Algebraic & Geometric Topology 7 (2007) 957–1006

Multiple bridge surfaces restrict knot distance

MAGGY TOMOVA

Suppose M is a closed irreducible orientable 3–manifold, K is a knot in M, P and Q are bridge surfaces for K and K is not removable with respect to Q. We show that either Q is equivalent to P or $d(K, P) \leq 2 - \chi(Q - K)$. If K is not a 2–bridge knot, then the result holds even if K is removable with respect to Q. As a corollary we show that if a knot in S^3 has high distance with respect to some bridge sphere and low bridge number, then the knot has a unique minimal bridge position.

57M25, 57M27, 57M50

1 Introduction and preliminaries

Distance is a generalization of the concept of weak and strong compressibility for bicompressible surfaces originally due to Hempel [5]. It has been successfully applied to study Heegaard splittings of 3–manifolds. For example in [4] Hartshorn shows that the Euler characteristic of an essential surface in a manifold bounds the distance of any of its Heegaard splittings. In [12] Scharlemann and Tomova show that the Euler characteristic of any Heegaard splitting of a 3–manifold similarly bounds the distance of any non-isotopic splitting.

A knot K in a 3–manifold M is said to be in bridge position with respect to a surface P if P is a Heegaard surface for M and K intersects each of the components of $M - P$ in arcs that are simultaneously parallel to P. If K is in bridge position with respect to P, we say that P is a bridge surface for K. The definition of distance has been extended to apply to bridge surfaces. In [2], Bachman and Schleimer prove that Hartshorn's result extends to the distance of a bridge surface, namely the Euler characteristic of an essential properly embedded surface in the complement of a knot bounds the distance of any bridge surface for the knot. In this paper we extend the ideas in [12] to show that the result there also extends to the case of a knot with two different bridge surfaces.

Corollary *Suppose K is a non-trivial knot in a closed, irreducible and orientable 3–manifold M and P is a bridge surface for K that is not a 4–times punctured sphere. If Q is also a bridge surface for K that is not equivalent to P, or if Q is a Heegaard surface for $M - \eta(K)$ then $d(K, P) \leq 2 - \chi(Q - K)$.*

Published: 20 June 2007 DOI: 10.2140/agt.2007.7.957

In this paper two Heegaard splittings are considered to be equivalent if one is isotopic to a possibly stabilized copy of the other. For bridge surfaces there are three obvious geometric operations that correspond to stabilizations and they are described in Section 10.

A knot K is said to be removable with respect to a bridge surface Q if K can be isotoped to lie in the spine of one of the handlebodies $M - Q$. Thus after the isotopy, Q is a Heegaard surface for $M - \eta(K)$. If we restrict our attention only to bridge surfaces with respect to which the knot is not removable, we may extend the above theorem also to 2–bridge knots.

Corollary *Suppose P and Q are two bridge surfaces for a knot K and K is not removable with respect to Q. Then either Q is equivalent to P or $d(P) \leq 2 - \chi(Q_K)$.*

The result proves a conjecture of Bachman and Schleimer put forth in [2].

Corollary *If $K \subset S^3$ is in minimal bridge position with respect to a sphere P such that $d(K, P) > |P \cap K|$ then K has a unique minimal bridge position.*

The basic idea of the proof of the above theorem is to consider a 2–parameter sweep-out of $M - K$ by the two bridge surfaces. We keep track of information about compressions by introducing labels for the regions of the graphic associated to the sweep-out. We are able to conclude that if particular combinations of labels occur we can deduce the desired result. Using a quadrilateral version of Sperner's lemma, we conclude that one of the label combinations we have already considered must occur.

2 Surfaces in a handlebody intersected by the knot in unknotted arcs

Throughout this paper we will use the following definitions and notation.

Notation 2.1 *Let M be a compact orientable irreducible 3–manifold. If $K \subset M$ is some properly embedded 1–manifold, let M_K denote M with a regular (open) neighborhood $N(K)$ of K removed. If X is any subset of M, let $X_K = M_K \cap X$.*

Definition 2.2 *Suppose $(F, \partial F) \subset (M, \partial M)$ is a properly embedded surface in a compact orientable irreducible manifold M containing a 1–manifold K such that F is transverse to K.*

- *We will say that F_K is* n–times punctured *if $|F \cap K| = n$. If F_K is 1–time punctured, we will call it* punctured.

- *A simple closed curve in F_K is* inessential *if it bounds a subdisk of F_K or it is parallel to a component of ∂F_K. Otherwise the curve is* essential.
- *A properly embedded arc $(\beta, \partial\beta) \subset (F_K, \partial F_K)$ is* essential *if no component of $F_K - \beta$ is a disk.*
- *A properly embedded disk $(D, \partial D) \subset (M_K, F_K)$ is a* compressing disk *for F_K in M_K if ∂D is an essential curve in F_K.*
- *A disk D^c in M is a* cut-disk *for F_K if $D^c \cap F_K = \partial D^c$, ∂D^c is essential in F_K and D^c intersects K in a single point. Thus $D^c \cap M_K$ is an annulus with one boundary component an essential curve in F_K and the other one a meridional curve for the torus boundary component of M_K.*
- *A* c–disk *D^* for F_K is either a cut-disk or a compressing disk.*
- *A surface F_K is called* incompressible *if it has no compressing disks,* cut-incompressible *if it has no cut-disks and* c–incompressible *if it has no c–disks.*
- *A surface F_K is called* essential *if it is incompressible and at least one of its components is not parallel to ∂M_K.*

Now we restrict our attention to the case when the 3–manifold we are considering is a handlebody and the 1–manifold K consists of "unknotted" properly embedded arcs. To make this more precise we use the following definition modeled after the definition of a K–compression body introduced by Bachman in [1].

Definition 2.3 *A K–handlebody, (A, K) is a handlebody A and a 1–manifold, $(K, \partial K) \subset (A, \partial A)$, such that K is a disjoint union of properly embedded arcs and for each arc $\kappa \in K$ there is a disk, $D \subset A$ with $\partial D = \kappa \cup \alpha$, where $D \cap K = \kappa$ and $D \cap \partial A = \alpha$. These disks are called bridge disks and the arcs are called bridges.*

A spine of a handlebody A is a properly embedded finite graph Σ_A in A (typically chosen to have no valence 1 vertices) so that $A - \Sigma_A \cong \partial A \times [0, 1)$. Given a spine Σ_A and a collection K of bridges in A, K can be isotