'') = \chi(\mathcal{F}'_v)$. Therefore, the number of components of \mathcal{F}'' that are not annuli disjoint from ∂X is at most $c(\mathcal{F}'_v)$. These components can increase the number of approximate types of interior vertices of S' by at most $c(\mathcal{F}'_v)$. \square

10.18. The approximate star types of tiles. Let T be a tile of S' with vertices v_1 and v_2 in its boundary, and let T' be another tile of S' with vertices v'_1 and v'_2 in its boundary. We say that T and T' have the same *approximate star type* if at least one of the following holds (after possibly relabelling v'_1 and v'_2):

- (1) the stars of v_1 and v'_1 are parallel in \mathcal{H}' , as are the stars of v_2 and v'_2 ;
- (2) all their vertices lie in the interior of S' , and the stars of v_1 and v'_1 are approximately parallel, as are the stars of v_2 and v'_2 .

11. THE EUCLIDEAN SUBSURFACE

We continue to consider the link K in an arc presentation, and an admissible partial hierarchy H for the exterior of K . We denote the exterior of H by M .

We consider a generalised admissible surface S in M with no annular tiles. We also assume that no component of S is a bigon tile. We suppose that S is part of nested admissible envelopes $(N, X, Y, P \cap N, \mathcal{H}, S)$ and $(N', X, Y', P \cap N', \mathcal{H}', S')$. In particular, S contains the π_1 -injective subsurface S' .

Let $\text{approx-type}(S')$ be the number of approximate star types of tiles of S' . We will suppose that the following quantity is minimised, up to pattern-preserving isotopy:

$$(w_\beta(S'), \text{approx-type}(S')).$$

As usual, such pairs are compared using lexicographical ordering. More specifically, suppose that there is no pattern-isotopy of M taking S and S' to S_2 and S'_2 , and nested admissible envelopes $(N_2, X_2, Y_2, P \cap N_2, \mathcal{H}_2, S_2)$ and $(N'_2, X_2, Y'_2, P \cap N'_2, \mathcal{H}'_2, S'_2)$ satisfying

$$(w_\beta(S'_2), \text{approx-type}(S'_2)) < (w_\beta(S'), \text{approx-type}(S')),$$

$$c_+(\mathcal{H}_2) \leq c_+(\mathcal{H}), \quad c(\mathcal{H}'_2) \leq c(\mathcal{H}').$$

Now define the *Euclidean subsurface* E of S' to be the result of removing the interior vertices with valence other than 4, the vertices v on $\partial S'$ with $2d_i(v) + d_b(v) \neq 4$, the intersection $P \cap S'$, the generalised interior saddles that are not interior saddles, and the generalised boundary saddles that are not boundary saddles. So, E includes:

- (1) The interior of each tile of S' .
- (2) The interior of each separatrix of S' .
- (3) Each interior vertex of S' with valence 4.

- (4) Each boundary vertex v with $2d_i(v) + d_b(v) = 4$ that is disjoint from the boundary pattern P .
- (5) Each interior saddle and boundary saddle of S' .

A similar Euclidean surface was considered in [17].

Each square tile in E has four arc separatrices in its boundary. We may therefore realise this tile as the interior of a Euclidean square with side length 1, where the separatrices form the sides. Each half tile of E is realised as a right-angled isosceles triangle with side lengths 1, 1 and $\sqrt{2}$, where the hypotenuse lies in $\partial S'$. Thus, E has a locally Euclidean Riemannian metric. The remainder $S' \setminus E$ is a collection of isolated points.

Define the *degree* of a point in $S' \setminus E$ to be the number of separatrices to which it is incident. The *degree* of $S' \setminus E$, denoted $\deg(S' \setminus E)$, is the sum of the degrees of its points. A key part of the argument is to find a bound on the area of E in terms of $\deg(S' \setminus E)$. A simple geometric argument, analogous to Lemma 7.2 of [17], gives the following.

Lemma 11.1. *Suppose that each point of E is at a distance at most R from $S' \setminus E$. Then the area of E is at most $(\pi/4)R^2 \deg(S' \setminus E)$.*

Proof. For each point y in E , there is a shortest path from y to $S' \setminus E$, which is a Euclidean geodesic with length at most R . Thus, E is covered by the union of the geodesics emanating from $S' \setminus E$ with length at most R . These geodesics sweep out a region of area at most $(\pi/4)R^2 \deg(S' \setminus E)$. Thus, this forms an upper bound for the area of E . \square

Remark 11.2. Lemma 7.2 in [17] asserts that the area of E is bounded above by $\pi(R+1)^2 \ell(\partial E)$, where $\ell(\partial E)$ is the ‘combinatorial length’ of ∂E . The definition of $\ell(\partial E)$ in [17] was however incorrect, and should instead have been the number of separatrices in ∂E plus the sum of the degrees of the isolated points in ∂E . With that corrected definition, the proof of Lemma 7.2 in [17] does not need to be adjusted, and its application in the proof of Theorem 7.1 of [17] is valid.

Thus, it remains to control the distance of each point of E from $S' \setminus E$. This is achieved by the following theorem, which is the main result of this section.

Theorem 11.3. *Let $(N, X, Y, P \cap N, \mathcal{H}, S)$ and $(N', X, Y', P \cap N', \mathcal{H}', S')$ be nested admissible envelopes. Suppose that there is no homeomorphism $h: N \rightarrow N'$ such that all the following hold:*

- (1) h is equal to the identity on $P \cap N$ and X (and so extends to M);
- (2) h is a Dehn twist along a torus or clean annulus properly embedded in M ;
- (3) $(N, X, Y, P \cap N, \mathcal{H}, h(S))$ and $(h(N'), X, h(Y'), P \cap N', h(\mathcal{H}'), h(S'))$ are nested admissible envelopes;
- (4) $w_\beta(h(S')) < w_\beta(S')$.

Suppose also that there is no pattern-isotopy taking S and S' to S_2 and S'_2 , and nested admissible envelopes $(N_2, X_2, Y_2, P \cap N_2, \mathcal{H}_2, S_2)$ and $(N'_2, X_2, Y'_2, P \cap N'_2, \mathcal{H}'_2, S'_2)$ satisfying

$$(w_\beta(S'_2), \text{approx-type}(S'_2)) < (w_\beta(S'), \text{approx-type}(S')),$$

$$c_+(\mathcal{H}_2) \leq c_+(\mathcal{H}), \quad c(\mathcal{H}'_2) \leq c(\mathcal{H}').$$

Let e be the number of approximate star types of tiles of $E \setminus N_2(S' \setminus E)$. Then, for each point of the Euclidean subsurface E of S' , one of the following holds:

- (1) it has distance at most $4784e$ from $S' \setminus E$, or
- (2) it lies in a toral or clean annular component of S' that lies in E and has binding weight at most $23552e^2$.

The proof of this theorem relies heavily on work of the author in Section 9 of [17]. We will need to spend some time recalling some of this material, but for more details, the reader should refer to [17].

Note that we may assume that there is no homeomorphism h as in Theorem 11.3 for the following reason.

Lemma 11.4. *Let (M, P) be a compact orientable irreducible boundary-irreducible 3-manifold with boundary pattern. Let S be a properly embedded surface. If (M, P) contains an essential clean annulus or essential torus, then suppose that S is the JSJ tori and annuli. Let $h: M \rightarrow M$ be a Dehn twist along a properly embedded torus or clean annulus. Then $h(S)$ is strongly equivalent to S .*

Proof. Let T be the torus or annulus along which the Dehn twist is performed. If T is a torus, then $h(S)$ is strongly equivalent to S . If T is an inessential annulus, then h is pattern-isotopic to the identity and so $h(S)$ is pattern-isotopic to S . On the other hand, if T is an essential clean annulus, then S is the JSJ tori and annuli, by hypothesis. Hence, S can be pattern-isotoped off T . Therefore, again $h(S)$ is pattern-isotopic to S . \square

So, suppose that there is a homeomorphism h as in Theorem 11.3. Then by the above lemma, we may replace S by $h(S)$. We may also replace the nested admissible envelopes by $(N, X, Y, P \cap N, \mathcal{H}, h(S))$ and $(h(N'), X, h(Y'), P \cap N', h(\mathcal{H}'), h(S'))$. These handle structures have the same complexity as the original ones, because $c(h(\mathcal{H}')) = c(\mathcal{H}')$. This process decreases the binding weight of S' by (4) of Theorem 11.3. So by choosing S' to have smallest binding weight, we can assume that there is no homeomorphism h .

11.1. A ‘Euclidean’ branched surface. Recall that S' is properly embedded in the manifold N' . Let M' be obtained from N' by attaching $M \setminus N$. Thus, M' is obtained from M by inserting wedges, but we suppose that when a wedge is inserted in a way that creates a cut vertex, then we do not quite insert the full wedge, in order to maintain M' as a 3-manifold.

The first step is to construct a branched surface B_E lying in M . The principal type of patch of B_E arises from the tiles of the Euclidean surface E of S' . More precisely, each approximate star type of tile in \mathcal{H} of $E \setminus N_2(S' \setminus E)$ gives rise a patch of B_E . But we also consider a slight enlargement of $E \setminus N_2(S' \setminus E)$, consisting of a regular neighbourhood of $E \setminus N_2(S' \setminus E)$ in E , and parallel pieces of this also form tiles of B_E . This enlargement of $E \setminus N_2(S' \setminus E)$ is carried by the branched surface B_E .

Let $N(B_E)$ be a thickening of B_E , and let $\pi: N(B_E) \rightarrow B_E$ be the map that collapses fibres. Lemma 8.1 in [17] gives that each cusp in $\partial N(B_E)$ is either an annulus or a disc. Moreover, the cusps that are discs intersect $\pi^{-1}(\partial B_E)$ in two arcs. (Lemma 8.1 in [17] has a couple of hypotheses which are easily checked for the 3-manifold that is a slight enlargement of $N(B_E)$.)

In Lemma 9.1 of [17], it was shown that there are at most $24n^2$ types of Euclidean tile (where n was the arc index of the arc presentation for K). This estimate does not apply here. But the following gives an alternative upper bound.

Lemma 11.5. *The number of approximate star types of tiles of $E - N_2(S' \setminus E)$ is at most*

$$e = 16c_+(\mathcal{H}) + 32c(\mathcal{H}').$$

Proof. Pick a maximal collection of tiles in $E - N_2(S' \setminus E)$, all having the same approximate star type. Consider an outermost tile T of a given approximate star type. One of its vertices must be outermost in its approximate star type. By Propositions 10.10 and 10.11, there are at most $4c_+(\mathcal{H}) + 8c(\mathcal{H}')$ such vertices. Since the tiles we are considering lying in $E - N_2(S' \setminus E)$, their vertices lie in E and so have valence at most 4. Hence, each such vertex is incident to at most 4 tiles. The required bound then follows. \square

Hence, our branched surface B_E has at most this many Euclidean patches.

Lemma 11.6. *The cusps of B_E are annuli and discs in $\partial N(B_E)$. The discs lie in N' .*

Proof. An application of Lemma 8.1 in [17] gives that each cusp is either an annulus or a disc. In the latter case, the disc is a copy of $I \times I$, where each $\{*\} \times I$ lies in a fibre of $N(B_E)$. It was also shown that $\partial I \times I$ lies in $\partial_h N(B_E)$. Now $N(B_E)$ is obtained from N' by adding the spaces between tiles of S' of the same star approximate type. Tiles of S' that are incident to $\partial S'$ have the same approximate type if and only if they are parallel in N' . Thus we deduce that $\partial I \times I$ lies in N' . Hence, $\partial I \times I$ is disjoint from the wedges making up $N' \setminus N$. These wedges intersect each cusp of $N(B_E)$ in regular neighbourhoods of arcs and curves that are transverse to the fibres and disjoint from the horizontal boundary. Hence if any wedge were to intersect a disc cusp of $N(B_E)$, it would do so in a regular neighbourhood of arcs, each lying entirely in the interior of the cusp. Consider the endpoint of one of these arcs. The tiles of S' above and below it are not of the same approximate star type, contradicting the definition of $N(B_E)$. Hence, we deduce that the disc components of the cusps of $N(B_E)$ lie in N' , as claimed. \square

11.2. The monodromy homomorphism. For each component B_0 of B_E , there is a homomorphism $\pi_1(B_0) \rightarrow O(2)$ where $O(2)$ is the group of orthogonal transformations of \mathbb{R}^2 . To define this, one must consider a loop ℓ in B_0 based at some basepoint. As ℓ runs from one patch to another, one may parallel translate a Euclidean 2-plane along with it. By the time that the loop has returned to the basepoint, this 2-plane has been subjected to an orthogonal transformation, which is the image of $[\ell]$ under this homomorphism. Full details can be found in Section 9.3 in [17].

11.3. A finite cover of the branched surface. It is clear that the image of the monodromy homomorphism lies in the subgroup of $O(2)$ that preserves the unit square centred at the origin. This is a dihedral group of order 8. Hence, the kernel of $\pi_1(B_0) \rightarrow O(2)$ is a finite index subgroup of $\pi_1(B_0)$. This corresponds to a finite cover \tilde{B}_0 of B_0 . Let \tilde{B}_E be the union of these covers, as B_0 varies over each component of B_E . This is a branched surface. Indeed, we may take a thickening $N(B_E)$ of B_E , and if we take the corresponding cover of this, the result is a thickening $N(\tilde{B}_E)$ of \tilde{B}_E . Thus \tilde{B}_E is a branched surface in the 3-manifold $N(\tilde{B}_E)$.

The principal type of patches of \tilde{B} are lifts of Euclidean patches of B_E . We term these Euclidean patches of \tilde{B}_E . Since the covering map $\tilde{B}_E \rightarrow B_E$ has degree at

most 8 on each component of \tilde{B}_E , we deduce that \tilde{B}_E has at most $8e$ Euclidean patches.

Let \tilde{S}' be the inverse image of S' in $N(\tilde{B}_E)$. The inverse image of E in $N(\tilde{B})$ is a surface \tilde{E} made up of Euclidean tiles.

One may define a monodromy homomorphism $\pi_1(\tilde{B}_0) \rightarrow O(2)$ for each component \tilde{B}_0 of \tilde{B}_E . But by the construction of \tilde{B}_E , this monodromy homomorphism is trivial. This implies that \tilde{B}_E is transversely orientable, since if it were not, then for some loop ℓ in \tilde{B}_E , the image of ℓ under the monodromy homomorphism would have determinant -1 . We now fix a transverse orientation on \tilde{B}_E , which induces a transverse orientation on \tilde{E} .

11.4. Strategy of the proof. The main result of Section 9 in [17] was Proposition 9.8, which was as follows. If there is a point in \tilde{E} with distance more than $8000n^2$ from $\partial\tilde{E}$ (where again n is the arc index of the link), then \tilde{E} has a torus summand, when viewed as a surface carried by \tilde{B}_E .

This then implied the result that each point in E has distance at most $8000n^2$ from ∂E . For if not, then some point of \tilde{E} would have distance more than $8000n^2$ from $\partial\tilde{E}$ and hence \tilde{E} would have a torus summand. The projection of \tilde{E} to $N(B_E)$ forms a non-zero multiple of E , and the torus projects to a torus carried by B_E . So, we deduce that some non-zero multiple of E has a torus summand. This then implied that some non-zero multiple of S has a torus summand, as a normal surface. (Here, S was a spanning disc for the unknot.) But, in [17], S was arranged so that no multiple of S has a torus summand, because S is, in [17], a ‘boundary-vertex surface’.

Unfortunately, the same strategy does not work in our setting, because we are not assuming that S' is a ‘boundary-vertex’ surface. In particular, there is no obvious reason to prevent some multiple of S' having a torus summand. Instead, we will show the following.

Theorem 11.7. *Suppose that there is some point in E that has distance more than $1196e$ from $S' \setminus E$. Then there is a torus or clean annulus A properly embedded in M that is normal in \mathcal{H} and a homeomorphism $h: M \rightarrow M$ that is a power of a Dehn twist along A that satisfies (1) - (4) of Theorem 11.3.*

11.5. Doubling. Let E_+ be the abstract double of $\text{cl}(E)$ along $\partial S' \cap E$. Let B_{E+} be the double of B_E along ∂B_E . Let $N(B_{E+})$ be the double of $N(B_E)$ along $\pi^{-1}(\partial B_E)$. Note that the cusps of $N(B_{E+})$ are all annuli.

The covers $\tilde{B}_E \rightarrow B_E$ and $\tilde{E} \rightarrow E$ induce covers $\tilde{B}_{E+} \rightarrow B_{E+}$ and $\tilde{E}_+ \rightarrow E_+$. Similarly, we have a cover $N(\tilde{B}_{E+}) \rightarrow N(B_{E+})$.

11.6. Grids and annuli. A *grid* is a subsurface of \tilde{E} or E that is a union of tiles glued together to form a square of odd side length. Since the side length is odd, the centre of this square is at the centre of some tile. For a positive integer r and a point x in \tilde{E} or E , we let $D(x, r)$ denote a grid centred at a tile containing x with side length $2r + 1$.

The proof of Proposition 9.8 in [17] divided into two cases: either there is a ‘short’ closed geodesic ‘far’ from $\partial\tilde{E}$, or there is not. In [17], ‘short’ and ‘far’ were measured in terms of the arc index n . Here, we instead measure in terms of e . We similarly divide into two cases, but we work with E instead \tilde{E} . For this, we require the following lemma.

Lemma 11.8. *Suppose that $E - N_{4784e}(S' \setminus E)$ is non-empty. Let E_+ be the abstract double of $\text{cl}(E)$ along $\partial S' \cap E$. Let W be the inverse image of $S' \setminus E$ in E_+ . Then one of the following holds:*

- (1) *there is a closed geodesic α in $E_+ - N_{1472e}(W)$ with length at most $736e$ such that $N_{1472e}(\alpha)$ is either invariant under the involution of E_+ or disjoint from its image under the involution;*
- (2) *there is a grid in $E_+ - N_{1472e}(W)$ with side length $512e + 1$ that is either invariant under the involution or disjoint from its image under the involution.*

Proof. Since $E - N_{4784e}(S' \setminus E)$ is non-empty, so is $E_+ - N_{4784e}(W)$. Let x' be some point in $E_+ - N_{4784e}(W)$. If x' lies within distance $1840e$ of the copy of $\partial S' \cap E$ in E_+ , then there is a point x on $\partial S' \cap E_+$ that lies in $E_+ - N_{2944e}(W)$. If on the other hand, x' is more than $1840e$ from $\partial S' \cap E_+$, then we set $x = x'$. So, x either lies on $\partial S' \cap E_+$ or is more than $1840e$ from $\partial S' \cap E_+$.

Let r be the injectivity radius at x of E_+ . If $r > 368e$, then there is an embedded grid with side length $512e + 1$ centred at x , since $512e + 1 < 368\sqrt{2}e$. When x lies on $\partial S' \cap E_+$, we may choose the grid so that it is invariant under the involution of E_+ . When x is more than $1840e$ from $\partial S' \cap E_+$ then the grid is disjoint from its image under the involution. Note that the grid lies in $E_+ - N_{2576e}(W)$, and hence (2) is satisfied.

So suppose that r is at most $368e$. Let γ_1 and γ_2 be distinct geodesics starting at x with length r and having the same endpoint y in E_+ . Since r is exactly the injectivity radius at x , these two geodesics patch together exactly at y to form a smooth geodesic there. However, it need not be the case that they patch together to form a smooth closed geodesic at x . Let γ be the concatenation $\gamma_1 \cdot \overline{\gamma_2}$, which is a geodesic based at x . The monodromy of γ lies in $O(2)$. Indeed it lies in $SO(2)$ because E_+ is orientable.

Case 1. $\mu(\gamma)$ is the identity.

Then γ is a closed geodesic that is smooth at x . It has length at most $736e$ and lies in $E_+ - N_{1472e}(W)$. Hence, when x is disjoint from $\partial S' \cap E_+$ and hence $N_{1472e}(\gamma)$ is disjoint from $\partial S' \cap E_+$, γ is the required geodesic.

However, when x lies on $\partial S' \cap E_+$, we need to verify that γ is invariant under the involution of E_+ . In this case let D be the open disc of radius r about x . This is a Euclidean disc that is invariant under the isometric involution of E_+ . Hence, the involution acts by reflection in a straight line through x . We deduce that $D \cap (\partial S' \cap E_+)$ is this straight line. If $\gamma'(0)$ points in the direction of this line or is orthogonal to it, then γ is invariant under the involution as required. So suppose, for a contradiction, that $\gamma'(0)$ makes an angle strictly between 0 and $\pi/2$ with $\partial S' \cap E_+$. Within D , the only intersection between γ and $\partial S' \cap E_+$ is at x . However, γ_1 lies in one half of E_+ and γ_2 lies in the other half, where these two halves are swapped by the involution. We deduce that $\partial S' \cap E_+$ must also lie at the endpoint of γ_1 . It is therefore tangent to ∂D at that point. However, the Euclidean subsurface is made of squares and $\partial S' \cap E_+$ follows the sides of these squares. We deduce that the only possible angles between $\partial S' \cap E_+$ and γ at the endpoint of γ_1 are multiples of $\pi/2$. This is a contradiction.

Case 2. $\mu(\gamma)$ is a rotation of order 2.

Then γ_1 and γ_2 are equal, contradicting the assumption that they are distinct.

Case 3. $\mu(\gamma)$ is a rotation of order 4.

Then γ^4 can be developed onto a closed curve in the Euclidean plane. We may shorten this curve in the plane, as shown in Figure 44. This is achievable within $E_+ - W$ unless the curve hits W . But in this case, x would lie within distance at most $1472e$ from W , which is contrary to assumption. But if we can shrink γ^4 to length 0, it is homotopically trivial in E_+ . Hence so is γ , since the fundamental group of E_+ is torsion free. This is a contradiction, since $\mu(\gamma)$ is non-trivial. \square

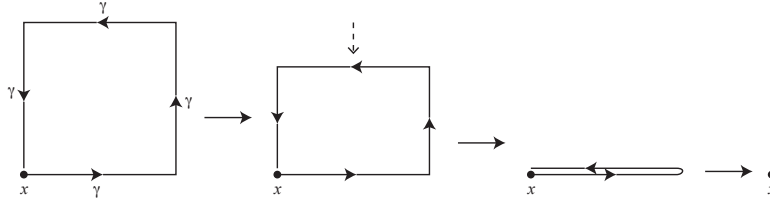


FIGURE 44. Shrinking γ^4

The proof of Theorem 11.3 divides into the two cases in the above lemma. We leave the first case for a while, and now consider the case where there is a grid in $E_+ - N_{1472e}(W)$ with side length $512e + 1$ and that is either invariant under the involution of E_+ or disjoint from its image. This lifts to a grid $D(\tilde{x}, 256e)$ in $\tilde{E}_+ - N_{1472e}(\tilde{W})$ also with side length $512e + 1$, where \tilde{E}_+ is the double of $\text{cl}(\tilde{E})$ along the inverse image of $\partial S' \cap E$, and \tilde{W} is the inverse image of W in \tilde{E}_+ .

11.7. First-return maps. For a subsurface F of \tilde{E} or \tilde{E}_+ , we may define a *first-return map* $r_F: \text{dom}(F) \rightarrow F$ as follows. Its domain of definition is a subsurface $\text{dom}(F)$ of F .

For a point y of F , one considers the fibre α_y of $N(\tilde{B}_E)$ or $N(\tilde{B}_{E_+})$ through y . This intersects F in a number of points. We consider the closest such point along α_y in the specified transverse direction on \tilde{B}_E or \tilde{B}_{E_+} . If there is no point of $\alpha_y \cap F$ in this transverse direction, then $r_F(y)$ is not defined. But otherwise, we take $r_F(y)$ to be this point.

We will consider surfaces F that are a union of tiles. In this situation, the interior of each tile of F is either in $\text{dom}(r_F)$ or is disjoint from it. Each tile of the latter sort is outermost in the patch of $N(\tilde{B}_E)$ or $N(\tilde{B}_{E_+})$ that contains it. Since there are at most $8e$ Euclidean patches of \tilde{B}_E , we deduce that all but at most $8e$ Euclidean tiles of $F \subset \tilde{E}$ lie in $\text{dom}(r_F)$. Similarly, all but at most $16e$ Euclidean tiles of $F \subset \tilde{E}_+$ lie in $\text{dom}(r_F)$.

11.8. The first-return map for large grids. As described above, we are currently considering the case where there is a grid $D(\tilde{x}, 256e)$ in $\tilde{E}_+ - N_{1472e}(\tilde{W})$ with side length $512e + 1$. We are assuming that this grid projects homeomorphically to a grid in E_+ and that it is invariant under the involution of E_+ .

Proposition 11.9. *Every point $\tilde{y} \in D(\tilde{x}, 174e)$ lies in $\text{dom}(r_{D(\tilde{y}, 81e)})$.*

This is very similar to Proposition 9.14 in [17], and their proofs follow the same lines. A key step is the following, which is an analogue of Proposition 9.15 in [17].

Proposition 11.10. *Let D_1, \dots, D_m be a collection of disjoint grids in \tilde{E}_+ , each with side length at least $64e + 1$. Suppose that the restriction of the covering map $\tilde{E}_+ \rightarrow E_+$ to each D_i is a homeomorphism onto its image. Assume also that each D_i is either invariant under the involution of \tilde{E}_+ or disjoint from its image under the involution. Let d_i be the supremal distance of the points in D_i from ∂D_i at which $r_{\tilde{E}_+}$ is not defined. Then $\sum_i d_i \leq 16e$.*

Note that Proposition 11.10 has a hypothesis that was not present in Proposition 9.15 in [17]. This is that the restriction of the covering map $\tilde{E}_+ \rightarrow E_+$ to each D_i is a homeomorphism onto its image. This hypothesis is required because here we not assuming that S or S' is a ‘boundary-vertex’ surface.

Proof. Consider the cusps of \tilde{B}_{E_+} , and let $\partial_- C$ denote the 1-manifold in \tilde{E}_+ lying directly above these cusps. By ‘above’, we mean that the transverse orientation on \tilde{E}_+ at $\partial_- C$ points towards the cusps. Note that $\partial_- C$ divides \tilde{E}_+ into two subsurfaces: $\text{dom}(r_{\tilde{E}_+})$ and $E_+ \setminus \text{dom}(r_{\tilde{E}_+})$. Transversely orient $\partial_- C$ in \tilde{E}_+ so that it points towards $\text{dom}(r_{\tilde{E}_+})$. The 1-manifold $D_i \cap \partial_- C$ consists of properly embedded arcs and simple closed curves.

We claim that each simple closed curve of $D_i \cap \partial_- C$ points into the disc in D_i that it bounds. This is an analogue of Claim 2 in the proof of Proposition 9.15 in [17]. Suppose that, on the contrary, there is a simple closed curve β_- in $D_i \cap \partial_- C$ that points out of the disc that it bounds in the grid D_i . This is an embedded curve in the grid D_i and hence it projects to an embedded curve $\bar{\beta}_-$ in E_+ . Therefore, the cusp of \tilde{B}_{E_+} directly below β_- projects homeomorphically to a cusp of B_+ . Let β_+ be the curve in \tilde{E}_+ directly below β_- . Then β_+ projects homeomorphically to a simple closed curve $\bar{\beta}_+$ in E_+ . This restricts to a simple closed curve or properly embedded arc γ_+ in E . Similarly, $\bar{\beta}_-$ restricts to a simple closed curve or properly embedded arc γ_- in E . The disc bounded by $\bar{\beta}_-$ restricts to a disc D_- in S' separated off by γ_- .

Suppose first that γ_+ and γ_- are simple closed curves. Thus the union of the portions of the fibres between them is an annulus A containing an annular cusp of $N(B_E)$. The union of $A \cup D_-$ is a disc bounded by γ_+ lying in M . Since S is incompressible, we deduce that γ_+ bounds a disc D_+ in S . The boundary of this disc lies in S' , and hence all of D_+ lies in S' . Then $D_- \cup A \cup D_+$ is a sphere that bounds a ball in M .

Now suppose that γ_+ and γ_- are arcs. Thus the union of the portions of the fibres between them is a disc D containing a disc cusp of $N(B_E)$. By Lemma 11.6, D lies in N' . Hence, after a small isotopy, $D_+ \cup D$ is a disc in M' , with boundary consisting of an arc in S' and an arc in $\partial M'$, and with interior disjoint from S' . By the boundary-incompressibility of S' , we deduce that γ_+ separates off a clean disc D_+ in S' . Hence, $D_- \cup A \cup D_+$ is a disc properly embedded in M' disjoint from the pattern. So by the boundary-irreducibility of M' , we deduce that $D_- \cup A \cup D_+$ separates off a 3-ball disjoint from the pattern. The interior of this ball is disjoint from S' and S .

We can assume that D_- and D_+ have the same binding weight, as otherwise we could isotope S' and S , replacing one disc by the other to reduce the binding weight of S' . Note that in the case where γ_+ and γ_- are simple closed curves, there may be components of intersection between $S \setminus S'$ and the ball bounded by

$D_- \cup A \cup D_+$. But we may replace these by discs parallel to D_- or D_+ . This might introduce new 2-sphere components of $S \setminus S'$, in which case these are discarded.

Now replace as many parallel copies of D_+ as possible by parallel copies of D_- . Eventually, this must stop when some approximate tile type of S' has been removed. Hence, we have found a surface S'_2 that is pattern-isotopic to S' having the same binding weight, but where $\text{approx-type}(S'_2) < \text{approx-type}(S')$, contrary to an assumption of Theorem 11.3. This proves the claim.

We claim that any point of $D_i \cap \partial_- C$ that is furthest from ∂D_i lies on an arc component of $D_i \cap \partial_- C$. This is an analogue of Claim 1 in the proof of in the proof of Proposition 9.15 in [17]. Cut D_i along the arc components of $D_i \cap \partial_- C$ and let D'_i be the disc containing the centre of D_i . Since the arc components have length at most $32e$, D'_i contains a grid with side length $32e + 1$. So, D'_i contains at least $(32e + 1)^2$ tiles. We will rule out the possibility that there are any simple closed curves of $\partial_- C$ in D'_i . Let δ be the union of those simple closed components of $D'_i \cap \partial_- C$ that are outermost, in other words, that do not lie within another simple closed component of $D'_i \cap \partial_- C$. The total length of δ is at most $32e$, and so the total number of tiles that it can bound is at most $(8e)^2$. But by the above claim, $\text{dom}(r_{\tilde{E}_+}) \cap D'_i$ lies within γ . So at least $(32e + 1)^2 - (8e)^2$ tiles do not lie in $\text{dom}(r_{\tilde{E}_+})$. However, $r_{\tilde{E}_+}$ is defined on all but at most $16e$ tiles of \tilde{E}_+ . This is a contradiction, proving the claim.

We are now in a position to complete the proof of Proposition 11.10. Consider a point on $D_i \cap \partial_- C$ that is furthest from ∂D_i . By the above claim, this lies on an arc component of $D_i \cap \partial_- C$, and hence its distance from ∂D_i is at most $16e$. So, we deduce that the grid with side length $32e + 1$ with the same centre as D_i has interior disjoint from $\partial_- C$ and so either lies in $\text{dom}(r_{\tilde{E}_+})$ or disjoint from $\text{dom}(r_{\tilde{E}_+})$. But this grid contains at least $(32e + 1)^2$ tiles whereas $r_{\tilde{E}_+}$ is defined on all but at most $16e$ tiles. So we deduce that a point in $D_i \setminus \text{dom}(r_{\tilde{E}_+})$ furthest from ∂D_i has distance from ∂D_i that is at most half the length of $\partial_- C \cap D_i$. So, d_i is at most half the length of $\partial_- C \cap D_i$. Thus, $\sum_i d_i$ is at at most half the length of $\partial_- C$, and hence at most $16e$. \square

Proof of Proposition 11.9. The proof is a direct analogue of that of Proposition 9.14 in [17]. Let \tilde{y} be a point in $D(\tilde{x}, 81e)$. We will show that there is a non-negative integer m and a map $D(\tilde{y}, 32e) \times [0, m' + 1] \rightarrow N(\tilde{B}_{E+})$ such that

- (1) the restriction to $D(\tilde{y}, 32e) \times [0, m' + 1]$ is an embedding;
- (2) the image of $D(\tilde{y}, 32e) \times \{0\}$ is $D(\tilde{y}, 32e)$;
- (3) the inverse image of \tilde{E}_+ is $D(\tilde{y}, 32e) \times ([0, m' + 1] \cap \mathbb{Z})$;
- (4) for each point z in $D(\tilde{x}, 32e)$, the image of $\{z\} \times [0, m' + 1]$ is a subset of a fibre of $N(\tilde{B}_{E+})$;
- (5) the image of $D(\tilde{y}, 32e) \times \{m' + 1\}$ has non-empty intersection with $D(\tilde{y}, 48e)$.

We define two increasing sequences m_i and k_i of non-negative integers and maps $D(\tilde{y}, 48e - k_i) \times [0, m_i] \rightarrow N(\tilde{B}_{E+})$ with properties analogous to (1)-(4) above. The procedure starts with $k_0 = 0$ and $m_0 = 0$. Then the following procedure is performed:

- (1) If $D(\tilde{y}, 48e - k_i) \times \{m_i\}$ intersects $D(\tilde{y}, 48e)$ and $m_i > 0$, then the procedure terminates; otherwise we pass to the next step.

- (2) If $r_{\tilde{E}_+}$ is defined on all of $D(\tilde{x}, 48e - k_i) \times \{m_i\}$, let $r_{\tilde{E}_+}(D(\tilde{x}, 48e - k_i) \times \{m_i\})$ be $D(\tilde{x}, 48e - k_i) \times \{m_i + 1\}$, increase m_i by 1, and between these two grids there is a product which is added to the product region. Then return to step (1).
- (3) Suppose that $r_{\tilde{E}_+}$ is not defined for all of $D(\tilde{x}, 48e - k_i) \times \{m_i\}$. This means that there is at least one cusp of $N(\tilde{B}_{E_+})$ below $D(\tilde{x}, 48e - k_i) \times \{m_i\}$. Let d_i be the maximal distance of such cusps from the boundary of $D(\tilde{x}, 48e - k_i) \times \{m_i\}$. Applying Proposition 11.10 to the images of $D(\tilde{x}, 48e - k_0) \times \{m_0\}, \dots, D(\tilde{x}, 48e - k_i) \times \{m_i\}$ gives that $\sum_{j=1}^i d_j$ is at most $16e$. Set $k_{i+1} = \sum_{j=1}^i d_j$, and define $m_{i+1} = m_i + 1$. Increase i by 1 and return to step (1).

Note that our application of Proposition 11.10 in step (3) is permitted. This is because restriction of the covering map $\tilde{E}_+ \rightarrow E_+$ to each of the grids $D(\tilde{x}, 48e - k_0) \times \{m_0\}, \dots, D(\tilde{x}, 48e - k_i) \times \{m_i\}$ is an embedding. For suppose that this was not the case, and that two points in one of these grids had the same image in E_+ . Then, translating these two points along the fibres towards $D(\tilde{x}, 48e)$, we would deduce that two distinct points in $D(\tilde{x}, 48e)$ had the same image in E_+ . However, the restriction of $\tilde{E}_+ \rightarrow E_+$ to $D(\tilde{x}, 48e)$ is a homeomorphism onto $D(x, 48e)$, and in particular is injective. Thus, we have shown that there is the required map $D(\tilde{y}, 32e) \times [0, m' + 1] \rightarrow N(\tilde{B}_{E_+})$.

Let \tilde{y} be a point in $D(\tilde{x}, 174e)$. Around \tilde{y} , there is the grid $D(\tilde{y}, 32e) \subset D(\tilde{x}, 256e)$. Let $D(\tilde{y}, 32e) \times [0, m' + 1] \rightarrow N(\tilde{B}_{E_+})$ be the above map. By (5), there is a point z in $D(\tilde{y}, 32e)$ such that $\{z\} \times \{m' + 1\}$ has image equal to a point in $D(\tilde{y}, 48e)$. Now $\{\tilde{y}\} \times \{m' + 1\}$ lies in the grid $D(\{z\} \times \{m' + 1\}, 32e)$ and hence lies in $D(\tilde{y}, 1 + 80e) \subset D(\tilde{y}, 81e)$. There may be other points of $D(\tilde{y}, 81e)$ on the fibre between $\{\tilde{y}\} \times \{0\}$ and $\{\tilde{y}\} \times \{m' + 1\}$. But the first return map for $D(\tilde{y}, 81e)$ is certainly defined at \tilde{y} . \square

11.9. Translation invariance of the first-return map. As in the previous subsection, let $D(\tilde{x}, 256e)$ be a grid in $\tilde{E}_+ - N_{1472e}(\tilde{W})$. We are assuming that $D(\tilde{x}, 256e)$ is either invariant under the involution of \tilde{E}_+ or disjoint from its image under the involution. We are also assuming that $D(\tilde{x}, 256e)$ projects homeomorphically to a grid in E_+ . By Proposition 11.9, for each $\tilde{y} \in D(\tilde{x}, 174e)$, the points \tilde{y} and $r_{D(\tilde{y}, 81e)}(\tilde{y})$ both lie in the disc $D(\tilde{x}, 256e)$ and so there is a well-defined vector $v_{\tilde{y}}$ in $T_{\tilde{x}}\tilde{E}_+$ taking \tilde{y} to $r_{D(\tilde{y}, 81e)}(\tilde{y})$.

Proposition 11.11. *The vector field $\{v_{\tilde{y}} : \tilde{y} \in D(\tilde{x}, 48e)\}$ is covariant constant.*

This was Proposition 9.16 in [17], with essentially the same proof, which we briefly sketch. First consider moving \tilde{y} within a tile. Then $D(\tilde{y}, 81e)$ remains the same grid as \tilde{y} varies. The map $r_{D(\tilde{y}, 81e)}$ sends \tilde{y} to a point in $N(\tilde{B}_{E_+})$ directly below it. Because \tilde{B}_{E_+} has trivial monodromy, the vector in $D(\tilde{y}, 80e)$ from \tilde{y} to its image remains unchanged.

Now consider what happens when \tilde{y} moves to a point \tilde{y}' in an adjacent tile. Then $D(\tilde{y}, 81e)$ also moves, and a new tile of $D(\tilde{y}', 81e)$ appears below \tilde{y}' . This is again because \tilde{B}_{E_+} has trivial monodromy. The image of \tilde{y}' under $r_{D(\tilde{y}', 81e)}$ lies in this tile, and again the vector in $D(\tilde{y}', 81e)$ from \tilde{y}' to its image remains unchanged.

This argument then gives the following result.

Proposition 11.12. *There is a positive integer m with the following property. For any point $\tilde{y} \in D(\tilde{x}, 174e)$, there is a map $D(\tilde{y}, 81e) \times [0, m+1] \rightarrow N(\tilde{B}_{E+})$ such that*

- (1) *the restriction to $D(\tilde{y}, 81e) \times [0, m+1]$ is an embedding;*
- (2) *the image of $D(\tilde{y}, 81e) \times \{0\}$ is $D(\tilde{y}, 81e)$;*
- (3) *the inverse image of \tilde{E}_+ is $D(\tilde{y}, 81e) \times ([0, m+1] \cap \mathbb{Z})$;*
- (4) *for each point z in $D(\tilde{x}, 81e)$, the image of $\{z\} \times [0, m+1]$ is a subset of a fibre of $N(\tilde{B}_{E+})$;*
- (5) *the image of $D(\tilde{y}, 81e) \times \{m+1\}$ has non-empty intersection with $D(\tilde{y}, 81e)$.*

11.10. The monodromy of paths. In Section 11.2, we defined the monodromy of a closed loop in B_{E+} . We now use this to define the monodromy of certain paths in E_+ .

Let p be a path in E_+ such that $\pi p(0) = \pi p(1)$ where $\pi: N(B_{E+}) \rightarrow B_{E+}$ is the projection map that collapses each fibre to a point. Then there is a unique embedded path ϵ in a fibre joining $p(1)$ to $p(0)$. We define the *monodromy* $\mu(p)$ of p to be the monodromy of $p.\epsilon$. Note that if p is a closed loop, then the two notions of monodromy coincide and so no confusion can arise. Note also that the monodromy of a path p is unchanged if a homotopy is applied to p in E_+ that fixes its endpoints.

11.11. Paths with close endpoints. Let p be a path in $D(x, 174e)$ such that $p(0)$ and $p(1)$ lies in the same fibre of $N(B_{E+})$ and with the property that p lies in $D(p(0), 81e)$. Let ϵ be the part of the fibre between these two points. Suppose that ϵ lies in the image of $D(p(0), 81e) \times [0, m+1]$ or $D(p(0), 81e) \times (0, m+1]$ in $N(B_{E+})$ for the map described in Proposition 11.12 composed with the covering map $N(\tilde{B}_{E+}) \rightarrow N(B_{E+})$. Then we say that p has *close endpoints*. Our aim is to prove the following.

Proposition 11.13. *No path p in $D(x, 81e)$ has close endpoints.*

It suffices to consider the case where p is a geodesic in $D(x, 174e)$.

Lemma 11.14. *Let p be a geodesic in $D(x, 174e)$ with close endpoints. Then the following are homotopically trivial loops in $N(B_{E+})$:*

$$\begin{aligned} p.\epsilon.(-\mu(p)p).-\epsilon & \quad \text{if } \det(\mu(p)) = 1; \\ p.\epsilon.(-\mu(p)p).\epsilon & \quad \text{if } \det(\mu(p)) = -1. \end{aligned}$$

Here, $-\mu(p)p$ denotes the geodesic that starts at $p(0)$ with initial vector $-\mu(p)p'(0)$ and with length equal to that of p . Also, $-\epsilon$ denotes the embedded arc in the same fibre as ϵ that ends at $\epsilon(1)$, that intersects E_+ the same number of times as ϵ and that intersects ϵ only at $\epsilon(1)$.

Proof. Let \tilde{p} be a lift of p to $D(\tilde{x}, 174e)$. Let $\tilde{\epsilon}$ be the lift of ϵ starting at the endpoint of \tilde{p} . Then $\tilde{p}.\tilde{\epsilon}$ lies in the image of $D(p(0), 81e) \times [0, m+1]$ or $D(p(0), 81e) \times (0, m+1]$. The lift of $(-\mu(p)p).-\epsilon$ or $(-\mu(p)p).\epsilon$ (as appropriate) starting at the endpoint of $\tilde{\epsilon}$ runs along a path parallel to \tilde{p} but in reverse and then closes up, ending at $\tilde{p}(0)$. Hence, the composition of these 4 paths is a homotopically trivial loop in $D(p(0), 81e) \times [0, m+1]$ or $D(p(0), 81e) \times (0, m+1]$. Thus, the image of these 4 paths is a homotopically trivial loop in $N(B_{E+})$. \square

Therefore, when p is a path in $D(x, 174e)$ with close endpoints, the following paths are homotopic relative to their endpoints:

$$\begin{aligned} p.\epsilon &\simeq \overline{-\epsilon.(-\mu(p)p)} & \text{if } \det(\mu(p)) = 1; \\ p.\epsilon &\simeq \bar{\epsilon}.(-\mu(p)p) & \text{if } \det(\mu(p)) = -1. \end{aligned}$$

Here, the overline denotes the reverse of a path.

11.12. Paths in grids with close endpoints have positive determinant.

Lemma 11.15. *Any path in $D(x, 81e)$ with close endpoints has monodromy with positive determinant.*

Proof. Suppose that the monodromy of p has negative determinant. Therefore, $p.\epsilon$ is a homotopically non-trivial loop in $N(B_{E+})$. Therefore, $p.\epsilon.p.\epsilon$ is also, since $N(B_{E+})$ has torsion free fundamental group (being a compact 3-dimensional submanifold of the 3-sphere). But

$$p.\epsilon.p.\epsilon \simeq p.\epsilon.\bar{\epsilon}.(-\mu(p)p) \simeq p.(-\mu(p)p).$$

However $p.(-\mu(p)p)$ remains within $D(x, 174)$ which is a grid and which therefore does not contain any homotopically non-trivial loops. \square

11.13. Paths in grids with close endpoints have trivial monodromy.

Lemma 11.16. *Any path in $D(x, 81e)$ with close endpoints has trivial monodromy.*

Proof. Let p be a path in $D(x, 81e)$ with close endpoints and non-trivial monodromy. By Lemma 11.15, this monodromy has positive determinant and is therefore a rotation of order 2 or 4. By assumption, the endpoints of p lie in the same fibre of $N(B_{E+})$. Indeed, the part of the fibre between $p(0)$ and $p(1)$ lies in the image of $D(\tilde{x}, 81e) \times [0, m+1]$ or $D(\tilde{x}, 81e) \times (0, m+1]$. We may homotope p relative to its endpoints to a geodesic. Let p_2 be the geodesic starting at $p(1)$, with length and direction given by $\mu(p)(p'(0))$ translated to $p(1)$. Then p_2 fellow travels with p , in the sense that they have the same images under the projection map $N(B_{E+}) \rightarrow B_{E+}$. Define p_3 and p_4 similarly. Then $p.p_2.p_3.p_4$ has distinct endpoints in the same fibre of $N(B_{E+})$. However, $p.p_2.p_3.p_4$ is a homotopically trivial loop in $D(x, 174e)$, and in particular, it ends where it starts. This is a contradiction. \square

Proof of Proposition 11.13. Suppose that on the contrary there is a path p in $D(x, 81e)$ with close endpoints. Let \tilde{p} be a lift to $D(\tilde{x}, 81e)$. Let ϵ be the path in $N(B)$ between $p(0)$ and $p(1)$. By Lemma 11.16, $p.\epsilon$ has trivial monodromy, and so lifts to a loop in $N(\tilde{B})$ starting at $\tilde{p}(0)$. So either $\tilde{p}(1)$ is a point in $D(\tilde{p}(0), 81e)$ that lies strictly between $\tilde{p}(0)$ and $r_{D(\tilde{p}(0), 81e)}(\tilde{p}(0))$, or $\tilde{p}(0)$ is a point in $D(\tilde{p}(1), 81e)$ that lies strictly between $\tilde{p}(1)$ and $r_{D(\tilde{p}(1), 81e)}(\tilde{p}(1))$. This contradicts the definition of the first-return map. \square

11.14. An embedded annulus in $N(\tilde{B}_+)$. We can now form an embedded annulus \tilde{C} in $N(\tilde{B}_+)$. This is constructed as follows.

Let $\tilde{\alpha}$ be the geodesic in $D(\tilde{x}, 81e)$ from \tilde{x} to $\tilde{x}' = r_{D(\tilde{x}, 81e)}(\tilde{x})$. Let $\tilde{\beta}$ be a geodesic starting at \tilde{x} with length $32e$ that is parallel to a side of $D(\tilde{x}, 81e)$ and that has angle between $\pi/2$ and $3\pi/4$ from $\tilde{\alpha}$. Let $\tilde{\beta}'$ be the result of translating $\tilde{\beta}$ using the vector $v_{\tilde{x}}$, so that it runs through \tilde{x}' . The region \tilde{Q} between $\tilde{\beta}$ and $\tilde{\beta}'$ is a Euclidean parallelogram. By the way that we have chosen $\tilde{\beta}$, \tilde{Q} lies within

$D(\tilde{x}, 81e)$. All of interior angles lie between $\pi/4$ and $3\pi/4$. Note also that when $\tilde{\alpha}$ is invariant under the involution of \tilde{E}_+ , so too is \tilde{Q} .

The two sides $\tilde{\beta}$ and $\tilde{\beta}'$ of \tilde{Q} have the same image in \tilde{B}_{E+} . Hence, we can insert a rectangle \tilde{R} between these two sides of \tilde{Q} , that is vertical in $N(\tilde{B}_{E+})$ in the sense that it is a union of arcs that lie in fibres of $N(\tilde{B}_{E+})$. The union of \tilde{Q} and \tilde{R} is an annulus \tilde{C} .

When $D(\tilde{x}, 256e)$ is invariant under the involution of $N(\tilde{B}_{E+})$, then \tilde{x} and \tilde{x}' are fixed points. Hence in this case, $\tilde{\alpha}$ is fixed pointwise by the involution and so \tilde{C} is invariant under the involution.

Lemma 11.17. *\tilde{C} is embedded in $N(\tilde{B}_{E+})$.*

Proof. The annulus \tilde{C} consists of the parallelogram \tilde{Q} and a vertical rectangle \tilde{R} . Both \tilde{Q} and \tilde{R} are embedded. So the only way that \tilde{C} could fail to be embedded is if there is a point z_1 in \tilde{Q} and a point z_2 in the interior of \tilde{R} that coincide in $N(\tilde{B}_{E+})$. Let z_3 be a point in \tilde{R} that lies in the same fibre as z_2 and that also lies in \tilde{Q} . Then $r_{\tilde{Q}}(z_3)$ is a point lying between z_3 and z_2 (possibly equalling z_2 but not z_3). Hence, $r_{D(z_3, 81e)}(z_3)$ also lies between z_3 and z_2 . So the number of points of \tilde{E}_+ strictly between z_3 and $r_{D(z_3, 81e)}(z_3)$ is less than m , the constant from Proposition 11.12, which is a contradiction. \square

Lemma 11.18. *\tilde{C} is disjoint from its covering translates under the action of the covering group for $N(\tilde{B}_{E+}) \rightarrow N(B_{E+})$.*

Proof. Suppose that on the contrary, \tilde{C} intersects some image $g(\tilde{C})$ under some non-trivial covering transformation g . This covering translate is made up of a covering translate $g(\tilde{Q})$ of \tilde{Q} and a covering translate $g(\tilde{R})$ of \tilde{R} . Let z in \tilde{C} be a point of intersection between \tilde{C} and $g(\tilde{C})$. There are several possible cases to consider.

Suppose first that z lies in \tilde{Q} and $g(\tilde{Q})$. Then $g^{-1}(z)$ lies in \tilde{Q} , and there is a path \tilde{p} in \tilde{Q} from $g^{-1}(z)$ to z . This projects a closed loop in $N(B_{E+})$ with non-trivial monodromy. However, this path lies in the image of \tilde{Q} , which lies in $D(x, 81e)$, and every closed loop in this grid has trivial monodromy. This is a contradiction.

Now suppose that z lies in \tilde{Q} and $g(\text{int}(\tilde{R}))$. Now $g(\tilde{R})$ consists of a union of arcs, each of which is a subset of a fibre. Let z_2 be an endpoint of the arc containing z , with z_2 lying in $g(\tilde{Q})$. So, $g^{-1}(z_2)$ lies in \tilde{Q} . Let \tilde{p} be a geodesic in \tilde{Q} joining z to $g^{-1}(z_2)$. Let p be the image of \tilde{p} under the projection map to $N(B)$. Then p is a path in $D(x, 81e)$ with close endpoints, contradicting Proposition 11.13. An analogous argument also rules out the case where z lies in $g(\tilde{Q})$ and $\text{int}(\tilde{R})$.

The final case is where z lies in $\text{int}(\tilde{R})$ and $g(\text{int}(\tilde{R}))$. However, if we move z along the fibre it lies in, then we can find a point z' of intersection between \tilde{C} and $g(\tilde{C})$ that lies in \tilde{Q} or $g(\tilde{Q})$ (or both). This has been ruled out in the previous cases. \square

Thus, \tilde{C} projects to an embedded annulus C in $N(B_{E+})$. This is either disjoint from its image under the involution of $N(B_{E+})$ or invariant under this involution.

11.15. Return to the annular case. In Lemma 11.8, we showed that either there is a closed geodesic in $E_+ - N_{1472e}(W)$ with length at most $736e$ or there is a grid $D(x, 256e)$ in $E_+ - N_{1472e}(W)$ with side length $512e + 1$. We have so far focused on the second case. However, we now consider the first case.

Let α be a simple closed geodesic in $E_+ - N_{1472e}(W)$ with length at most $736e$, such that $N_{1472e}(\alpha)$ is either invariant under the involution or disjoint from its image under the involution.

The following is a version of Lemma 9.9 in [17], which essentially the same proof, which is omitted.

Lemma 11.19. *Suppose that α is a simple closed geodesic in $E_+ - N_{1472}(W)$ with length at most $736e$. Suppose also that the component of E_+ containing α is not a torus with binding weight at most $23552e^2$. Then for all $r \leq 16e$, $N_r(\alpha)$ is a Euclidean annulus with core curve α .*

Note that the monodromy along α is in $SO(2)$ because E_+ is orientable. Note also that the monodromy preserves the unit tangent vector to α since α is a closed geodesic. So the monodromy along α is in fact trivial. Hence, the inverse image of α in \tilde{E}_+ is a disjoint union of copies of α . Let $\tilde{\alpha}$ be one such component.

Let \tilde{x} be a point on $\tilde{\alpha}$, and let $\tilde{\beta}$ be a geodesic through \tilde{x} that is orthogonal to $\tilde{\alpha}$ and that runs for distance $16e$ in both directions from \tilde{x} . Let \tilde{Q} be the rectangle that results from cutting $N_{16e}(\tilde{\alpha})$ along $\tilde{\beta}$. Let Q be its image in E_+ . We refer to the edges of \tilde{Q} that are identified in \tilde{E}_+ by $\tilde{\beta}$ and $\tilde{\beta}'$. Note that \tilde{Q} is either invariant under the involution of $N(\tilde{B}_{E+})$ or disjoint from its image under the involution.

We now consider the two cases of Lemma 11.8 simultaneously.

11.16. Another vertical rectangle. The following was proved in Section 9.8 of [17].

Lemma 11.20. *There is a translate $\tilde{\gamma}$ of $\tilde{\alpha}$ that starts on $\tilde{\beta}$ and ends on $\tilde{\beta}'$ and a map $\tilde{\gamma} \times [0, K + 1] \rightarrow N(\tilde{B}_{E+})$, for some integer $K \geq 0$, with the following properties:*

- (1) *the restriction to $(\tilde{\gamma} - \partial\tilde{\gamma}) \times [0, K + 1]$ is an embedding;*
- (2) *if x_1 and x_2 are the start and endpoints of $\tilde{\gamma}$, then the restriction to each $x_i \times [0, K + 1]$ is an embedding;*
- (3) *the images of (x_1, y_1) and (x_2, y_2) coincide if and only if $y_1 - y_2$ is the integer m from Proposition 11.12;*
- (4) *the image of $\tilde{\gamma} \times \{0\}$ is $\tilde{\gamma}$;*
- (5) *the inverse image of \tilde{E}_+ is $\tilde{\gamma} \times ([0, K + 1] \cap \mathbb{Z})$;*
- (6) *for each point z in $\tilde{\gamma}$, the image of $\{z\} \times [0, K + 1]$ is a subset of a fibre of $N(\tilde{B}_{E+})$;*
- (7) *the image of $\tilde{\gamma} \times [0, K + 1]$ is disjoint from \tilde{Q} ;*
- (8) *there is a point y in the interior of $\tilde{\gamma}$ such that the image of $y \times \{K + 1\}$ is equal to a point y' in \tilde{Q} .*

Moreover when $\tilde{\alpha}$ is invariant under the involution of $N(\tilde{B}_{E+})$, then so are $\tilde{\gamma}$, y and y' .

Let \tilde{V} be the image of $\tilde{\gamma} \times [0, K + 1]$ in $N(\tilde{B}_{E+})$. By construction, its interior is disjoint from \tilde{C} .

Lemma 11.21. *$\tilde{C} \cup \tilde{V}$ is disjoint from its covering translates.*

Proof. Suppose that, on the contrary, the image of $\tilde{C} \cup \tilde{V}$ under some non-trivial covering transformation g has non-empty intersection with $\tilde{C} \cup \tilde{V}$. Let z be a point in the intersection.

We know by Lemma 11.18 that $g(\tilde{C}) \cap \tilde{C}$ is empty. So, there are three remaining cases: z can lie in $g(\tilde{C}) \cap \tilde{V}$, $g(\tilde{V}) \cap \tilde{C}$ or $g(\tilde{V}) \cap \tilde{V}$. In the third case, we can slide z vertically in \tilde{V} until it lies in $g(\tilde{C})$ or \tilde{C} . So if the third case holds, then so does one of the first two cases. Furthermore, the first two cases are essentially identical. So, we may assume that $z \in g(\tilde{C}) \cap \tilde{V}$.

Consider a horizontal path \tilde{q} in \tilde{V} starting at z , running in the direction of $\tilde{\alpha}$. (See Figure 45.) Its image $g^{-1}(\tilde{q})$ starts at $g^{-1}(z) \in \tilde{C}$ and runs along a horizontal path in \tilde{Q} . Since we have applied a non-trivial covering transformation, this path $g^{-1}(\tilde{q})$ has had a non-trivial element of $O(2)$ applied to it. By the time we have reached the end of \tilde{q} , we reach a point that is in the same fibre as the start of \tilde{q} and there are m points of \tilde{E} directly between the start and end of \tilde{q} , where m is the integer from Proposition 11.12. Hence, the same is true of $g^{-1}(\tilde{q})$. But $g^{-1}(z)$ lies in the grid $D(\tilde{x}, 81e)$. So, by Proposition 11.11, the first point of $D(g^{-1}(z), 81e)$ that lies in the same fibre as $g^{-1}(z)$ in the specified transverse direction differs from $g^{-1}(z)$ by translation in the direction of $\tilde{\alpha}$. So, $g^{-1}(\tilde{q})$ must run in the same direction as $\tilde{\alpha}$. Hence, g must preserve the $\tilde{\alpha}$ direction and the specified transverse orientation on $N(\tilde{B}_{E+})$. But this implies that g represents the trivial element of $O(2)$, and hence g was the identity covering transformation. This is a contradiction. \square

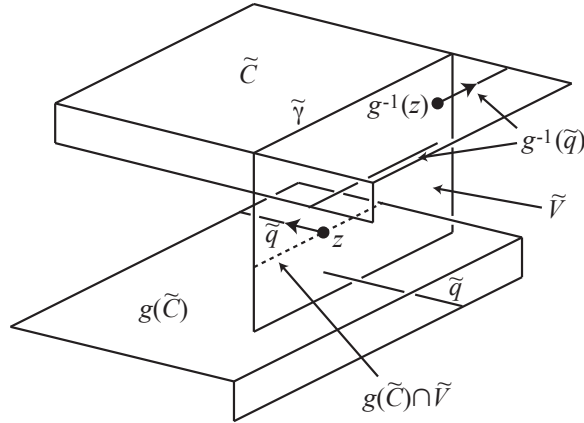


FIGURE 45. An arrangement from the proof of Lemma 11.21

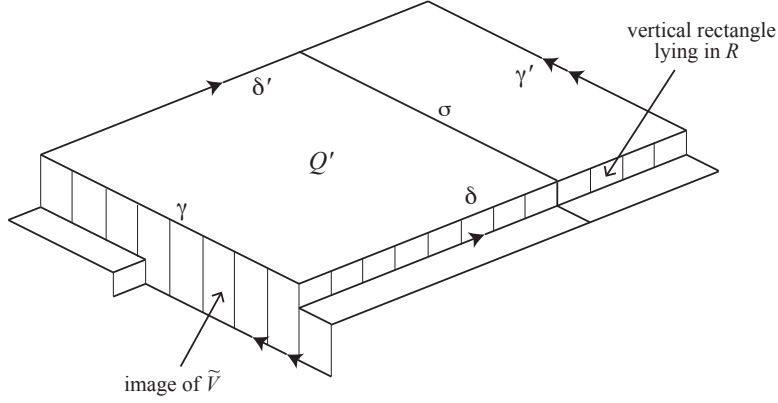
11.17. An embedded torus in $N(B_+)$. Let γ be the image of $\tilde{\gamma}$ in Q . Let γ' be its translate that goes through y' . Let δ be the subset of β that lies between $\beta \cap \gamma$ and $\beta \cap \gamma'$. Define $\delta' \subset \beta'$ similarly.

Then $\gamma \cup \delta \cup \gamma' \cup \delta'$ forms an embedded parallelogram Q' in $N(B_{E+})$. The sides δ and δ' follow the same itinerary in $N(B_{E+})$, and have a portion of the vertical rectangle R between them. Thus, these close up to form an embedded annulus.

The opposite sides of this annulus have the image of \tilde{V} between them. Attaching \tilde{V} to these two boundary components gives an embedded torus T in $N(B_{E+})$. (See Figure 46.)

Let σ be a curve in T that is the union of a parallel copy of γ and a fibre in R .

This torus T is either disjoint from its image under the involution of $N(B_{E+})$ or invariant under the involution. In the case where T is disjoint from its image, let

FIGURE 46. The torus T

T' be the union of T and this image. When T is invariant under the involution, let $T' = T$.

11.18. Removing a multiple of the summand T' . This torus or tori T' is a summand of E_+ , because of the following lemma, which is Lemma 8.2 in [17].

Lemma 11.22. *Let S_1 and S_2 be normal surfaces carried by a branched surface. Suppose that in each patch of the branched surface, the weight of S_2 is at least the weight of S_1 . Then S_1 is a summand of S_2 .*

In the case where \tilde{x} lies in the grid $D(\tilde{x}, 256e)$, let m be the number of sheets of \tilde{E}_+ strictly between \tilde{x} and $r_{D(\tilde{x}, 81e)}(\tilde{x})$. (This is the constant from Proposition 11.12.) In the case where $\tilde{\alpha}$ is a closed geodesic, let m be zero.

Lemma 11.23. $(m+1)T'$ is a summand of E_+ .

Proof. When $m = 0$, this is a consequence of Lemma 11.22. So let us assume that \tilde{x} lies in the grid $D(\tilde{x}, 256e)$. By Lemma 11.22, it suffices to show that $(m+1)T'$ is carried by the branched surface $N(B_{E_+})$ and that in each patch, the weight of E_+ is at least that the weight of $(m+1)T'$. At each point z in \tilde{Q} , $r_{D(z, 81e)}(z)$ and z lie in the same patch of $N(\tilde{B}_{E_+})$. If any part of \tilde{E}_+ lies between them, this also lies in the same patch. Just as \tilde{Q} was used to create the torus T' , so these parts of E_+ therefore may be used to create the tori $(m+1)T'$, which is then a summand of E_+ . (See Figure 47.) \square

Let \overline{E}_+ be the surface carried by B_{E_+} satisfying $E_+ = \overline{E}_+ + (m+1)T'$.

11.19. Dehn twisting along T' .

Proposition 11.24. *The surface \overline{E}_+ is obtained from E_+ by Dehn twisting some number of times around T' .*

We will need the following.

Lemma 11.25. *Let \overline{S} be a surface carried by a branched surface B , and let T be a torus carried by B . Suppose that $\overline{S} \cap T$ consists of k curves that are essential on T and so that the summation directions are all coherently oriented around T . Then*

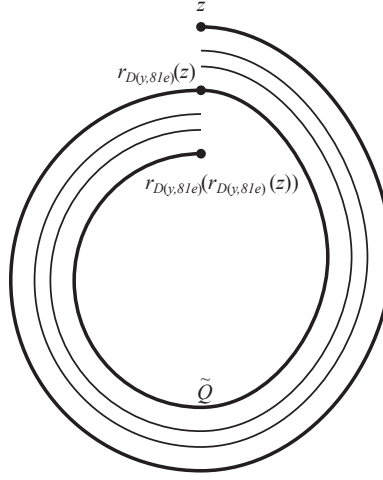


FIGURE 47. Other parts of E_+ may lie between z and $r_{D(z,81e)}(z)$ but these remain parallel to Q

for any multiple n of k , $\bar{S} + nT$ is obtained from \bar{S} by Dehn twisting some number of times around T .

Proof. Consider a simple closed curve on T that intersects each curve of $\bar{S} \cap T$ once. Then in a regular neighbourhood of this curve, the surfaces \bar{S} and T sit as shown in Figure 48. It is now clear that, when n is a multiple of k , then $\bar{S} + nT$ is obtained from \bar{S} by Dehn twisting about T n/k times. \square

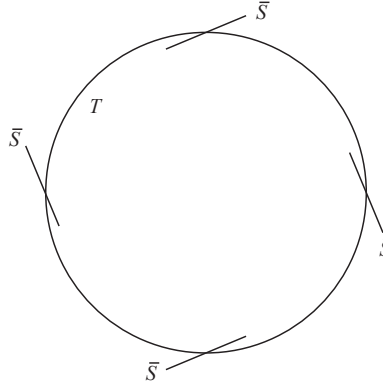


FIGURE 48. The case $k = 4$

Thus, to prove Proposition 11.24, we just need to show that $\bar{E}_+ \cap T'$ consists of essential curves in T' and that, for each component T of T' , $|\bar{E}_+ \cap T|$ is a multiple of $m+1$. The following lemma will be used to deal with the number of these curves.

Lemma 11.26. *Let ζ' be a collection of parallel essential simple closed curves on a torus T . Let ζ be a simple closed curve of T with slope other than that of ζ' and that intersects each component of ζ' in points of the same sign. Then $|\zeta'|$ divides $|\zeta \cap \zeta'|$.*

Proof. The assumption that intersects each component of ζ' in points of the same sign implies that the number of such points is just the modulus of the intersection between the slopes. Hence, $|\zeta \cap \zeta'|$ is just $|\zeta'|$ times this intersection number. \square

Proof of Proposition 11.24. Let σ be the curve on the torus T defined in Section 11.17. We see that all the intersection points between σ and $\overline{E}_+ \cap T$ have the same sign and respect the direction of the normal summation. Hence, $\zeta = \sigma$ and $\zeta' = \overline{E}_+ \cap T$ satisfy the hypotheses of Lemma 11.26. So for each component T of T' , $|E_+ \cap T|$ divides $|\sigma \cap E_+|$, which is $m + 1$. Hence, by Lemma 11.25, \overline{E}_+ is obtained from E_+ by Dehn twisting some number of times about T' . \square

11.20. The case where T is disjoint from its image. Suppose now that T is disjoint from its image under the involution of $N(B_{E_+})$. Hence, it restricts to a torus, which we will also call T , in $N(B_E)$. This is disjoint from the boundary of M . The Dehn twist defined in Proposition 11.24 restricts to a Dehn twist about T in M . Let \overline{E} be the result of removing $(m + 1)$ summands of T from E , viewed as a surface carried by B_E . Then \overline{E} is obtained from E by applying some power h of this Dehn twist about T . Let S'_2 be the surface $(S' \setminus N(B_E)) \cup \overline{E}$. Then S'_2 is obtained from S' by applying h . The surface $S' \setminus S'$ may intersect $N(B)$, but near T , it runs parallel to S' . Hence, we may also obtain a surface S_2 from S by applying h . Similarly, we can obtain the 3-manifold $h(N')$ and handle structure $h(\mathcal{H}')$. So, $(N, X, h(Y), P \cap N, \mathcal{H}, h(S))$ and $(h(N'), X, h(Y'), P \cap N', h(\mathcal{H}'), h(S'))$ are nested admissible envelopes. (See Figure 49.) Note that the binding weight $w_\beta(h(S'))$ is strictly less than the binding weight $w_\beta(S')$, since we have removed $(m + 1)w_\beta(T)$, which is positive. This contradicts one of the assumptions of Theorem 11.3. So, the theorem is proved in this case.

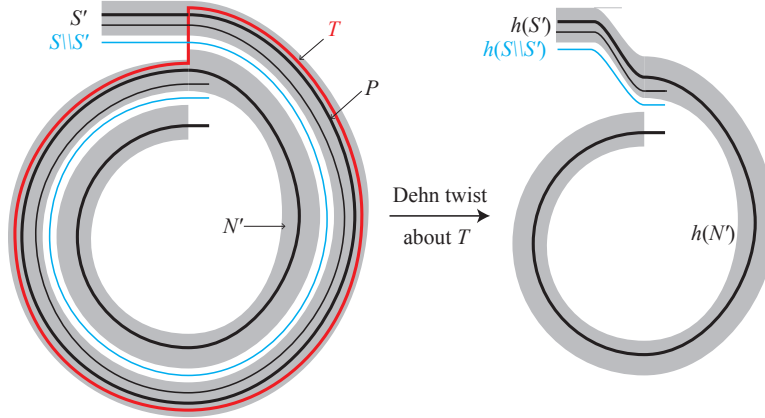


FIGURE 49. Dehn twisting along the torus T in M

11.21. The case where T is invariant under the involution. In this case, T restricts to an annulus A in $N(B_E)$. The Dehn twist about T restricts to a Dehn twist about A . Since we doubled $N(B_E)$ along $\partial M' \cap N(B)$, this annulus A has boundary lying in $\partial M'$.

When T was disjoint from its image under the involution, there may have been parts of $S' \setminus S'$ lying between the sheets of E near T . However, in the current case,

this does not arise, for the following reason. If there was any part of $S \setminus S'$ between the sheets of E near T , then they would run parallel to E , and so would have to intersect $\partial M'$. Consider such a point of intersection. Above and below this point would be tiles of S' that are of the same approximate type. However, these tiles intersect $\partial M'$, and so must actually be parallel in N' , and so there cannot in fact be $S \setminus S'$ between them.

If any component of ∂A does not lie in ∂M , then we may extend A to an annulus \tilde{A} properly embedded in M by attaching a collar in the wedge region $N \setminus N'$. This annulus \tilde{A} lies in N and is carried by the branched surface of which N is thickening. The Dehn twist about A extends to a Dehn twist h about \tilde{A} . Again, $(N, X, h(Y), P \cap N, \mathcal{H}, h(S))$ and $(h(N'), X, h(Y'), P \cap N', h(\mathcal{H}'), h(S'))$ are nested admissible envelopes. Again, the binding weight $w_\beta(h(S'))$ is strictly less than the binding weight $w_\beta(S')$, since we have removed $(m+1)w_\beta(A)$, which is positive. Again this proves Theorem 11.3 in this case.

12. THE THIN PART OF THE SURFACE

In Section 10.12, the cut vertices of S' were defined. These are the non-manifold points of S' . Modifications to S' were given, but it was important that these took place away from the cut vertices. As a result, the surface S'_{cut} was defined in Section 10.12, which is obtained from S' by cutting along its cut vertices. The remnants of the cut vertices in S'_{cut} are not forgotten. We can view the modifications in Section 10 as being applied to S'_{cut} . We will see later that we can apply these modifications until the binding weight of the non-bigon components of S'_{cut} is small, in the sense that it is bounded above by a polynomial function of the initial arc index of the link. However, this does not imply that the binding weight of S' is small, since there may be parts of S' that consist of long lines of bigons joined up in a row. We call these the ‘thin’ parts of S' . In this section, we explain how to modify these thin parts so that they too have small binding weight.

12.1. The thin part. The *thin part* of the surface S' consist of the union of the bigon components of S'_{cut} , together with the cut vertices of S' that are incident to two bigons and no other tiles. See Figure 50. We denote the thin part by S'_{thin} .

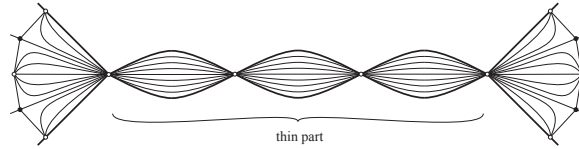


FIGURE 50. The thin part of S'

Recall that S' is properly embedded in the manifold N' . Let M' be obtained from N' by attaching $M \setminus N$. Thus, M' is obtained from M by inserting wedges, but we suppose that when a wedge is inserted in a way that creates a cut vertex, then we do not quite insert the full wedge, in order to maintain M' as a 3-manifold.

The following is the main result of this section.

Proposition 12.1. *Suppose that the component of M' containing S' is not an I -bundle over a compact surface with P lying in the vertical boundary as a collection of essential simple closed curves. Suppose also that if the component of M' containing*

S' has any JSJ annuli, then S' is the JSJ annuli. Let e_{bigon} be an upper bound for the number of \mathcal{H}' -parallelism types of bigons in S' . There is a pattern-isotopy of S' taking it to a nearly admissible surface S''' with binding weight at most

$$(2e_{\text{bigon}} + w_\beta(\partial M') + 1) w_\beta(S' \setminus S'_{\text{thin}}).$$

Remark 12.2. This proposition refers to the number of \mathcal{H}' -parallelism types of bigons in S' . This is at most the number of approximate star types of tiles of S' . This is because when two bigons have the same approximate star type, then they are actually parallel in \mathcal{H}' , as their vertices lie on $\partial M'$.

12.2. The thin bundle. The *thin bundle* \mathcal{B} is a 3-dimensional submanifold of M' , constructed as follows. Take a regular neighbourhood of the thin part of S' . Between any two bigon tiles of S'_{thin} that are \mathcal{H}' -parallel, insert the space between them. Between any two vertices of the thin part of S' that are \mathcal{H}' -parallel, insert the space between them. This is an I -bundle over a surface, for the following reason. Each bigon tile of S'_{thin} is a product $I \times I$, where the second factor is given by the θ value. A regular neighbourhood of each vertex in the thin part of S' is again a product $I \times I$, when this vertex has been suitably enlarged so that the surface is properly embedded in M' . When we form regular neighbourhoods of these bigon tiles and these vertices, the result is still an I -bundle, and it remains so when we attach the spaces between them.

12.3. Annular simplifications. We now describe a modification to M' . Suppose that there is a clean annulus A properly embedded in M' . Suppose also that there is a clean annulus A' in $\partial M'$ such that $\partial A = \partial A'$. Assume that $A \cup A'$ bounds a 3-manifold G such that

- (1) either G is product region between A and A' , or G lies in a 3-ball in M' ;
- (2) the intersection between G and the thin bundle \mathcal{B} is a union of components of \mathcal{B} ;
- (3) A is a vertical boundary component of \mathcal{B} .

Removing the interiors of G and A' from M' is called an *annular simplification*.

Note that an annular simplification can be achieved by an isotopy of an earlier surface in the partial hierarchy, as follows. Since A' is disjoint from the boundary pattern, it lies within some surface S_i of the partial hierarchy. If we replace S_i by the surface $(S_i \setminus A') \cup A$, then this new surface is isotopic to S_i . All the vertices of A are vertices of A' , and so this does not increase the binding weight of the hierarchy. This isotopy of the surface has the effect of removing G from M' . However, G is added to the part of the 3-manifold on the other side of S_i . Note that when constructing the partial hierarchy, we always ensured that it was adequately separating, and we therefore guaranteed that on the two sides of A' , there were different components of M' . Thus, whenever we perform an annular simplification to the component of M' containing S' , we reduce the number of components of \mathcal{B} in that component of M' .

So let us suppose that no further annular simplifications can occur to the component of M' containing S' .

12.4. The generalised thin bundle. By using arguments analogous to ones introduced in [16], we will now enlarge \mathcal{B} to a bigger I -bundle \mathcal{B}_+ , which we call the *generalised thin bundle*. Throughout this enlargement process, \mathcal{B}_+ will have the following properties:

- (1) \mathcal{B}_+ is an I -bundle over a compact surface;
- (2) the horizontal boundary of \mathcal{B}_+ is $\mathcal{B}_+ \cap \partial M'$, and is clean;
- (3) the vertical boundary of \mathcal{B}_+ is a union of vertical boundary components of the thin bundle.

Proposition 12.3. *Suppose that \mathcal{B}_+ is maximal with the above properties. Suppose also that the component of M' containing S' admits no annular simplifications. Then each component of \mathcal{B}_+ either is an I -bundle over a disc or has incompressible vertical boundary.*

Proof. Let \mathcal{B}'_+ be the union of the components of \mathcal{B}_+ that are not I -bundles over discs. We will show that the vertical boundary of \mathcal{B}'_+ is incompressible.

Suppose that, on the contrary, it has a compression disc D . Let A be the vertical boundary component containing ∂D . By the definition of \mathcal{B}'_+ , D does not lie entirely in \mathcal{B}'_+ . Its interior is disjoint from \mathcal{B}'_+ (by the definition of a compression disc), but it may intersect $\mathcal{B}_+ - \mathcal{B}'_+$.

Now, A compresses along D to give two discs D'_1 and D'_2 properly embedded in M' . Their boundaries are clean. Since (M', P) is irreducible and boundary-irreducible, D'_1 and D'_2 are parallel to discs D_1 and D_2 in $\partial M' - P$, via 3-balls B_1 and B_2 . There are two cases to consider: where B_1 and B_2 are disjoint and where they are nested.

Let us suppose first that they are disjoint. Then, $A \cup D_1 \cup D_2$ bounds a 3-ball B , by the irreducibility of M' . Since the interior of D is disjoint from \mathcal{B}'_+ , this ball B does not lie in \mathcal{B}'_+ . So, we may extend the I -bundle structure of \mathcal{B}'_+ over B , contradicting the maximality of \mathcal{B}_+ .

Let us now suppose that B_1 and B_2 are nested; say that B_2 lies in B_1 . Let A' be $D_1 \setminus D_2$. Then, A' is a clean annulus in $\partial M'$ such that $\partial A = \partial A'$. Let G be the 3-manifold bounded by $A \cup A'$. This lies in the 3-ball B_1 . So, M' admits an annular simplification, which is a contradiction.

Thus, the vertical boundary of \mathcal{B}'_+ is incompressible. \square

Proof of Proposition 12.1. Let $\partial_v \mathcal{B}_+$ be the vertical boundary of the generalised thin bundle. It is a union of components of the vertical boundary of \mathcal{B} .

We claim that the binding weight of $\partial_v \mathcal{B}_+$ is at most $2e_{\text{bigon}}$, where e_{bigon} is an upper bound for the number of \mathcal{H}' -parallelism types of bigons in S' . It is composed of a union of bigon tiles of S' , and each such tile is outermost in its \mathcal{H}' -parallelism class. The claim follows immediately.

We will show that, after a pattern-isotopy, S' intersects \mathcal{B}_+ in discs that are vertical with respect to the I -bundle structure on \mathcal{B}_+ . Furthermore, this does increase the binding weight of the surface outside of \mathcal{B}_+ and it does not increase the number of components of $S' \cap \partial_v \mathcal{B}_+$. First note that, because the vertical boundary of \mathcal{B}_+ is equal to a union of vertical boundary components of \mathcal{B} , and S' intersects \mathcal{B} in bigons, we may perform a small isotopy of S' so that afterwards it intersects $\partial_v \mathcal{B}_+$ in vertical arcs. Suppose now that $S' \cap \mathcal{B}_+$ has a boundary compression disc D that intersects $\partial \mathcal{B}_+$ in $\partial_h \mathcal{B}_+$. Since S' is boundary-incompressible (with respect to the pattern), the arc $S' \cap D$ separates off a clean disc in S' . This and D separate off a ball disjoint from the pattern, and we may isotope S' across this ball. This has the effect of reducing the number of components of $S' \cap \partial_v \mathcal{B}_+$. It also does not increase the binding weight of S' outside of \mathcal{B}_+ . Note also that $S' \cap \mathcal{B}_+$ is incompressible. For suppose that it had a compression disc D . Since S' is incompressible, ∂D bounds

a disc D' in S' . This misses the boundary of M' , and hence does not intersect $\partial_v \mathcal{B}_+$. But it must then lie in \mathcal{B}_+ , which is a contradiction, proving the claim. Thus $S' \cap \mathcal{B}_+$ is now incompressible and has no boundary compression disc that intersects $\partial \mathcal{B}_+$ in $\partial_h \mathcal{B}_+$. The only such surfaces are, up to isotopy preserving $\partial_h \mathcal{B}_+$ and $\partial_v \mathcal{B}_+$, vertical or horizontal in \mathcal{B}_+ . However, since $S' \cap \partial_v \mathcal{B}_+$ is vertical, we deduce that $S' \cap \mathcal{B}_+$ must be vertical, as required.

Let \mathcal{B}_+^a be those components of \mathcal{B}_+ that are I -bundles over annuli, and where both vertical boundary components are inessential in M' . We now isotope S' off $\mathcal{B}_+ - \mathcal{B}_+^a$. We do this first for the components of $\mathcal{B}_+ - \mathcal{B}_+^a$ that are I -bundles over discs. In any such $D^2 \times I$ component, its intersection with S' is isotopic to $\alpha \times I$, where α is a collection of arcs properly embedded in D^2 . Pick an outermost such arc, and replace it by a disc parallel to a subset of $\partial_v \mathcal{B}_+$. Repeating this process, we eventually isotope S' off the $D^2 \times I$ components of \mathcal{B}_+ .

Now consider the components of $\mathcal{B}_+ - \mathcal{B}_+^a$ that are not I -bundles over discs. By Proposition 12.3, their vertical boundary is a union of incompressible annuli. A component B of $\mathcal{B}_+ - \mathcal{B}_+^a$ cannot have all its vertical boundary components being inessential in M' , for the following reason. If any is inessential, then it is parallel to an annulus in $\partial M'$ that intersects the pattern in a (possibly empty) union of core curves. If all the product regions between these annuli and the vertical boundary are disjoint from the interior of B , then this component of M' is an I -bundle, and the boundary pattern intersects each vertical boundary component in a (possibly empty) collection of parallel core curves, which is contrary to assumption. On the other hand, if one of these product regions contains B , then $\partial_v B$ is an essential subsurface of an annulus, and hence is a collection of annuli. This implies that B must be an I -bundle over an annulus, which is a contradiction. Hence, $\mathcal{B}_+ - \mathcal{B}_+^a$ is non-empty, then we deduce that M' contains a clean essential annulus. In this situation, we are assuming that S' consists of the JSJ annuli. Therefore, S' can be pattern-isotoped off the incompressible vertical boundary of \mathcal{B}_+ . So, there is a product region between S' and this vertical boundary. We can pattern-isotope S' across this product region, so that part of it runs parallel to a subset of the vertical boundary. It is possible that this product region does not lie in \mathcal{B}_+ , in which case, the new part of S' lies in \mathcal{B}_+ . However, it remains the case that S' intersects \mathcal{B}_+ in vertical discs, or possibly annuli parallel to some vertical boundary components of \mathcal{B}_+ . It also remains the case that S' is disjoint from the components of \mathcal{B}_+ that are I -bundles over discs. Hence, we may repeat this process until S' is disjoint from the vertical boundary of \mathcal{B}_+ . If any components of S' are parallel to a vertical boundary component of \mathcal{B}_+ , they may be pattern-isotoped off \mathcal{B}_+ .

At the end of this process, the resulting surface S''' is disjoint from $\mathcal{B}_+ - \mathcal{B}_+^a$. In each component of B^a of \mathcal{B}_+^a , the surface is vertical. If any component S''' is inessential in B_a , we may isotope it out of B_a , so that it runs parallel to a subset of the vertical boundary. The remaining essential components of $S''' \cap B_a$ may be arranged so that they intersect each vertex in $\partial_h B_a$ at most once. The resulting surface differs by an isotopy and some power of a Dehn twist about one of the boundary components of $\partial_v B_a$. This Dehn twist is pattern-isotopic to the identity on M' .

Each component of $S''' \setminus S'$ in $\mathcal{B}_+ - \mathcal{B}_+^a$ runs parallel to a vertical subset of $\partial_v \mathcal{B}_+$. It therefore has binding weight at most $2e_{\text{bigon}}$. Each such component starts and ends on a vertex of $S' \setminus S'_{\text{thin}}$ that is not in its thin part. Hence, the

total binding weight of $S''' \setminus S'$ in $\mathcal{B}_+ - \mathcal{B}_+^a$ is at most $2e_{\text{bigon}} w_\beta(S' \setminus S'_{\text{thin}})$. In \mathcal{B}_+^a , each component of $S''' \setminus S'$ also starts and ends on a vertex of $S' \setminus S'_{\text{thin}}$ that is not in its thin part. So the binding weight of those components is at most $w_\beta(\partial M') w_\beta(S' \setminus S'_{\text{thin}})$. The remainder of S''' is $S' \setminus S'_{\text{thin}}$, and so we obtained our required upper bound on the binding weight of S''' . \square

13. THE EULER CHARACTERISTIC ARGUMENT

Theorem 13.1. *Let (M, P) be a compact orientable irreducible boundary-irreducible 3-manifold with boundary pattern. Suppose that M is a subset of S^3 with ∂M in alternative admissible form and P in an arc presentation. Let S be a properly embedded surface in M in generalised admissible form. Let $(N, X, Y, P \cap N, \mathcal{H}, S)$ and $(N', X, Y', P \cap N', \mathcal{H}', S')$ be nested admissible envelopes. Suppose that there is no homeomorphism h as in Theorem 11.3. Let e be the number of approximate star types of tiles of S' . Let $S' \setminus S'_{\text{thin}}$ be obtained from S' by cutting along its cut vertices and then discarding any bigon tiles. Let S'' be obtained from S'_{thin} by discarding any toral or annular components that lie in E and have binding weight at most $23552e^2$. Let S_+ be the double of S'' , which inherits a singular foliation. Then the number of non-cut vertices of S_+ with valence at most 3 is at least*

$$\frac{w_\beta(S'')}{31\pi(2392e)^2} + \frac{72}{31}\chi(S) - \frac{5}{31}|S \cap P|.$$

Proof. Recall that the cut-adjusted Euler characteristic χ_c was defined in Section 10.12 and that $\chi_c(S'') = \chi_c(S' \setminus S'_{\text{thin}}) = \chi_c(S'_{\text{cut}}) = \chi(S')$. Also define the *cut-adjusted Euler characteristic* $\chi_c(S_+)$ of S_+ to be $\chi(S_+)$ minus its number of cut vertices. Then $\chi_c(S_+) = 2\chi_c(S'') = 2\chi(S')$. Note that S' is homeomorphic to S plus possibly some clean discs, and so $\chi(S') \geq \chi(S)$. Form a cell complex on S_+ by declaring that each vertex of S_+ is a 0-cell and each generalised saddle is a 0-cell. Each separatrix is a 1-cell. Each tile is a 2-cell. Let E and F be the edges and faces of this cell complex. Let C be the cut vertices of S_+ . Since the vertices V and generalised saddles X form the 0-cells of this complex, then obviously

$$|V| + |X| - |E| + |F| - |C| = \chi_c(S_+) \geq 2\chi(S).$$

Since each tile of S_+ has four separatrices in its boundary, $2|E| = 4|F|$, and so $|E| = 2|F|$. Each edge runs from a vertex to a generalised saddle, and so contributes 1 to $d(v)/2 + d(x)/2$ for some vertex v and generalised saddle x , where d is their valence. So

$$|E| = \frac{1}{2} \sum_{v \in V} d(v) + \frac{1}{2} \sum_{x \in X} d(x).$$

Therefore,

$$\begin{aligned} 2\chi(S) &\leq |V| + |X| - |E| + |F| - |C| \\ &= |V| + |X| - |E|/2 - |C| \\ &= \sum_{v \in V} \left(1 - \frac{d(v)}{4}\right) + \sum_{x \in X} \left(1 - \frac{d(x)}{4}\right) - |C|. \end{aligned}$$

We now introduce the *cut-adjusted valence* of a vertex v in S_+ to be

$$d_c(v) = \begin{cases} d(v) + 4 & \text{if } v \text{ is a cut vertex} \\ d(v) & \text{otherwise.} \end{cases}$$

So,

$$2\chi(S) \leq \sum_{v \in V} \left(1 - \frac{d_c(v)}{4}\right) + \sum_{x \in X} \left(1 - \frac{d(x)}{4}\right).$$

For each integer $k \geq 2$, let v_k be the number of vertices in S_+ with cut-adjusted valence k . Then

$$3v_1 + 2v_2 + v_3 = \sum_{\substack{v \in V \\ d_c(v) < 4}} (4 - d_c(v)) \geq 8\chi(S) + \sum_{\substack{v \in V \\ d_c(v) > 4}} (d_c(v) - 4) + \sum_{x \in X} (d(x) - 4).$$

Note that

$$\sum_{\substack{v \in V \\ d_c(v) > 4}} (d_c(v) - 4) = \sum_{k > 4} \sum_{\substack{v \in V \\ d_c(v) = k}} (d_c(v) - 4) = \sum_{k > 4} v_k(k - 4) \geq \frac{1}{5} \sum_{k > 4} kv_k.$$

Similarly, because each generalised saddle has valence that is an even integer at least 4,

$$\sum_{x \in X} (d(x) - 4) \geq \frac{1}{3} \sum_{\substack{x \in X \\ d(x) \neq 4}} d(x).$$

So,

$$3v_1 + 2v_2 + v_3 \geq \left(\frac{1}{5} \sum_{k > 4} kv_k + \frac{1}{3} \sum_{\substack{x \in X \\ d(x) \neq 4}} d(x) \right) + 8\chi(S).$$

In particular,

$$(1) \quad 3v_1 + 2v_2 + v_3 \geq \frac{1}{5} \sum_{k > 4} kv_k + 8\chi(S) \geq \sum_{k > 4} v_k + 8\chi(S),$$

and

$$3v_1 + 2v_2 + v_3 \geq \frac{1}{3} \sum_{\substack{x \in X \\ d(x) \neq 4}} d(x) + 8\chi(S).$$

We now wish to bound $\deg(S_+ \setminus E_+)$ from above in terms of the valence of the non-Euclidean saddles and vertices. Let $v_4^{E_+}$ be the number of 4-valent non-cut vertices in the interior of E_+ . Each of these vertices in the interior of E_+ contributes 1 to the area of E_+ and so by Lemma 11.1 and Theorem 11.3,

$$v_4^{E_+} \leq \pi(2392e)^2 \deg(S_+ \setminus E_+).$$

The remaining 4-valent non-cut vertices lie in the copy of $S \cap P$ in S_+ , and therefore

$$v_4 \leq v_4^{E_+} + |S \cap P|.$$

Each point in $S_+ \setminus E_+$ is a generalised saddle that is not a saddle, a vertex with cut-adjusted valence other than 4 or a point in $S \cap P$ with cut-adjusted valence 4.

So, we deduce that

$$\begin{aligned}
& \deg(S_+ \setminus E_+) \\
& \leq 4|S \cap P| + \sum_{k \neq 4} kv_k + \sum_{\substack{x \in X \\ d(x) \neq 4}} d(x) \\
& \leq 4|S \cap P| + 3v_1 + 2v_2 + v_3 + \sum_{k > 4} kv_k + \sum_{\substack{x \in X \\ d(x) \neq 4}} d(x) \\
& \leq 4|S \cap P| + 3v_1 + 2v_2 + v_3 + 15v_1 + 10v_2 + 5v_3 - 40\chi(S) + 9v_1 + 6v_2 + 3v_3 - 24\chi(S).
\end{aligned}$$

Therefore,

$$\begin{aligned}
(2) \quad 27v_1 + 18v_2 + 9v_3 & \geq \deg(S_+ \setminus E_+) + 64\chi(S) - 4|S \cap P| \\
& \geq \frac{v_4^{E_+}}{\pi(2392e)^2} + 64\chi(S) - 4|S \cap P| \\
& \geq \frac{v_4}{\pi(2392e)^2} + 64\chi(S) - 5|S \cap P|.
\end{aligned}$$

Adding (1) to (2), we deduce that

$$31v_1 + 21v_2 + 11v_3 \geq \frac{v_4}{\pi(2392e)^2} + \sum_{k \neq 4} v_k + 72\chi(S) - 5|S \cap P|.$$

Therefore,

$$31v_1 + 21v_2 + 11v_3 \geq \frac{|V|}{\pi(2392e)^2} + 72\chi(S) - 5|S \cap P|.$$

Finally note that each vertex in S'' gives rise to one or two vertices in V , and each vertex in V arises in this way. Therefore,

$$w_\beta(S'') \leq |V| \leq 2w_\beta(S'').$$

Hence, the number of non-cut vertices of S_+ with valence at most 3 is at least

$$\frac{w_\beta(S'')}{31\pi(2392e)^2} + \frac{72}{31}\chi(S) - \frac{5}{31}|S \cap P|.$$

□

14. MODIFICATIONS IN THE CASE OF A BROKEN SECTION

In the hierarchy described in Section 3, the final surface might be a broken section. Recall that this is where the exterior of the partial hierarchy is a collection of 3-balls and solid tori with clean longitudes. The solid tori patch together to form a circle bundle. The intersection between any two such solid tori is a collection of annuli that are unions of circle fibres. The intersection of any three such solid tori is a union of finitely many fibres. We let M denote the circle bundle, we let A be these annuli and we let C be the union of the fibres meeting three solid tori and the fibres meeting ∂M and two solid tori. The broken section is obtained from a section of M by perturbing it a little in each component of $M \setminus A$. Thus in each component of $M \setminus A$, it is a meridian disc and the discs are all disjoint from each other. Let P be the boundary pattern for M .

In this section, we present the minor modifications to our previous arguments that are needed to deal with broken sections.

Some components of M are dealt with in a different way from others. We let M_{st} be the union of the solid toral components of M , and let M_{nst} be the union of the remaining components.

14.1. The handle structure \mathcal{K} . We are considering the situation where the penultimate manifold of the hierarchy is $M \setminus A$ plus some 3-balls. As previously, we create the polyhedral decomposition Δ for the exterior of K that contains the partial hierarchy H has a subcomplex. This also collapses to a triangulation \mathcal{T} . This restricts to a triangulation \mathcal{T}_M for M . The annuli A are simplicial in \mathcal{T} and \mathcal{T}_M . As in Section 6.5, we enlarge \mathcal{T}_M to $\hat{\Delta}_M$, and we let \mathcal{K} be the dual handle structure for M . Since A was simplicial in \mathcal{T}_M , it naturally intersects \mathcal{K} in the 1-handles, 2-handles and 3-handles. In particular, for each 1-handle, A is either disjoint from it or A intersects the 1-handle in a co-core.

14.2. The components that are not solid tori. For the moment, we focus on M_{nst} .

Lemma 14.1. *Each component of $\partial M_{\text{nst}} \setminus P$ is a disc.*

Proof. Each component of $M \setminus A$ is a solid torus with a clean longitudinal annulus in its boundary. Hence, $\partial M_{\text{nst}} \setminus P$ consists of discs and annuli, where the annuli are a union of fibres. So if any component of $\partial M_{\text{nst}} \setminus P$ is an annulus, it lies in $\partial N(K)$, since otherwise it would be incident to another solid toral component and hence would lie in A rather than in ∂M_{nst} . So there is an incompressible clean annulus A' properly embedded in M_{nst} that is not parallel to a clean annulus in ∂M_{nst} . A core curve of A' is a fibre of the circle bundle, and hence is π_1 -injective in the exterior of K . So, A' is incompressible in the exterior of K . It is disjoint from the JSJ tori of the exterior of K , since we cut along those in an earlier step in the partial hierarchy, and it does not lie in the Seifert pieces, since these have been decomposed into solid tori by the partial hierarchy. Hence, A' is boundary parallel in the exterior of K . Consider the product region between A' and the relevant annulus in $\partial N(K)$. The first surface in the partial hierarchy to intersect this region would have to be inessential, contrary to the construction of the hierarchy. This contradiction proves the lemma. \square

14.3. The branched surface B and its natural handle structure. According to Proposition 2.23, some section S of M_{nst} can be pattern-isotoped, via an isotopy preserving A , to a weakly fundamental surface in \mathcal{T}_M . Moreover, this surface intersects P as few times as possible, and among such surfaces, it has least extended weight. As in Section 8.5, the branched surface B is constructed, which has a patch for each tile type of S . The annuli A intersect $N(B)$ in a union of fibres. Let \mathcal{H} denote the natural handle structure on $N(B)$.

14.4. The arc index of $\partial S \cap M_{\text{nst}}$.

Lemma 14.2. *The arc index of $\partial S \cap M_{\text{nst}}$ is at most $|S \cap P|96w(w - \sum_{i=1}^n \chi(S_i))$, where w is binding weight of the partial hierarchy before S .*

Proof. The curves $\partial S \cap M_{\text{nst}}$ are cut into arcs by P . Each such arc lies in a disc component of $\partial M_{\text{nst}} \setminus P$. Since S has minimal extended weight up to pattern-isotopy, each arc of $\partial S \cap M_{\text{nst}}$ must intersect therefore each edge of \mathcal{T}_M at most once. The number of edges of \mathcal{T}_M is at most $96w(w - \sum_{i=1}^n \chi(S_i))$ by Lemma 6.3. \square

14.5. New boundary pattern. Define P' to be the boundaries of the tiles of the partial hierarchy in ∂M_{nst} . Then $|S \cap P'|$ is at most the arc index of $\partial S \cap M_{\text{nst}}$, which is bounded above by Lemma 14.2.

14.6. Modifications arising from low-valence interior vertices. In Section 9, various modifications were made to the next surface in the partial hierarchy. In this case, we make these modifications to the section S . Note that we never insert a wedge, since this is only performed around low-valence vertices in ∂S that are disjoint from the pattern. However, every vertex in ∂S is a vertex of ∂M , and is therefore in the pattern P' .

The modifications that we make will move the section S , but they will not move A . Hence, the relative position of S and A may change. In particular, $S \setminus A$ need not remain meridian discs for the solid tori $M \setminus A$. We will adjust A suitably in Section 14.10.

As in Section 10, we keep track of an admissible envelope $(N, X, Y, P \cap N, \mathcal{H}, S)$. In order to maintain consistency with the early argument, we will also consider another surface S' and another admissible envelope $(N', X, Y', P \cap N', \mathcal{H}', S')$. However, since we will not be performing any wedge insertions, S', N', Y' and \mathcal{H}' will be equal to S, N, Y and \mathcal{H} .

The annuli A initially intersect $N(B)$ in a union of fibres, and we will maintain this property for the intersection with N .

When we consider interior 2-valent vertices as in Section 10.15, we need to be careful also to consider the annuli A . The only situation where care is required here is when A intersects the star of each vertex s^i in an arc that runs from s_1^i to s_2^i via s^i . But after the modification in Section 10.15, there remains a copy of s_1^i and s_2^i , and we simply ensure that A includes the fibre between them.

The modification involving 3-valent interior vertices is also easy to handle. There a tile that is incident to the vertex is collapsed, thereby combining two generalised saddles. If A intersects that tile, it does so in a fibre, and we can ensure that after the modification, A runs through the new combined generalised saddle.

14.7. The Euclidean subsurface. As before, we must consider the Euclidean subsurface of $S' = S$. This is defined in the same way as previously, except that E also does not include the intersection $S \cap C$. The branched surface B_E is disjoint from $\partial M \cup C$, as is $E \setminus N_2(S' \setminus E)$. It therefore lies in the manifold $M_C = M \setminus N(C)$. This is a circle bundle over the circle, and we can view the annuli A as properly embedded in M_C .

We need to ensure that the conclusion of Theorem 11.3 still holds. That theorem had two hypotheses, which we vary to the following. The first hypothesis is that there is no homeomorphism $h: N \rightarrow N$ such that all the following hold:

- (1) h is equal to the identity on $P \cap N$ and X (and so extends to M_C and M);
- (2) h is a Dehn twist along a torus properly embedded in M ;
- (3) $(N, X, Y, P \cap N, \mathcal{H}, h(S))$ is an admissible envelope;
- (4) $w_\beta(h(S)) < w_\beta(S)$.

The second hypothesis is there is no pattern-isotopy taking S to a surface S_2 , and a nested admissible envelope $(N_2, X_2, Y_2, P \cap N_2, \mathcal{H}_2, S_2)$ satisfying

$$(w_\beta(S_2), \text{approx-type}(S_2)) < (w_\beta(S), \text{approx-type}(S)),$$

$$c_+(\mathcal{H}_2) \leq c_+(\mathcal{H}), \quad c(\mathcal{H}_2) \leq c(\mathcal{H}).$$

We now explain why the conclusion of Theorem 11.3 remains true with these modified hypotheses.

The second hypothesis is used in the proof of Proposition 11.10. There, a collection of disjoint grids in \tilde{E}_+ are considered. In the proof, a disc D_- in one of these grids is analysed. Its boundary is a boundary component of an annulus that is vertical in N . The other boundary component of this annulus lies in S and bounds a disc D_+ in S . If the discs have different binding weights, then we may replace one by the other to reduce the binding weight of S . On the other hand, if they have the same binding weight, then we can replace as many parallel copies of D_+ as possible by parallel copies of D_- . This reduces the number of approximate star types of tiles of S .

The proof proceeds by constructing a torus T carried by $N(B_+)$. In our situation, T is actually carried by N . Then a power of a Dehn twist h about T reduces the binding weight of S . This contradicts our first hypothesis.

14.8. The thin part. In our previous argument, the insertion of wedges may introduce cut vertices to S' . As a result, the surface S'_{cut} was defined. However, in our situation, no wedges are inserted, and therefore no cut vertices are introduced. Hence, $S'_{\text{cut}} = S' = S$, and the thin part of S' is empty.

14.9. The Euler characteristic argument. This is almost unchanged. The only variation is that we have modified the boundary pattern from P to P' , and we have also declared that the Euclidean subsurface is disjoint from C . Hence, the revised statement of Theorem 13.1 is that the number of interior vertices of S with valence at most 3 is at least

$$\frac{w_\beta(S)}{31\pi(2392e)^2} + \frac{72}{31}\chi(S) - \frac{5}{31}|S \cap P'| - \frac{5}{31}|S \cap C|.$$

14.10. Controlling the annuli A . As we will see in Section 15, we will modify S until its weight is bounded. These modifications all take place away from C and so are contained in the manifold $M_C = M \setminus N(C)$. Let $F = S \cap M_C$ and let F' be the resulting surface in M_C . We have not moved A , and so $F \cup A$ might not be isotopic to $F' \cup A$. We now explain how to remedy this. We will consider the following modifications.

Let S_1 and S_2 be two surfaces properly embedded in a 3-manifold M_C and intersecting transversely. Suppose there is a simple closed curve of $S_1 \cap S_2$ that bounds a disc in one of the surfaces. Suppose that the disc is $D_1 \subset S_1$. Suppose also that the interior of D_1 is disjoint from S_2 . Assume also that ∂D_1 bounds a disc D_2 in S_2 , which might have interior intersecting S_1 . Then removing D_2 from S_2 and replacing it by a parallel copy of D_1 , thereby removing ∂D_1 from $S_1 \cap S_2$, is known as *disc swap*.

Lemma 14.3. *Let M_C be a compact orientable 3-manifold that is a circle bundle over the circle. Let A be a collection of properly embedded vertical annuli so that $M_C \setminus A$ is a collection of solid tori. Let F be a section of M_C . Let F' be obtained from F by a homeomorphism of M_C supported in the interior of M_C , and suppose also that F' intersects A transversely. Suppose that F'' and A'' are obtained from F' and A by a sequence of disc swaps until no more such disc swaps are possible. Then there is a homeomorphism of M_C , supported in the interior of M_C and taking $F'' \cup A''$ to $F \cup A$.*

Proof. Since F is a section, it intersects each component of ∂A exactly once. As F'' is obtained from F by a homeomorphism of M_C supported in the interior of M_C , F'' also intersects each component of ∂A exactly once. Hence, F'' intersects each component of A in an essential arc plus possibly some inessential simple closed curves. However, because no further disc swaps are possible, there are no simple closed curves of $F'' \cap A$. By a further isotopy preserving A , we may assume that the arcs $F'' \cap A$ are transverse to the circle fibres.

Let $N(A \cup \partial M_C)$ be a thin regular neighbourhood of $A \cup \partial M_C$. We may assume that $F'' \cap N(A \cup \partial M_C)$ is homeomorphic to $F \cap N(A \cup \partial M_C)$. Now, F is obtained from $F'' \cap N(A \cup \partial M_C)$ by attaching discs. Since F and F'' have the same Euler characteristic, F'' must also be obtained from $F'' \cap N(A \cup \partial M_C)$ by attaching discs. These must be one such disc in each component of $M_C \setminus A$, forming a meridian disc for that component. Thus, there is an isotopy preserving A and supported in the interior of M_C taking F'' to a section. Hence there is a homeomorphism of M_C , supported in the interior of M_C and taking $F'' \cup A''$ to $F \cup A$. \square

Thus, the procedure that we use is as follows. We have applied a homeomorphism of M_C taking F to F' , until it has controlled weight. We have not moved A , and so its weight is unchanged. Suppose that F' and A admit a disc swap. Then it is possible to perform a disc swap without increasing the binding weight of the surfaces, as follows. Consider a simple closed curve of $F' \cap A$ that bounds a disc D in one of the surfaces, and suppose that among all such curves, the binding weight of D is minimal. Then we may assume that the interior of D is disjoint from $F' \cap A$. By the incompressibility of the surfaces, ∂D bounds a disc D' in the other surface. By our minimality assumption, the binding weight of D' is at least the binding weight of D . Hence, the disc swap that removes D' and replaces it by D does not increase the binding weight of the surfaces.

So, we may perform disc swaps to F' and A without increasing the binding weight, until no more disc swaps are possible. Then by Lemma 14.3, the resulting surfaces are strongly equivalent to $F \cup A$.

14.11. The solid toral components. The above argument dealt with M_{nst} . The solid toral components must be handled in a different way, since they do not necessarily contain a clean incompressible annulus that is not parallel to a clean annulus in ∂M . However, unlike other circle bundles, a solid torus has the advantage that cutting along a section decomposes it into a ball. Hence, we simply remove A and C from M_{st} , let S_{st} be a section in M_{st} , and then treat S_{st} in the same way as the remaining surfaces in the hierarchy.

15. REALISING THE HIERARCHY WITH POLYNOMIALLY BOUNDED BINDING WEIGHT

In this section, we prove the following result, which arranges an exponentially controlled hierarchy into admissible form with polynomially bounded binding weight. This key result is central to this paper, and relies on all the machinery developed so far, including normal surface theory and the structure of arc presentations.

Theorem 15.1. *Let K be a non-split link in S^3 , other than the unknot. Assign boundary pattern to $\partial N(K)$ as described in Theorem 3.3. Let $H = \{S_1, \dots, S_\ell\}$ be an adequately separating, $(\lambda, 2^{15})$ -exponentially controlled hierarchy for the exterior*

of K , where

$$\lambda = \sum_{i=1}^{\ell} -3\chi(S_i) + 6|S_i| + 6|S_i \cap P_i|.$$

Here, P_i is the boundary pattern in the exterior of the first $i-1$ surfaces. Let H' be obtained from $S_1 \cup \dots \cup S_{\ell}$ by the procedure in Section 14, which at the final stage removes the intersection between H and the interior of the solid toral components of the relevant circle bundle and then inserts weakly fundamental meridian discs. Let

$$q(x) = (10^{80}\lambda^8)^{\sum_{i=0}^{\ell} 22^i} x^{22^{\ell}}.$$

Let D be an arc presentation for K with arc index n . Then there is a sequence of at most $n^2 q(n)$ exchange moves and cyclic permutations, and a homeomorphism of the exterior of K that equals the identity on $\partial N(K)$ and that takes $\partial N(K) \cup H'$ to an admissible hierarchy with binding weight at most $q(n)$.

Proof. Let k be a positive integer that is at most the length of the hierarchy H . For $k < \ell$, let H_k be the partial hierarchy consisting of the first k surfaces of H , including $S_0 = \partial N(K)$. Let $H_{\ell} = H'$. Let

$$q_k(x) = (10^{80}\lambda^8)^{\sum_{i=0}^k 22^i} x^{22^k}.$$

We will show by induction on k that there is a sequence of at most $n^2 q_k(n)$ exchange moves and cyclic permutations and an isotopy of the exterior of K that takes $\partial N(K) \cup H_k$ to an admissible partial hierarchy with weight at most $q_k(n)$. Setting k to be the length ℓ of the hierarchy will clearly then prove the theorem.

The induction starts by performing at most $n+1$ stabilisations to D so that, for each component of $\partial N(K)$ with non-empty boundary pattern, the slope of this pattern equals the writhe of the relevant component of K . We then place $\partial N(K)$ into admissible form. We may do this simply by taking a thin regular neighbourhood of K . This requires no exchange moves or cyclic permutations, and the resulting binding weight of $\partial N(K)$ is at most $4n+2$. Note that $q_0(n) \geq 4n+2$.

We now establish the inductive step. Let $\{S_0, \dots, S_k\}$ be the surfaces of the partial hierarchy H_k , where $k < \ell$, and let $S = S_{k+1}$ be the next surface. Let P be the boundary pattern on the exterior of H_k . Suppose that H_k is admissible, and has binding weight at most w . We will show that there is a sequence of at most $10^{40}\lambda^4 w^{11}$ steps that takes $\partial N(K) \cup H_{k+1}$ to an admissible hierarchy with binding weight at most $10^{80}\lambda^8 w^{22}$. Thus, the binding weight is at most

$$10^{80}\lambda^8 (q_k(n))^{22} = 10^{80}\lambda^8 (10^{80}\lambda^8)^{\sum_{i=1}^{k+1} 22^i} n^{22^{k+1}} = (10^{80}\lambda^8)^{\sum_{i=0}^{k+1} 22^i} n^{22^{k+1}} = q_{k+1}(n).$$

Each step will use at most n^2 exchange moves and cyclic permutations. So, the number of exchange moves and cyclic permutations so far is at most

$$n^2 q_k(n) + n^2 10^{40}\lambda^4 (q_k(n))^{11} < n^2 10^{80}\lambda^8 (q_k(n))^{22} = n^2 q_{k+1}(n).$$

This will establish the inductive step.

So, we now suppose that H_k is in admissible form with binding weight at most w . In particular, it has no annular tiles.

Step 1. Modify H_k so that it is well-spaced.

We use the procedure described in Section 7.2 to make the hierarchy well-spaced. This increases its binding weight to at most $7w$.

Step 2. Constructing a triangulation for the exterior of H_k .

A triangulation for the exterior M of H_k is described in Section 6.2. Lemma 6.3 states that it has at most $t = 112w(7w - \sum_{i=1}^k \chi(S_i)) \leq 112w(7w + \lambda)$ tetrahedra.

Step 3. Realising S as an exponentially controlled normal surface.

By assumption, S can be realised as a $(\lambda, 2^{15})$ -exponentially controlled surface, and so the weight of S is at most $\lambda 2^{15t}$.

Step 4. Homotope S to a nearly embedded surface \hat{S} with boundary in an arc presentation.

As in Section 6.7, we now homotope S to a nearly embedded alternative admissible surface \hat{S} with binding weight at most $200(7w)^3(7w + \lambda)^2 \lambda 2^{15t}$. It is normal with respect to the handle structure \mathcal{K} .

Step 5. Forming the branched surface B that carries S .

In Section 8.5, a branched surface B was described that carries S . Let $N = N(B)$ be its thickening and let \mathcal{H} be its natural handle structure. We set X to be $\partial_h N$, and set Y to be empty. In Section 10.2, we defined the admissible envelope $(N, X, Y, P \cap N, \mathcal{H}, S)$. We also initially set $(N', X, Y', P \cap N', \mathcal{H}', S')$ to be equal to $(N, X, Y, P \cap N, \mathcal{H}, S)$ and note that these form nested admissible envelopes. In Section 10.6, we defined complexities for these handle structures. By Lemma 10.4, these complexities satisfy

$$c(\mathcal{H}) \leq c_+(\mathcal{H}') \leq 11648(7w)^2(7w + \lambda).$$

We will modify the envelopes as described in Section 10, but their complexities will not increase. The surface S' may have some cut vertices. We let $S' \setminus S'_{\text{thin}}$ be the result of cutting S' along these cut vertices and discarding any bigon tiles. Let E be the Euclidean subsurface of $S' \setminus S'_{\text{thin}}$. Let S'' be obtained from $S' \setminus S'_{\text{thin}}$ by removing any toral or clean annular components of S' lying in E with at most $23552e^2$ vertices. Let e be the number of approximate star types of tiles of S'' . By Propositions 10.10 and 10.11,

$$e \leq 16c_+(\mathcal{H}) + 32c(\mathcal{H}') \leq 559104(7w)^2(7w + \lambda).$$

Step 6. Finding many vertices with small valence.

Let S_+ be the double of S'' . By Theorem 13.1, when $k < \ell$, the number of non-cut vertices of S_+ with valence at most 3 is at least

$$\frac{w_\beta(S'')}{31\pi(2392e)^2} + \frac{72}{31}\chi(S) - \frac{5}{31}|S \cap P| \geq \frac{w_\beta(S'')}{31\pi(2392e)^2} - \lambda.$$

At the final step, we might be dealing with a broken section, in which case, the number of non-cut vertices of S_+ with valence at most 3 is (in the terminology of Section 14) at least

$$\begin{aligned} & \frac{w_\beta(S'')}{31\pi(2392e)^2} + \frac{72}{31}\chi(S) - \frac{5}{31}|S \cap P'| - \frac{5}{31}|S \cap C| \\ & \geq \frac{w_\beta(S'')}{31\pi(2392e)^2} - \lambda - \frac{5}{31}|S \cap P|96(7w)(7w + \lambda) \\ & \geq \frac{w_\beta(S'')}{31\pi(2392e)^2} - 388w^2\lambda. \end{aligned}$$

Their stars come in at most e approximate types. Thus, in S'' , we may find a collection of at least

$$\frac{1}{e} \left(\frac{w_\beta(S'')}{62\pi(2392e)^2} - 194w^2\lambda \right)$$

low-valence vertices all with parallel stars.

Step 7. Modifying S and S' and their nested admissible envelopes.

In Section 10, procedures were described which reduced the binding weight. Each was applied to a collection of parallel stars with low valence. Thus, the binding weight of S'' reduces by at least

$$\frac{1}{e} \left(\frac{w_\beta(S'')}{62\pi(2392e)^2} - 194w^2\lambda \right) \geq \frac{1}{e} \left(\frac{w_\beta(S'')}{10^{10}e^2} - 194w^2\lambda \right)$$

So $w_\beta(S'') - 194 \cdot 10^{10}\lambda w^2e^2$ decreases by at least a factor $(1 - 1/(10^{10}e^3))$, for the following reason:

$$\begin{aligned} & w_\beta(S'') - 194 \cdot 10^{10}\lambda w^2e^2 - w_\beta(S'')/(10^{10}e^3) + 194\lambda w^2/e \\ & \leq (1 - 1/(10^{10}e^3))(w_\beta(S'') - 194 \cdot 10^{10}\lambda w^2e^2). \end{aligned}$$

We perform these operations until $w_\beta(S'') - 194 \cdot 10^{10}\lambda w^2e^2 < 1$, in other words until $w_\beta(S'') \leq 194 \cdot 10^{10}\lambda w^2e^2$.

Step 8. Reintroducing the toral and annular components.

When forming S'' , toral and annular components of S'_{thin} were removed. Each had binding weight at most $23552e^2$. The number of such components is at most $2e$, since this is an upper bound for the maximal number of disjoint surfaces carried by B_E . The above modifications have not increased their binding weight. So, we may add them back to the surface, giving a surface with binding weight at most

$$194 \cdot 10^{10}\lambda w^2e^3 + 47104e^3 \leq 2 \cdot 10^{12}\lambda w^2e^3.$$

Step 9. Counting the number of iterations.

Let x be the number of times that we applied Steps 6 and 7. Each time such a step was repeated, $w_\beta(S'') - 194 \cdot 10^{10}\lambda w^2e^2$ was multiplied by a factor of at most $(1 - 1/(10^{10}e^3))$. The initial binding weight was at most

$$200(7w)^3(7w + \lambda)^2\lambda 2^{15t} \leq 200(7w)^3(7w + \lambda)^2\lambda 2^{1792w(7w + \lambda)} \leq 2^{11760w(w + \lambda)}.$$

Since, after $x - 1$ steps, the binding weight is still more than 1, we deduce that

$$2^{11760w(w + \lambda)}(1 - 1/(10^{10}e^3))^{x-1} \geq 1$$

and hence that

$$(\log 2)11760w(w + \lambda) + (x - 1)\log(1 - 1/(10^{10}e^3)) \geq 0.$$

Rearranging and using the inequality $\log(1 - y) \leq -y$ for $0 < y < 1$ gives

$$(x - 1)/(10^{10}e^3) \leq (\log 2)11760w(w + \lambda).$$

Hence,

$$x \leq 1 + 10^{10}(\log 2)11760(559104)^3 7^9 w^7 (w + \lambda)^4 \leq 10^{40}\lambda^4 w^{11}.$$

The final inequality uses that $x + y \leq xy$ for integers $x, y \geq 2$. As explained above, the total number of exchange moves and cyclic permutations is therefore at most $n^2 q_{k+1}(n)$.

Step 10. Bounding the binding weight of the partial hierarchy.

Each step increases the binding weight of the hierarchy $\{S_0, \dots, S_k\}$ by at most 6. Hence, after these steps, it has binding weight at most $w + 6 \cdot 10^{40} \lambda^4 w^{11} \leq 7 \cdot 10^{40} \lambda^4 w^{11}$.

Step 11. Controlling the thin part of S' .

By Proposition 12.1 and Remark 12.2, there is a pattern-isotopy of $S'' \cup S'_{\text{thin}}$ taking it to a surface S_{k+1} with binding weight at most

$$\begin{aligned} (2e + 1 + 7 \cdot 10^{40} \lambda^4 w^{11}) w_\beta(S'') &\leq (2e + 1 + 7 \cdot 10^{40} \lambda^4 w^{11}) \cdot 2 \cdot 10^{12} \lambda w^2 e^3 \\ &\leq 10^{54} \lambda^5 w^{13} e^3 \leq 10^{79} \lambda^8 w^{22}. \end{aligned}$$

In the case of where S_{k+1} is a broken section, we need to do disc swaps with S_{k+1} and the annuli A so that $S_{k+1} \setminus A$ is meridian discs for the solid tori forming the circle bundle. As explained in Section 14.10, this can be achieved without increasing the binding weight of S_{k+1} and A . Furthermore, since this isotopy takes place in the exterior of K , no exchange moves or cyclic permutations are required. Finally, the section is perturbed to a broken section, which might increase its weight by a factor of at most 3, but it remains at most $10^{80} \lambda^8 w^{22}$. \square

16. ISOTOPING A LINK THROUGH A HANDLE STRUCTURE

Let \mathcal{H} be a handle structure of the 3-sphere. We say that a link K *respects* the handle structure \mathcal{H} if the following all hold:

- (1) K lies in the union of the 0-handles and 1-handles of \mathcal{H} ;
- (2) it intersects each 0-handle in a properly embedded 1-manifold;
- (3) it intersects each 1-handle $D^1 \times D^2$ in arcs, each of which is of the form $D^1 \times \{*\}$ for some point $* \in \text{int}(D^2)$.

Our aim in this section is to start with a link K that respects \mathcal{H} and to perform an ambient isotopy of K taking it into a single 0-handle. This can clearly be done. However, we wish to perform this isotopy in a controlled way. We therefore introduce the following moves. Each starts and ends with K respecting the handle structure.

Move 1: Sliding across a 2-handle

Let H_2 be a 2-handle of \mathcal{H} , and let $N(H_2)$ be a small regular neighbourhood of it. So, for each 0-handle and 1-handle to which H_2 is attached, $N(H_2)$ intersects that handle in a collection of balls. Note that $N(H_2)$ has a natural product structure as $D^2 \times D^1$. Suppose that $K \cap N(H_2)$ consists of a single arc of the form $\alpha \times \{*\}$, where α is an arc in ∂D^2 and $\{*\}$ is a point in $\text{int}(D^1)$. Suppose also that this arc $K \cap N(H_2)$ lies in a single 0-handle. Then we isotope $\alpha \times \{*\}$ across the disc $D^2 \times \{*\}$, replacing it with the arc $(\partial D^2 \setminus \alpha) \times \{*\}$. (See Figure 51.)

Move 2: Sliding along a 1-handle

Let H_1 be a 1-handle of \mathcal{H} , and let $N(H_1)$ be a regular neighbourhood of H_1 . This has a product structure $D^1 \times D^2$. Suppose that all but one arc of $K \cap N(H_1)$ is of the form $D^1 \times \{*\}$ for some point $* \in \text{int}(D^2)$. Suppose that the remaining arc γ is a concatenation of three arcs $\alpha \times \{p_1\}$, $\alpha \times \{p_2\}$ and $\{q\} \times \beta$, where α is an arc in D^1 such that $\alpha \cap \partial D^1$ is a single point, where q is the remaining point $\partial \alpha \setminus \partial D^1$, and where β is an arc in the interior of D^2 joining p_1 to p_2 . Suppose also that γ intersects H_1 in two arcs. Then, we replace γ by an arc γ' that is disjoint from H_1 . Once again, γ' is the concatenation of three arcs $\alpha' \times \{p_1\}$, $\alpha' \times \{p_2\}$ and

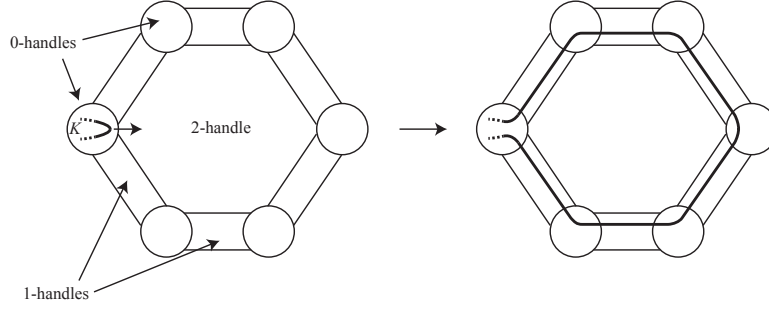


FIGURE 51. Sliding across a 2-handle

$\{q'\} \times \beta$, where α' is a subarc of α . We also permit the reverse of this move. (See Figure 52.)

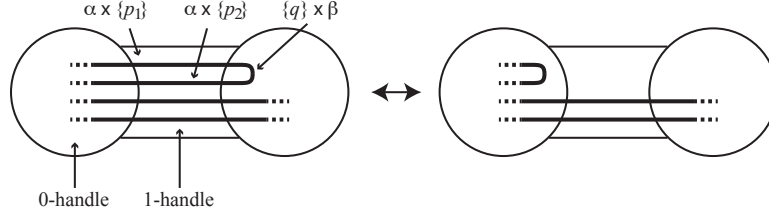


FIGURE 52. Sliding across a 1-handle

Move 3: Handle-preserving isotopy

The final move consists of an ambient isotopy of the 3-sphere, that preserves each handle, and such that K respects \mathcal{H} throughout.

Proposition 16.1. *Let K be a link that respects a handle structure \mathcal{H} on the 3-sphere. Then there is a sequence of the above moves, after which K lies in the interior of a single 0-handle.*

Proof. We start by defining a system of arcs and discs embedded in \mathcal{H} , which we term the *specified arcs and discs*. The specified arcs are easily defined: they are just the co-cores of the 2-handles. In other words, for a 2-handle $D^2 \times D^1$, the specified arc is $\{0\} \times D^1$, where $\{0\}$ is the central point in D^2 .

We now need to define the specified discs. Each disc will include the co-core of a 1-handle, which is $\{*\} \times D^2$ for some point $\{*\}$ in $\text{int}(D^1)$. The specified discs will also include parts lying in the 2-handles, as follows. For a 2-handle $D^2 \times D^1$, the intersection with the co-cores of the 1-handles is of the form $P \times D^1$, where P is a finite collection of points in ∂D^2 . For each point p in P , there is a radial arc α in D^2 running in a straight line from p to 0. We term $\alpha \times D^1 \subset D^2 \times D^1$ a *strip*. (See Figure 53.) For each co-core of a 1-handle, we attach all the strips to which it is incident. The result is a *specified disc*. (See Figure 54.)

Using the specified arcs and discs, one can form a handle structure \mathcal{H}' for the 3-sphere as follows. The 3-handles of \mathcal{H}' and \mathcal{H} are the same. A thin regular neighbourhood of each specified arc forms a 2-handle of \mathcal{H}' . A thickening of each specified disc forms a 1-handle. The remainder of the manifold is the 0-handles of

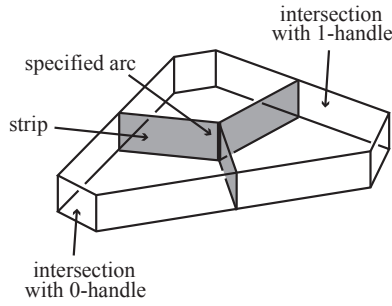


FIGURE 53. The specified arc and strips within a 2-handle

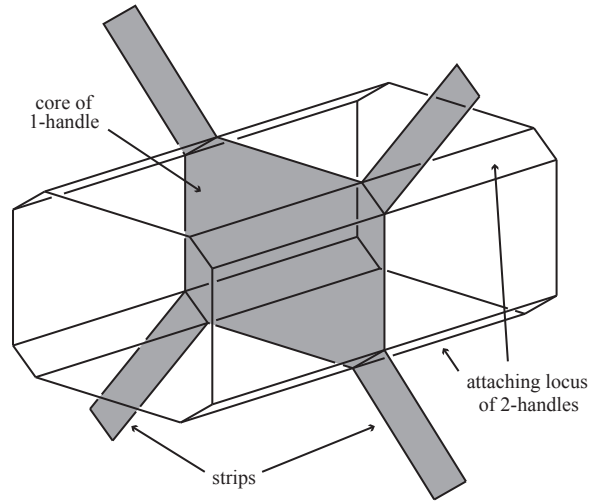


FIGURE 54. A specified disc associated with a 1-handle

\mathcal{H}' . Now \mathcal{H} and \mathcal{H}' are, in some sense, the same handle structure. More precisely, there is a homeomorphism $h: S^3 \rightarrow S^3$ taking each handle of \mathcal{H}' to a handle of \mathcal{H} . This homeomorphism expands each 2-handle of \mathcal{H}' until it fills the 2-handle of \mathcal{H} that contains it. It also lengthens each 1-handle of \mathcal{H}' .

We now pick an isotopy of S^3 that takes $h^{-1}(K)$ into a 0-handle of \mathcal{H}' . By general position, we may assume that the link misses the specified arcs, except at finitely many points in time, when it crosses through some specified arc transversely. Hence, we may also arrange that, at this point in time, it moves across a 2-handle of \mathcal{H}' in a single slide. Similarly, the link intersects each specified disc transversely in finitely many points, except at finitely many instances, where one of two possible things happens. One possibility is that the link approaches the specified disc, then just touches it, and then goes through it, creating two new points of intersection. Hence, the link moves across a 1-handle of \mathcal{H}' in a slide. The other possibility is the reverse of this. This isotopy of S^3 gives a 1-parameter family of homeomorphisms of S^3 that starts at the identity. Applying h to each of these gives an isotopy, that moves K into the interior of a 0-handle of \mathcal{H} , and that is of the required form. \square

17. FROM THE ISOTOPY TO REIDEMEISTER MOVES

In this section, we complete the proof of our main theorem. We are given a diagram D of a link, with link type K , where D has c crossings. We pick a fixed diagram D' for K . We need to prove that there is a polynomial p_K such that D and D' differ by a sequence of at most $p_K(c)$ Reidemeister moves.

Let μ be the constant in Theorem 3.6. We first pick a fixed $(\mu, 2^{15})$ -exponentially controlled hierarchy H for the exterior of K , which exists by Theorem 3.6. We also ensure that it is adequately separating.

We isotope D to form a rectangular diagram. By Lemma 5.1, we can arrange that this rectangular diagram has arc index n where $n \leq (81/20)c$. Since only an isotopy of the plane was used, no Reidemeister moves are needed at this stage. This rectangular diagram determines an arc presentation also with arc index n . Theorem 15.1 provides a polynomial q , which depends only on K and H , such that, after at most $n^2q(n)$ cyclic permutations and exchange moves, the hierarchy has been placed in admissible form with binding weight at most $q(n)$.

17.1. The handle structure arising from the hierarchy. The hierarchy H determines a handle structure for the exterior of K , as follows. The union of $\partial N(K)$ and H is a 2-complex. Thicken this to a handle structure. The remainder of the exterior of K is a collection of 3-balls, which we take to be 3-handles of the handle structure.

We now extend this to a handle structure \mathcal{H} for S^3 , as follows. Pick a small regular neighbourhood of K , denoted $N_{\text{small}}(K)$, lying in the interior of $N(K)$. So $N(K) \setminus N_{\text{small}}(K)$ is a disjoint union of copies of $T^2 \times I$, one for each component of K . We give each component of $N_{\text{small}}(K)$ a handle structure as a 0-handle and a 1-handle. Pick a meridional simple closed curve on each component of $\partial N(K)$, as well as a simple closed curve on $\partial N(K)$ that has slope equal to a longitude plus some number of meridians. Let m and l be these meridian and integral curves. On each component of $\partial N(K)$, pick m and l so that they respect the handle structure on this torus, and so that they intersect in a single point p in some 0-handle. Between p and a point on the 0-handle of $N_{\text{small}}(K)$, attach a 1-handle that is vertical in the $T^2 \times I$ region. Now attach a 2-handle that runs along $m - p$, then along this 1-handle, then around the 0-handle intersecting K , and then back along the 1-handle. Similarly attach a 2-handle that runs along $l - p$, then along the vertical 1-handle, then along the 1-handle intersecting K , then back along the vertical 1-handle. These two 2-handles and the vertical 1-handle have decomposed the $T^2 \times I$ region into a ball, which we take to be a 3-handle.

We term \mathcal{H} as the handle structure *arising from H* . It depends on H as well as the choice of curves m and l . Note that K respects \mathcal{H} , in the sense of Section 16.

We now pick another fixed copy K' of K lying within a single 0-handle of \mathcal{H} . By Proposition 16.1, there is a sequence of moves, as described in Section 16, which take K to K' . Since H has polynomially bounded weight, it is not very surprising that one should be able to pass from the rectangular diagram for K to a diagram for K' using a polynomially bounded number of Reidemeister moves. So, if D' and K' are chosen so that a suitable projection of K' is D' , then this will complete the proof of Theorem 1.1. In the remainder of this section, we give some further details of this argument.

17.2. From an arc presentation to a rectangular diagram. We have K sitting inside the 3-sphere in a specific way, arising from its arc presentation. We now recall the construction from [6] that arranges K so that its projection to a plane is the associated rectangular diagram.

The 3-sphere is the join $S_\phi^1 * S_\theta^1$, where S_ϕ^1 is the binding circle. Since K is an arc presentation, it intersects the binding circle at finitely points, the vertices of K , and it intersects each page \mathcal{D}_t either in the empty set or in a single open arc joining distinct vertices of K . We can arrange that each such open arc is a concatenation of two arcs, each of which runs straight from a vertex of K to the point $\mathcal{D}_t \cap S_\theta^1$. We may also assume that the vertices of K avoid the value $\phi = 0$, and that the pages \mathcal{D}_t intersecting K avoid $\theta = 0$.

We now perturb K off S_ϕ^1 and S_θ^1 as follows. We view each page \mathcal{D}_t as an open unit disc. We replace an arc of intersection $K \cap \mathcal{D}_t$, by removing its intersection with some small disc around the origin in \mathcal{D}_t and replacing it by an arc going around the circumference of this disc. We choose the arc so that it avoids taking the value $\phi = 0$. We do something similar near the binding circle. Instead of going straight to a vertex on the binding circle and then away again, we replace this by an arc transverse to the pages with constant ϕ value. We choose this arc so that it avoid the value $\theta = 0$.

Thus, K is now disjoint from $S_\theta^1 \cup S_\phi^1$. The complement of $S_\theta^1 \cup S_\phi^1$ is a copy of $S_\theta^1 \times S_\phi^1 \times (0, 1)$, where the first two factors are parametrised by the θ and ϕ coordinates respectively. Projection onto the first two factors will give the diagrammatic projection map for K . In fact, because K avoids $\theta = 0$ and $\phi = 0$, its image lies in $(0, 2\pi) \times (0, 2\pi)$.

We have arranged that K is a concatenation of three types of arc. One type has constant θ and ϕ values and so projects to a point in $(0, 2\pi) \times (0, 2\pi)$. One type, lying in a page and staying near S_θ^1 , has constant θ value. So it projects to a vertical arc in $(0, 2\pi) \times (0, 2\pi)$. The third type of arc lies near S_ϕ^1 and has constant ϕ value, and so projects to a horizontal arc. Thus, we have a projection of K to $(0, 2\pi) \times (0, 2\pi)$ consisting of a concatenation of horizontal and vertical arcs. Moreover, when two such arcs cross, the vertical arc is the over-arc. This is the required rectangular diagram for K .

When K is in this form, it inherits a natural framing from the diagram, and hence each component has a writhe, which is the integral slope of its framing. We perform some initial stabilisations to the rectangular diagram, so that the writhe of each component is equal to the integral slope of the relevant component of l .

17.3. A projection of the 1-skeleton. The hierarchy H determines a cell structure for the exterior of K . We will now create a projection for its 1-skeleton.

Each 0-cell lies on the binding circle S_ϕ^1 . It therefore has a specific ϕ -value s . So it corresponds to an arc $(0, 2\pi) \times \{s\}$ in $(0, 2\pi) \times (0, 2\pi)$. A regular neighbourhood of this 0-cell is a 0-handle of \mathcal{H} . We may initially assume that the projection of this 0-handle to the plane is a small regular neighbourhood of this arc. We may also identify the 0-handle with $D^2 \times I$, so that its projection onto the first factor is the diagrammatic projection map.

Each 1-cell of the cell structure is an arc in the boundary of some surface of the hierarchy. It is therefore in arc presentation. Hence, it projects to a concatenation of horizontal and vertical arcs in the plane. We thicken each 1-cell so that it projects to a regular neighbourhood of the union of these arcs.

This specifies the projection of the 0-cells and 1-cells in the exterior of K . In addition, there are the 0-handles and 1-handles that form $N_{\text{small}}(K)$ and the 1-handles that are vertical in $N(K) \setminus N_{\text{small}}(K)$. We may assume that each 0-handle is a regular neighbourhood of a vertex of K . Each 1-handle in $N_{\text{small}}(K)$ runs along the remainder of a component of K , which is in an arc presentation. So this specifies its projection. Each vertical 1-handle can be chosen so that it is a regular neighbourhood of an arc with constant ϕ -value. So it projects to a regular neighbourhood of a horizontal arc in the plane.

17.4. Location of the 2-handles. Aside from the 2-handles lying in $N(K)$, each other 2-handle is a thickened 2-cell, and this 2-cell is made up of a union of tiles. So it suffices to explain the location of each tile. A tile has a product foliation. Each leaf has a constant θ -value and runs between two vertices. Thus, as long as $\theta \neq 0$, the leaf specifies a vertical arc in the diagram. Thus, when the leaf where $\theta = 0$ is removed, the tile is a union of such arcs.

The two 2-handles in each component of $N(K)$ interpolate between an arc in $\partial N(K)$ of slope m or l and a corresponding arc in $\partial N_{\text{small}}(K)$. Initially, before the hierarchy is constructed, K is in an arc presentation with arc index n and $N(K)$ is a thin regular neighbourhood of it. If this had remained the case, it would be easy to realise the 2-cells in $N(K)$ as a union of tiles with binding weight at most $2n$. However, as we build the hierarchy H , we modify $N(K)$ in two ways. One modification is to perform exchange moves and cyclic permutations to K , and these have a corresponding effect on $N(K)$ and $N_{\text{small}}(K)$. This modification has no effect on the binding weight of the 2-cells in $N(K)$. The other modification is to perform wedge insertions along $\partial N(K)$. These expand $N(K)$ but leave $N_{\text{small}}(K)$ unchanged. So at each wedge insertion, the 2-handles in $N(K)$ might expand. This expansion increases the binding weight of their constituent 2-cells by at most 2. So the total binding weight after these modifications is at most $2q(n)$, because $q(n)$ is an upper bound for the number of wedge insertions.

17.5. Shrinking the 0-handles. As described above, it is natural to take each 0-handle so that it projects to the regular neighbourhood of a horizontal arc in the plane. It is convenient to shrink this 0-handle, so that it actually projects to a little disc, with the property that the interior of the disc is disjoint from the projection of the interior of the 1-handles. We need to extend the 1-handles attached to this 0-handle, but this can be achieved by adding a regular neighbourhood of a horizontal arc to each endpoint of each 1-handle.

17.6. From slides to Reidemeister moves. We have picked a sequence of slides across handles and handle-preserving isotopies that takes the initial copy of K in \mathcal{H} to the copy K' that lies in a 0-handle. We can realise each of these moves in 3-space. Projecting, we obtain a sequence of diagrams that starts with D and ends with D' . We therefore need to bound the number of Reidemeister moves that relate successive diagrams in this sequence. We will arrange that these diagrams are actually rectangular diagrams.

It will be useful to have an upper bound d on the arc index of in each rectangular diagram of this sequence. Each diagram arises from a fixed realisation of the link in the handle structure \mathcal{H} . Each 0-handle of \mathcal{H} is a copy of $D^2 \times I$ and is embedded in $(0, 2\pi) \times (0, 2\pi) \times (0, 1)$ in a way that respects their product structures. So, the intersection of the link K with this 0-handle has a specific projection to $(0, 2\pi) \times$

$(0, 2\pi)$. In particular, this projection consists of a bounded number of straight arcs, and we may assume that each of these is vertical or horizontal. This bound depends only on the arrangement of K in \mathcal{H} , and does not depend on the way that \mathcal{H} is embedded in 3-space. We now consider the intersection of the link with a 1-handle of \mathcal{H} . As observed in Section 17.3, this 1-handle is a regular neighbourhood of an arc α that is in an arc presentation, with arc index bounded above by the binding weight of the hierarchy. Now, K respects the product structure on this 1-handle, and therefore each arc of intersection between K and the 1-handle runs in a thin regular neighbourhood of α . Its diagrammatic projection is therefore a concatenation of at most w horizontal and at most w vertical arcs, where w is the binding weight. We may also ensure that the projections of the 0-handles are disjoint from each other. We can also arrange that the projection of each 1-handle avoids the projection of each 0-handle, except at the points where the 1-handle is attached to a 0-handle. However, distinct 1-handles may have projections that overlap. We therefore deduce that each rectangular diagram in this sequence has arc index at most $d = c_1w + c_2$, where c_1 and c_2 are constants depending only on K , and w is the binding weight of the hierarchy.

We now find an upper bound on the number of Reidemeister moves that result from each of the moves described in Section 16. We start with the process of sliding along a 1-handle, because this is the simplest. Here, the knot remains in the union of the 0-handles and 1-handles. Moreover, its intersection with the 0-handles barely changes. In particular, the projections of the part of the link that lies in the 0-handles does not change, apart from near a small regular neighbourhood of the 1-handle that is involved in the move. As K is slid through the 1-handle, a series of type 2 Reidemeister moves are performed. These occur when the core curve α of the 1-handle and some other point of K have the same projection. Therefore, the number of type 2 Reidemeister moves is at most the number of points of intersection between the projections of α and K , which is at most $2wd$.

We next consider the case of handle-preserving isotopy. This changes the position of the link within the 0-handles and the 1-handles. However, we can perform these isotopies separately, first of all changing the position with the 0-handles and then changing the position within a small regular neighbourhood of the 1-handles. The isotopies supported within the 0-handles have the effect of isotoping the link, and in so doing, result in Reidemeister moves. The way that the diagram changes within the image of a 0-handle depends only on the choice of isotopy within \mathcal{H} , and not on the way that it is embedded. Thus, we deduce that the number of Reidemeister moves for this part of the process is a constant, independent of w . The next part of the process isotopes the link in a regular neighbourhood of the 1-handles, respecting the product structure on the 1-handles throughout. This may be achieved using a Reidemeister 2 move, followed by several Reidemeister 3 moves. The number of type 3 moves is equal to the number of points of projection between the core α of the 1-handle and the projection of the link in the remaining handles. Hence, the number of moves is at most a quadratic function of w .

Finally, we consider the situation where the link is slid across a 2-handle. It is convenient to consider the reverse of this move. This will require the same number of Reidemeister moves. The 2-handle is a thickening of a 2-cell C that is nearly admissible. Initially, K runs around ∂C . We will isotope K across C , one tile at a time. Let C_- be the union of the tiles that have not been moved across. Then we

will arrange that C_- is a disc (except after the final step, when it is empty). We will also arrange that K runs around ∂C_- .

We may pick a tile T in C_- , such that after removing this tile from C_- , the result is still a disc. This is for the following reason. Consider any tile that shares at least one edge with ∂C_- . If we were to remove this tile from C_- and create something other than a disc, then at least one edge of the tile would start and end on ∂C_- and with its interior lying in the interior of C_- . Consider the union of all such arcs, over all tiles sharing at least one edge with ∂C_- . Each is properly embedded in C_- and the interiors of these arcs are disjoint. So we may find such an arc β that is outermost in C_- . It separates off a sub-disc of C_- with interior disjoint from the arcs. Pick any tile in that disc that shares at least one edge with ∂C_- . (Avoid the tile containing β , unless it is the only tile in that sub-disc.) Then we can let T be this tile.

Suppose first that T has exactly two edges of ∂C_- in its boundary, and that these intersect at a saddle of C . Then we may slide K across this tile, maintaining it is an arc in a page throughout. In the diagram, each such arc is vertical. Hence, we obtain a 1-parameter family of vertical arcs. As the arc passes over some other vertical arc, this is achieved by an exchange move. By Lemma 5.2, this can be performed using at most d Reidemeister moves. It might be the case that one of the arcs in the tile has $\theta = 0$. At this point, the arc in the diagram switches from a vertical arc with θ just above 0 to one just below 2π , or vice versa. This is a cyclic permutation, using at most d^2 Reidemeister moves. Hence, in total, the number of Reidemeister moves is at most $2d^2$.

Now suppose that T has exactly one edge in ∂C_- . This runs between a saddle and a vertex. Consider the other edge of the tile emanating from this saddle. Then we may introduce a finger to K so that it runs along this edge and back again. This is achieved by sliding across a 1-handle. In doing so, we need at most $2wd$ type 2 Reidemeister moves. We can now slide K across the tile, as above.

The situation where there are three or four edges of ∂C_- is similar. We can slide K across the tile, and then perform the reverse of a finger move. When the tile has two edges of ∂C_- in its boundary, but they intersect in a vertex, we introduce a finger running along one the remaining edges, then we slide across the tile, and then we remove a finger. This requires at most $2d^2 + 4wd$ type 2 Reidemeister moves.

Each 2-handle is a thickening of a 2-cell C , and this consists of at most $4w$ tiles, for the following reason. The number of tiles in C is at most the number of half-tiles plus twice the number of full square tiles. This is equal to $4x^i(C)$ and $2x^b(C)$, where $x^i(C)$ and $x^b(C)$ denote the number of interior saddles and boundary saddles in the 2-cell. By Lemma 5.5, this is equal to $4v^i(C) + 2v^b(C) - 2$ where $v^i(C)$ and $v^b(C)$ denote the number of vertices in the interior and boundary of C . This is at most $4w$.

Sliding across each tile uses at most $2d^2 + 4wd$ Reidemeister moves, we have used at most $8d^2w + 16w^2d$ Reidemeister moves in total, which is at most a cubic function of w .

Thus, we have proved the following.

Proposition 17.1. *Let K be a non-split link, and let H be a hierarchy for the exterior of K . Let D' be a fixed diagram for K . Then there is a constant k , depending only on K , H and D' with the following property. If K is an arc presentation, and H is in admissible form with binding weight w , then there is a sequence of at most*

kw^3 Reidemeister moves taking the rectangular diagram for K that arises from the arc presentation to D' .

17.7. Split links. We are nearly in a position to prove the main theorem of this paper, Theorem 1.1. But first we explain how to deal with links that are split.

Theorem 17.2. *Let L be a split link that is a distant union $L_1 \sqcup \cdots \sqcup L_s$ of non-split links. Let D be a rectangular diagram for L with arc index n . Then there is a sequence of at most $(11n)^{11}$ exchange moves, cyclic permutations and destabilisations taking D to a disconnected rectangular diagram, where the components are diagrams of L_1, \dots, L_s .*

Proof. Theorem 1.4 of [17] provides a sequence of at most $(14n)^{10}$ exchange moves, cyclic permutations and destabilisations taking D to a disconnected rectangular diagram. Note that these moves do not increase the arc index of the rectangular diagram, and so this is still at most n . If any of the components of this diagram represents a split link, then at most $(14n)^{10}$ further such moves makes this disconnected. Repeating in this way, we eventually reach the required rectangular diagram where the components are diagrams of L_1, \dots, L_s . Since each component has arc index at least 2, the number of times we repeat is at most $n/2$. So the total number of moves is at most $(n/2)(14n)^{10} \leq (11n)^{11}$. \square

17.8. Proof of the main theorem.

Proof of Theorem 1.1. Let D be some diagram for K with crossing number c . In the case of the theorem we will consider $D = D_1$ or $D = D_2$. Let D' be some fixed diagram for K . We will provide an upper bound on the number of Reidemeister moves needed to relate D to D' . Hence, this will give an upper bound on the number of Reidemeister moves needed to relate D_1 to D_2 .

If K is split, then write it as a distant union $L_1 \sqcup \cdots \sqcup L_s$ of non-split links. We may assume that D' is a distant union of diagrams for L_1, \dots, L_s . The given diagram D can be isotoped to a rectangular diagram with arc index $n \leq (81/20)c$, by Lemma 5.1. By Theorem 17.2, this may be converted to a disconnected rectangular diagram, where the components are diagrams of L_1, \dots, L_s , using at most $(11n)^{11}$ exchange moves, cyclic permutations and destabilisations. It therefore suffices to consider these diagrams individually.

Thus, we may assume that K is non-split. We may also assume that K is not the unknot, since the main result of [17] establishes that $p_K(c) = (236c)^{11}$ works in this case.

Let H be a fixed adequately separating, $(\mu, 2^{15})$ -exponentially controlled hierarchy for the exterior of K , which exists by Theorem 3.3 and Remark 3.7. Here, μ is the constant from Theorem 3.6. This determines a handle structure for the exterior of K . Also let m and l be fixed curves on each component of $\partial N(K)$ that intersect once, that have meridional and integral slope, and that respect the handle structure. These then determine an associated handle structure \mathcal{H} for the 3-sphere. Let f be the sum of the absolute values of the integer slopes of l .

As above, the diagram D can be isotoped to a rectangular diagram with arc index at most $(81/20)c$, by Lemma 5.1. We then perform stabilisations so that each component of K has writhe equal to the framing of the relevant component of l . This requires at most $f + c$ stabilisations. Let $n \leq f + (101/20)c$ be the resulting arc index. Theorem 15.1 provides a polynomial q , which depends only on K , H

and D' , such that, after at most $n^2q(n)$ cyclic permutations and exchange moves, the hierarchy has been placed in admissible form with binding weight at most $q(n)$. Each of these modifications to the diagram requires at most n^2 Reidemeister moves.

By Proposition 17.1, this diagram can be converted to D' using at most $k(q(n))^3$ Reidemeister moves, where k depends only on K , H and D' .

So we have converted D to D' using at most the following number of Reidemeister moves:

$$\begin{aligned} & f + c + (f + (101/20)c)^4 q(f + (101/20)c) + k(q(f + (101/20)c))^3 \\ & \leq 2k(q(f + (101/20)c))^3. \end{aligned}$$

So we may set $p_K(c) = 2k(q(f + (101/20)c))^3$. \square

Remark 17.3. The constant k in the above proposition is algorithmically computable. Using the hierarchy H and the choice of meridians m and integral curves l , the handle structure \mathcal{H} for the 3-sphere may be constructed. The initial position of the link K within \mathcal{H} is determined. There is certainly some isotopy of the 3-sphere taking K to a link K' that lies within a 0-handle of \mathcal{H} . Such an isotopy can always be found eventually, since the isotopy is a level-preserving simplicial map $S^3 \times [0, 1] \rightarrow S^3 \times [0, 1]$ and all possible such simplicial maps can be enumerated. Eventually one will be found that forms an isotopy taking K into a 0-handle. We can then let K' be the resulting link in the 0-handle. The 0-handle has a product structure as $D^2 \times [0, 1]$, and the projection onto D^2 gives a diagram D'' for K , possibly after a small perturbation of K' . The procedure that is used above for creating the moves from the isotopy is deterministic, and so it specifies a sequence of Reidemeister moves taking D to D'' . Finally, since D'' and D' are both diagrams of the same link K , and so some sequence of Reidemeister moves relating them may always be found.

The final step is algorithmic, but it does not provide a good *a priori* bound for the number of Reidemeister moves relating D'' and D' . However, if instead we had initially chosen D' to be equal to D'' , then this step would not have been required.

18. EXPLICIT POLYNOMIALS

We will want to determine an actual value for the constant k for the figure-eight knot and for torus knots. In order to do so, we will need a more quantified argument.

In the previous section, we defined the handle structure of the 3-sphere arising from a hierarchy. It will be helpful to introduce the following slightly more general version. As previously, we start with the handle structure on the exterior of K arising from the hierarchy and then attach handles in $N(K)$. Also as before, we pick a union l of integral curves, one on each component of $\partial N(K)$, that respects the handle structure on the boundary. But now we pick multiple disjoint meridians m on $\partial N(K)$, where we require each component of $\partial N(K)$ to contain at least one of these meridians, but maybe more than one. Each of the meridians respects the handle structure and we require that each intersects l exactly once.

We now give $N_{\text{small}}(K)$ a handle structure consisting of $|m|$ 0-handles and $|m|$ 1-handles, where $|m|$ is the number of meridians. In $N(K) \setminus N_{\text{small}}(K)$, we insert $|m|$ vertical 1-handles, each running from one of the 0-handles of $N_{\text{small}}(K)$ to a point of intersection between m and l . Now insert $|m|$ 2-handles, each running around

some component of m , down a vertical 1-handle, around a 0-handle in $N_{\text{small}}(K)$ and then back up the vertical 1-handle. Also insert $|m|$ 2-handles, each running along a component of $l \setminus m$, then along a vertical 1-handle, then along a 1-handle in $N_{\text{small}}(K)$ and then back up another vertical 1-handle. Finally, we insert $|m|$ 3-handles into $N(K) \setminus N_{\text{small}}(K)$. We also say that this is a handle structure on the 3-sphere *arising from the hierarchy*. As before, it also depends on extra data, specifically the choice of integral curves l and meridians m .

We will also make an extra hypothesis on the handle structure \mathcal{H} on the 3-sphere, which we now describe.

Let \mathcal{H} be a handle structure of 3-manifold satisfying Convention 4.1. A handle structure \mathcal{H}' is obtained from \mathcal{H} by a *collapse* if:

- (1) there are handles H_- and H_+ of \mathcal{H} , where the index of H_+ is one more than the index of H_- , and where the core of H_+ intersects the co-core of H_- once; $H_- \cup H_+$ is therefore a 3-ball; let D_- be the union of points in $\partial(H_- \cup H_+)$ where H_- and H_+ are attached to handles of lower index, and let D_+ be $\partial(H_- \cup H_+) \setminus D_-$;
- (2) the handles H_- and H_+ are removed;
- (3) any handles of higher index that are incident to D_+ are instead attached along D_- using the natural identification between D_- and D_+ .

Thus, we permit handle pairs to be collapsed even when handles of higher index are attached to them. We say that a collapse is *elementary* if, in the above definition, D_+ lies in the boundary of the 3-manifold, or equivalently, H_- and H_+ are disjoint from any other handles of higher index. We say that \mathcal{H} is *completely collapsible* if there is a sequence of collapses that takes it to a single 0-handle.

Proposition 18.1. *If a handle structure \mathcal{H} of a 3-manifold is collapsible, then there is a sequence of elementary collapses that takes it to a single 0-handle.*

Proof. This is by induction on the number of handles. The induction starts trivially when \mathcal{H} has a single handle, which must be a 0-handle, and so no elementary collapses are required.

So consider a collapsible handle structure \mathcal{H} , and suppose that the lemma is true for all collapsible handle structures with fewer handles than \mathcal{H} .

By assumption, \mathcal{H} has a sequence of collapses taking it to a 0-handle. We claim that there is such a sequence where all the $3/2$ collapses happen first, then all the $2/1$ collapses and then all the $1/0$ collapses. This just follows from the easy observation that we may interchange the order of successive collapses if a $1/0$ collapse happens just before a $3/2$ or $2/1$ collapse or a $2/1$ collapse happens before a $3/2$ collapse. The handle structure that results after performing such a pair of collapses does not depend on the order in which they happen. So, we now assume that the collapses to \mathcal{H} are ordered in this way.

We may similarly interchange the order of the $3/2$ collapses, and we will feel free to do so, for the following reason. We form a graph, using only the 2-handles and 3-handles that are removed in the $3/2$ collapses. The vertices of the graph correspond to the 3-handles and to the components of intersection between the boundary of the manifold and the 2-handles that are involved in the $3/2$ collapses. The edges of the graph correspond to the 2-handles that are removed in these collapses. Each 2-handle that is removed in the $3/2$ collapses must be incident to at least one 3-handle. When a $3/2$ collapse is performed, the homotopy type of this

graph is unchanged, except when a $3/2$ collapse removes an entire component of the graph. Hence, the graph is forest. In particular, it has more vertices than edges. However, there is the same number of 2-handles as 3-handles involved in the $3/2$ collapses. Therefore, there is a vertex corresponding to a component of intersection between the manifold boundary and a 2-handle involved in a $3/2$ collapse. Hence, this 2-handle and the adjacent 3-handle admit an elementary collapse. In this way, there is a sequence of elementary collapses that removes all the 3-handles.

We now consider the first $2/1$ collapse. Say this involves handles H_2 and H_1 , and let \mathcal{H}' be the resulting handle structure. If this was the final $2/1$ collapse, then it was in fact an elementary collapse. Then, inductively, \mathcal{H}' admits a sequence of elementary collapses taking it to a 0-handle. So, we may assume that the next collapse is also a $2/1$ collapse. Inductively, \mathcal{H}' admits a sequence of elementary collapses taking it to a 0-handle. So, we may assume that the next collapse to \mathcal{H}' is an elementary collapse, involving handles H'_2 and H'_1 . Note that H'_1 and H'_2 are also handles in \mathcal{H} .

Suppose first $H'_2 \cup H'_1$ is disjoint from $H_2 \cup H_1$ in \mathcal{H} . Then we may interchange the order of the collapses. Let \mathcal{H}'' be the handle structure obtained from \mathcal{H} by collapsing H'_2 and H'_1 . Note that \mathcal{H}'' is obtained from \mathcal{H} by an elementary collapse. Since \mathcal{H}'' collapses to \mathcal{H}' , it is also totally collapsible. Hence, by induction, there is a sequence of elementary collapses taking \mathcal{H}'' to a 0-handle, as required.

So suppose now that $H'_2 \cup H'_1$ intersects $H_2 \cup H_1$. Since handles of the same index are disjoint, there are two possibilities: H'_2 runs over H_1 , or H_2 runs over H'_1 .

Suppose that H'_2 runs over H_1 , but that H_2 does not run over H'_1 . The collapse of H_2 and H_1 does not change the components of intersection between H'_1 and the 2-handles. So, in \mathcal{H} , H'_1 only has H'_2 running over it once and is not incident to any other 2-handles. Hence, H'_1 and H'_2 admit an elementary collapse in \mathcal{H} . The resulting handle structure is completely collapsible, since we may collapse H_2 and H_1 and then follow the remaining collapses that were applied to \mathcal{H} . Hence, by induction, it completely collapses using only elementary collapses. So the same is true of \mathcal{H} .

Now suppose that H_2 runs over H'_1 . The collapse of H_1 and H_2 then introduces new components of intersection between H'_1 and the 2-handles. But in \mathcal{H}' , H'_1 intersects the 2-handles just once. We therefore deduce that in \mathcal{H} , H_1 either is disjoint from the 2-handles other than H_2 or has just a single component of intersection with these handles. In the former case, the collapse of H_1 and H_2 is an elementary collapse, and the proof is complete then. On the other hand, when H_1 intersects the other 2-handles exactly once, this intersection must be with H'_2 . We deduce that H_2 runs over H_1 once and H'_1 once, and that H'_2 runs over H_1 once. Furthermore, H'_1 intersects the 2-handles just once, in H_2 . So we perform an elementary collapse to H_2 and H'_1 . The resulting handle structure is completely collapsible, since we may then collapse H_1 and H'_2 , and then perform the remaining collapses that were applied to $\mathcal{H}' \setminus (H'_1 \cup H'_2)$. Hence, again inductively, \mathcal{H} admits a sequence of elementary collapses taking it to a 0-handle.

Thus, we are reduced to the final case where only $1/0$ collapses are performed. Thus, \mathcal{H} is a thickened tree. In particular, it may collapsed to a single 0-handle by collapsing the leaves of the tree one at time. This is a sequence of elementary collapses. \square

Proposition 18.2. *Let K be a non-split link, and let H be a hierarchy for the exterior of K . Let l be a collection of integral curves on $\partial N(K)$, one on each component of $\partial N(K)$. Let m be a collection of disjoint meridians, with at least one on each component of $\partial N(K)$. We require that these curves respect the handle structure and that each component of m intersects l in a single point. Let \mathcal{H} be the associated handle structure of the 3-sphere. Suppose that, when some 3-handle is removed from \mathcal{H} , the result is completely collapsible. This determines an isotopy of the 3-sphere taking the complement of the 3-handle to a 0-handle of \mathcal{H} . Let K' be the image of K after this isotopy, and suppose that the projection of the 0-handle $D^2 \times I$ onto the first factor gives a diagram D' for K' . Let a be the number of 1-cells in this cell structure, and let b be the maximal number of 1-cells that any 2-cell runs over, counted with multiplicity. Then the constant k in Proposition 17.1 may be taken to be $216a^4b^{3a+3}$.*

Proof. By Proposition 18.1, we may remove the 3-handle from \mathcal{H} and then perform a sequence of elementary collapses to reach a single 0-handle.

When a 2-handle and a 3-handle are collapsed, this has no effect on the link, since the link respects the handle structure. This collapse also does not affect the number of 1-cells that the remaining 2-cells run over. So we ignore these collapses. Each of the remaining collapses removes a 1-cell, and so there are a of these.

After the i th such collapse, each 0-handle intersects K in a tangle, which will project to a union of vertical and horizontal arcs. Let t_i be the total number of these arcs. Initially, exactly $|m|$ of the tangles contain three vertical and horizontal arcs, and the rest are empty. So initially, $t_0 = 3|m| \leq 3a$.

Let m_i be an upper bound for the number of times that the link runs over each 1-handle after the i th collapse. This only increases when a 1-handle and a 2-handle are collapsed. Before such a collapse, the link needs to be slid off the 1-handle. Each such slide may increase the number times it runs over the other 1-handles by at most b . Initially, $m_0 = 1$. Then $m_{i+1} \leq bm_i$. Hence, after i collapses, we have $m_i \leq b^i$.

After all the 3-handles have been removed, we may then perform all the collapses involving a 1-handle and a 2-handle. As discussed above, before the collapse, we first slide at most m_i arcs of K across the 2-handle. Each slide introduces at most b components to the tangles, and each such component is a concatenation of three vertical and horizontal arcs. Hence, $t_{i+1} \leq t_i + 3bm_i \leq t_i + 3b^{i+1}$. So inductively, t_i is at most $3a + 3b^{i+1}$, which is at most $3ab^{i+1}$. Therefore, the arc index is at most $t_i + 2m_iw \leq 3ab^{i+1}w$. Set d_i to be this upper bound $3ab^{i+1}w$.

Consider the case where a 1-handle and a 2-handle are collapsed. Before the collapse, we first slide at most m_i arcs of K across the 2-handle. As explained in the proof of Proposition 17.1, the number of Reidemeister moves performed in each such slide is at most $8d_i^2w + 16w^2d_i \leq 24d_i^2w \leq 216a^2b^{2i+2}w^3$.

Consider now when a 0-handle and a 1-handle are collapsed. We retract the 1-handle, taking the 0-handle with it. In doing so, we pass at most t_i arcs in the projection of the 0-handle past at most d_i other arcs, at most w times. This is achieved using at most t_id_iw exchange moves, each of which requires at most d_i Reidemeister moves. Hence, the number of Reidemeister moves is at most $t_id_i^2w \leq 27a^3b^{3i+3}w^3$. After this, the two 0-handles are merged into a single 0-handle, and the diagrams in these 0-handles are merged into a single diagram. This does not increase t_i or m_i .

We deduce that the total number of Reidemeister moves is at most $216a^4b^{3a+3}w^3$. \square

Remark 18.3. Each 1-cell of the cell structure, other than the 1-cells with interior in $\text{int}(N(K))$, arises where three surfaces of the hierarchy meet or where a surface in the hierarchy meets $\partial N(K)$. At the endpoints of such a 1-cell are points where some S_i meets the pattern P_i . Also taking into account the 1-cells with interior in $\text{int}(N(K))$, we have

$$a = 2|m| + 2 \sum_i |S_i \cap P_i|.$$

Moreover, each 2-cell lying in the hierarchy or $\partial N(K)$ can run over each 1-cell at most once. There are other 2-cells lying in $N(K)$, and these run over the 1-handles at most $\text{length}(l) + 3$ and $\text{length}(m) + 2$ times. Hence,

$$b \leq \max\{a, \text{length}(l) + 3, \text{length}(m) + 2\}.$$

Combining these estimates with Theorem 15.1, Proposition 18.2 and the proof of Theorem 1.1, we obtain the following.

Theorem 18.4. *Let K be a non-split link other than the unknot. Let S_1, \dots, S_ℓ be a hierarchy for the exterior of K as in Proposition 18.2 and that is adequately separating and $(\lambda, 2^{15})$ -exponentially controlled, where*

$$\lambda = \sum_{i=1}^{\ell} -3\chi(S_i) + 6|S_i| + 6|S_i \cap P_i|.$$

Let m and l be curves on $\partial N(K)$ as in Proposition 18.2, and let f be the sum of the absolute values of the integral slopes of l . Suppose that the resulting handle structure of the 3-sphere is totally collapsible once some 3-handle is removed. Let

$$a = 2|m| + \sum_{i=1}^{\ell} 2|S_i \cap P_i|,$$

$$b = \max\{a, \text{length}(l) + 3, \text{length}(m) + 2\}$$

Then we may set

$$p_K(c) = 512a^4b^{3a+3}(10^{80}\lambda^8)^3 \sum_{i=0}^{\ell} 22^i (f + (101/20)c)^{3 \cdot 22^i}.$$

19. THE FIGURE-EIGHT KNOT

Theorem 19.1. *For K the figure-eight knot, we may set $p_K(c) = (10^{108}c)^{15460896}$.*

Proof. The figure-eight knot is shown in Figure 55 (i). We now construct an adequately separating, $(\lambda, 2^{15})$ -exponentially controlled hierarchy for its exterior M , using the procedure given in Section 3. The length ℓ of this hierarchy will be 5. The constants λ and a defined in Theorem 18.4 will be $\lambda = 492$ and $a = b = 128$. The framing f is 4. The resulting handle structure on S^3 will be totally collapsible, once a suitable 3-handle has been removed. Hence, by Theorem 18.4,

$$512a^4b^{3a+3}(10^{80}\lambda^8)^3 \sum_{i=0}^{\ell} 22^i (f + (101c/20))^{3(22)^i}$$

would work as $p_K(c)$. This is bounded above by $(10^{108}c)^{15460896}$, which we may set as $p_K(c)$.

Since M is hyperbolic, we may jump straight to Step (3) in Section 3. Let S_1 be the orientable double cover of the surface shown in Figure 55 (ii). Note that

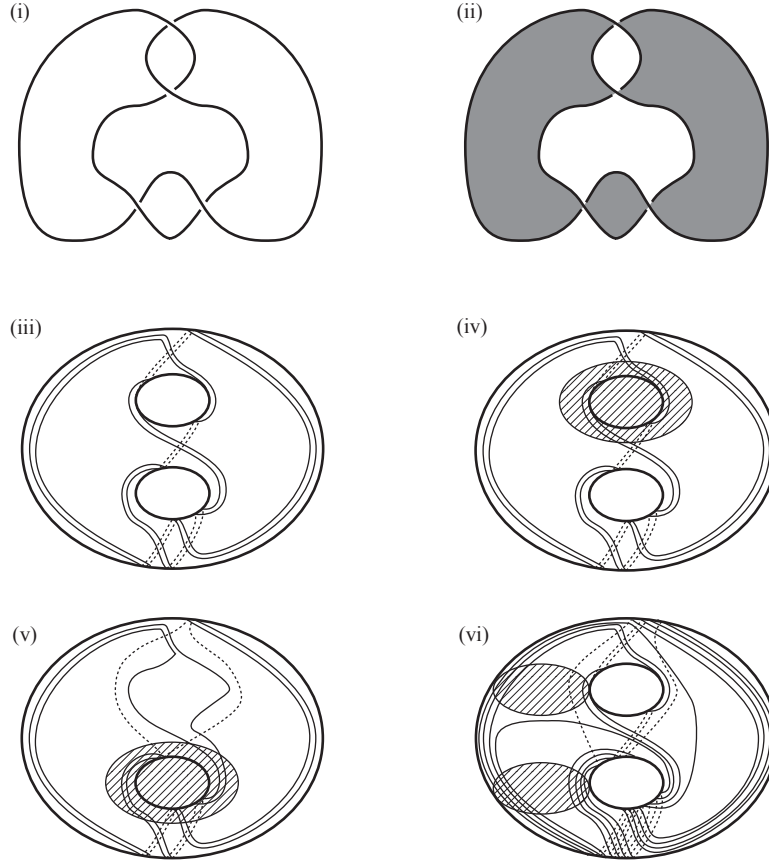


FIGURE 55. Some of the surfaces used to build a weakly fundamental hierarchy for the figure-eight knot exterior

$-\chi(S_1) + 2|S_1| + 2|S_1 \cap P_1| = 4$. It was shown by Hatcher and Thurston [11] that S_1 is incompressible and boundary-incompressible. Note that it is not a fibre in a fibration of M over the circle, since it is separating and connected. Thus, we are permitted to cut along this surface in Step (3). This surface intersects $\partial N(K)$ into two curves, each with integral slope 4.

This creates a manifold M_2 , which has two components. One component, which we denote by M'_2 , consists of the exterior of the surface shown in Figure 55 (ii). The other component, which we denote by M''_2 is a regular neighbourhood of this checkerboard surface. Now M''_2 is an I -bundle chamber, and so it cannot be decomposed at this point. The component M'_2 is shown in Figure 55 (iii). It has a non-trivial JSJ decomposition, as follows. It contains a properly embedded disc that intersects the pattern four times, as shown in Figure 55 (iv). A regular neighbourhood of this disc is an I -bundle over the disc. Attach onto this I -bundle a regular neighbourhood of $\partial M'_2 \cap \partial N(K)$. The result is an I -bundle over a surface, which we denote by M'_3 . One can compute that the base surface is a once-punctured Möbius band. Let M''_3 be $M''_2 \setminus M'_3$. This is shown in Figure 55 (v). Then M'_3 and M''_3 intersect in an annulus S_2 , which we claim is the JSJ annulus for M'_2 . Note

that M_3'' is a solid torus, and its boundary pattern consists of two curves, each of which winds three times around the solid torus. Thus, M_3'' is Seifert fibred, whereas M_3' is an I -bundle, and so these do indeed form the pieces of the JSJ decomposition of M_2' . As required by Step 3 (iii), we let this annulus S_2 be the next surface in the hierarchy.

Since M_3'' is a Seifert fibred solid torus with boundary pattern equal to a union of fibres, it is simple. Hence, Step 4(v) applies to it. We therefore let S_3 be two parallel copies of a meridian disc for this solid torus, each of which can be chosen to intersect the pattern 6 times.

The I -bundle M_2'' has now inherited some boundary pattern in its horizontal boundary. One can check that this contains no clean essential annulus, and hence is simple. Thus, Step 4(v) can be applied to it. We decompose along S_4 , two parallel copies of the discs shown in Figure 55 (vi). Two of these discs intersects the pattern 8 times and two intersects the pattern 6 times.

The only remaining component that is not a 3-ball is M_3' , which is an I -bundle over a once-punctured Möbius band. Again one can check that this contains no clean annuli. We can decompose along S_5 , which is two parallel copies of two vertical discs. Each of these discs each intersects the pattern 4 times. We choose the location of these discs carefully as follows. As observed above, $\partial N(K)$ intersected S_1 in two curves, each with slope 4. These divide $\partial N(K)$ into two annuli. There is a product structure on $N(K)$ as the product of an annulus A and an interval I for which $A \times \partial I$ is a copy of each of these annuli. When further decompositions were made, one of these annuli inherited further boundary pattern. Specifically, when we decompose along S_4 , one of the annuli inherits 8 new arcs of pattern, which come in four parallel pairs. We choose the eight discs S_5 so that the eight arcs $S_5 \cap (A \times \partial I)$ are almost parallel to the existing eight arcs of pattern, via the product structure $A \times I$.

After these decompositions, the resulting manifolds are 3-balls and so the hierarchy terminates.

We may set l to be a curve of slope 4 running along one of the components of ∂S_1 . Its length is 16. The framing f is therefore 4. We pick m to consist of 8 meridians, each running along an arc in ∂S_5 , an arc in ∂S_4 and two short arcs in ∂S_1 . We chose the location of S_5 carefully so that we could set m to be these 8 meridians.

We now briefly explain how to verify the collapsibility hypothesis in Proposition 18.2. We must first pick a 3-handle to be removed. We pick one of 3-handles lying in M_3'' , the one containing the point at infinity in Figure 55.

We will shortly give the sequence of collapses that takes the resulting handle structure to a single 0-handle. But first we present a simplified example. Consider the handle structure of the 3-ball shown in Figure 56. This is a thickening of the following cell structure. Start with a standardly embedded closed orientable genus two surface. Then attach compression discs for its complementary handlebodies. We use two compression discs in each handlebody, as shown in Figure 56. These discs decompose the ‘inside’ handlebody into a 3-ball, which we take to be a 3-cell. The boundaries of these discs decompose the genus two surface into an annulus. We pick an essential arc in this annulus, as shown in Figure 56, and declare that this is a 1-cell. Thus, the surface has now been decomposed to a disc, which we take to be a 2-cell of the cell structure. It is not hard to collapse this cell structure, as

follows. First, collapse the 2-cell which lies in the genus two surface with the inside 3-cell. Then collapse each compression disc along with a 1-cell in its boundary. Then finally retract the resulting tree to a point.

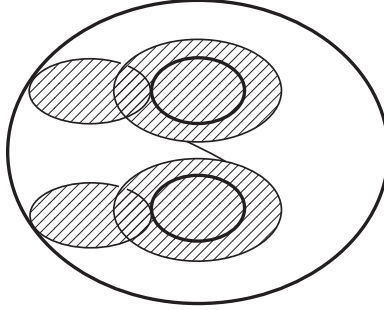


FIGURE 56. A collapsible handle structure of a 3-ball

Our aim is to retract the given handle structure of the 3-ball to the one shown above. We start with the handle structure on $N(K)$. We collapse the eight vertical 1-handles in $N(K) \setminus N_{\text{small}}(K)$ with the eight 0-handles intersecting K , and collapse the 1-handles through which K runs with the 2-handles incident to l . We also collapse the sixteen 1-handles that contain the arcs of intersection between m and ∂S_1 . After these collapses, the handle structure on $N(K)$ is quite simple. It consists of sixteen 0-handles and sixteen 1-handles, which combine to give two thickened circles running along ∂S_1 . It consists of 16 further 1-handles, eight lying in each component of $A \times \partial I$. It also consists of eight 2-handles, each forming a meridian, as well as eight further 2-handles in each component of $A \times \partial I$. Finally, it contains eight 3-handles. We may collapse this handle structure first of all onto one component of $A \times \partial I$, and then onto a single component of ∂S_1 . After this, it just consists of eight 0-handles and eight 1-handles glued in a circular fashion.

We now consider M'_3 , which was an I -bundle that has been decomposed along eight vertical discs. After this decomposition, it has become seven balls, each of which is an I -bundle over a disc. One of these balls contains the disc shown in Figure 55 (iv), and in the boundary of that ball, there are two copies of that disc. We collapse each of the remaining six balls onto one of their components of intersection with $S_1 \cap S_2$. We then collapse the remaining 3-ball onto one of the copies of the disc shown in Figure 55 (iv).

The two discs in S_3 are parallel, with a 3-handle between them. We collapse one of these discs with the 3-handle. Similarly, the discs S_4 come in two pairs, with each pair cobounding a 3-handle. We collapse these two 3-handles with two of the discs.

The handle structure that we have reached is very close to the simple one described above. We have the genus two surface, to which four compression discs as shown in Figure 56 are attached. The genus two surface is subdivided into 2-cells by its intersection with the four discs, but also by other remaining cells, for example the ones running along $N(K)$. But we can collapse these extraneous cells, until we have exactly the handle structure in Figure 56. As described above, this is completely collapsible. \square

20. TORUS KNOTS

Theorem 20.1. *When K is a torus knot, we may set $p_K(c) = (10^{80}c)^{21962}$.*

Proof. Let K be an (r, s) -torus knot with $0 < r < s$. We present an exponentially controlled hierarchy for the exterior of K with length 3. The exterior of K is Seifert fibred and so we give $\partial N(K)$ boundary pattern that is a single longitude. The torus knot K lies on a standardly embedded torus. The intersection between this torus and the exterior of K is an annulus, which we take to S_1 . It intersects the longitude $2rs$ times. By Proposition 2.16, this is weakly fundamental and hence, by Corollary 2.7, it is $(2 + 4rs, 2^{15})$ -exponentially controlled. This decomposes the exterior of K into two solid tori. Pick one of them, and let S_2 be two copies of a meridian disc for it that intersects the pattern as few times as possible. This number is $2r$, say. By Theorem 2.8, it can be made weakly fundamental and hence $(4 + 8r, 2^{15})$ -exponentially controlled. The other solid torus now inherits a boundary pattern, where every complementary region is a disc. A meridian disc intersects this pattern $2s$ times. Let S_3 be two copies of this surface. This is weakly fundamental by Theorem 2.8 and hence is $(2 + 8s, 2^{15})$ -exponentially controlled. For this hierarchy, $\lambda = 6 + 12rs + 24r + 24s$. Thus, the polynomial q provided by Theorem 15.1 is

$$\begin{aligned} q(x) &= (10^{80}(42 + 12rs + 24r + 24s)^8)^{\sum_{i=0}^3 22^i} x^{22^i} \\ &= (10^{80}(6 + 12rs + 24r + 24s)^8)^{11155} x^{10648}. \end{aligned}$$

Thickening this hierarchy gives a handle structure on the exterior of K . We extend this to a handle structure \mathcal{H} on S^3 , using the method described in Section 17.1. This requires a choice of simple closed curves m and l on $\partial N(K)$. For l , we pick a curve lying in $\partial N(S_1)$, which has slope equal to a longitude plus rs meridians. For m , we take it to be some meridional curve that respects the handle structure and that intersects l once.

We will not use the estimate in Proposition 18.2, since this will not be strong enough. Instead, we will construct an isotopy of K through \mathcal{H} explicitly. We first slide K across the 2-handle that is vertical in $N(K) \setminus N_{\text{small}}(K)$ and that runs along l . The 2-handle is a thickening of a 2-cell with binding weight at most $2q(n)$, and hence comprised of at most $8q(n)$ tiles. As we slide over the 2-handle, the arc index of K remains at most $d \leq n + 8q(n) \leq 9q(n)$. Each slide across a tile uses at most $2d^2 + 4wd$ Reidemeister moves, where w is the binding weight of the hierarchy. So, the number of Reidemeister moves is at most $1296(q(n))^3 + 288w(q(n))^2$.

We then slide K across the 1-handle that is vertical in $N(K) \setminus N_{\text{small}}(K)$. This requires at most $2wd$ Reidemeister moves. After this, the knot runs along the 0-handles and 1-handles containing l . It runs along each such handle once.

Now consider one of the discs in S_2 and one of the discs in S_3 . Extend them a little so that their boundaries are on the standard torus containing K . Their boundaries intersect once and hence the union of these boundaries is a wedge W of two circles. We will isotope the knot so that afterwards it lies in a small regular neighbourhood of W . We can think of the standard torus as a square with side identifications, with the sides of the square form W . Currently, the knot lies on this torus as a curve of slope (r, s) . The two curves ∂S_1 , together with ∂S_2 and ∂S_3 divide the torus into discs. We can slide the knots across these discs until it lies in the boundary W of the square. There are $r + s + 1$ of these discs. The knot is divided up into $r + s$ arcs. Each arc of K is slid across each 2-cell at most once.

So we need to perform at most $(r + s + 1)(r + s)$ slides of L across 2-cells to get into a regular neighbourhood of the wedge of circles. It then goes r times round one circle in the wedge and s times round the other. We now slide K across the discs bounded by these circles. We need to perform $r + s$ slides. As described in Section 17.6, each slide across a 2-handle is achieved by at most $24w^3$ Reidemeister moves. Here, w is the binding weight of the hierarchy. After this, the knot lies in a single 0-handle as required. Its projection is in fact the standard diagram D' for the torus knot with crossing number $(r - 1)s$.

Thus, as argued in the proof of Theorem 1.1, the number of Reidemeister moves needed to convert an arbitrary diagram D to a standard diagram D' is at most

$$rs + c + (rs + (101/20)c)^4 q(rs + (101/20)c) + k(q(rs + (101/20)c))^3,$$

where $k = 1602 + 24(r + s)(r + s + 2)$. We can now use the fact that the crossing number of the torus knot is $(r - 1)s$ and so $c \geq (r - 1)s$. So $c \geq rs/2$ and $c \geq r + s + 2$. So, $k \leq 24c^2$. Thus, the number of Reidemeister moves is at most

$$\begin{aligned} & rs + c + (rs + (101/20)c)^4 q(rs + (101/20)c) + k(q(rs + (101/20)c))^3 \\ & \leq 3c + (141/20)^4 c^4 (10^{80} (6 + 12rs + 24r + 24s)^8)^{11155} (141c/20)^{10648} \\ & \quad + (1602 + 24c^2) (10^{80} (6 + 12rs + 24r + 24s)^8)^{33465} (141c/20)^{31944} \\ & \leq (10^{11} c)^{299666}. \end{aligned}$$

So we may set $p_K(c) = (10^{11} c)^{299666}$, as claimed. \square

REFERENCES

- [1] John A. Baldwin and Steven Sivek. On the complexity of torus knot recognition. *Trans. Amer. Math. Soc.*, 371(6):3831–3855, 2019.
- [2] Daniel Bennequin. Entrelacements et équations de Pfaff. In *Third Schnepfenried geometry conference, Vol. 1 (Schnepfenried, 1982)*, volume 107-108 of *Astérisque*, pages 87–161. Soc. Math. France, Paris, 1983.
- [3] Joan S. Birman and William W. Menasco. Studying links via closed braids. IV. Composite links and split links. *Invent. Math.*, 102(1):115–139, 1990.
- [4] Ryan Budney. JSJ-decompositions of knot and link complements in S^3 . *Enseign. Math. (2)*, 52(3-4):319–359, 2006.
- [5] Alexander Coward and Marc Lackenby. An upper bound on Reidemeister moves. *Amer. J. Math.*, 136(4):1023–1066, 2014.
- [6] Peter R. Cromwell. Embedding knots and links in an open book. I. Basic properties. *Topology Appl.*, 64(1):37–58, 1995.
- [7] M. Culler and P. B. Shalen. Bounded, separating, incompressible surfaces in knot manifolds. *Invent. Math.*, 75(3):537–545, 1984.
- [8] I. A. Dynnikov. Arc-presentations of links: monotonic simplification. *Fund. Math.*, 190:29–76, 2006.
- [9] Wolfgang Haken. Some results on surfaces in 3-manifolds. In *Studies in Modern Topology*, volume Vol. 5 of *Studies in Mathematics*, pages 39–98. Math. Assoc. America., 1968.
- [10] Joel Hass and Jeffrey C. Lagarias. The number of Reidemeister moves needed for unknotting. *J. Amer. Math. Soc.*, 14(2):399–428, 2001.
- [11] A. Hatcher and W. Thurston. Incompressible surfaces in 2-bridge knot complements. *Invent. Math.*, 79(2):225–246, 1985.
- [12] Geoffrey Hemion. On the classification of homeomorphisms of 2-manifolds and the classification of 3-manifolds. *Acta Math.*, 142(1-2):123–155, 1979.
- [13] William Jaco and J. Hyam Rubinstein. PL equivariant surgery and invariant decompositions of 3-manifolds. *Adv. in Math.*, 73(2):149–191, 1989.
- [14] Klaus Johannson. *Homotopy equivalences of 3-manifolds with boundaries*, volume 761 of *Lecture Notes in Mathematics*. Springer, Berlin, 1979.

- [15] Greg Kuperberg. Algorithmic homeomorphism of 3-manifolds as a corollary of geometrization. *Pacific J. Math.*, 301(1):189–241, 2019.
- [16] Marc Lackenby. The crossing number of composite knots. *J. Topol.*, 2(4):747–768, 2009.
- [17] Marc Lackenby. A polynomial upper bound on Reidemeister moves. *Ann. of Math. (2)*, 182(2):491–564, 2015.
- [18] Sergei Matveev. *Algorithmic topology and classification of 3-manifolds*, volume 9 of *Algorithms and Computation in Mathematics*. Springer, Berlin, second edition, 2007.
- [19] Kunio Murasugi. On the braid index of alternating links. *Trans. Amer. Math. Soc.*, 326(1):237–260, 1991.
- [20] Friedhelm Waldhausen. Eine Klasse von 3-dimensionalen Mannigfaltigkeiten. I, II. *Invent. Math.*, 3:308–333; *ibid.* 4 (1967), 87–117, 1967.
- [21] Friedhelm Waldhausen. Gruppen mit Zentrum und 3-dimensionale Mannigfaltigkeiten. *Topology*, 6:505–517, 1967.
- [22] Friedhelm Waldhausen. On irreducible 3-manifolds which are sufficiently large. *Ann. of Math. (2)*, 87:56–88, 1968.

MATHEMATICAL INSTITUTE, UNIVERSITY OF OXFORD,
WOODSTOCK ROAD, OXFORD OX2 6GG, UNITED KINGDOM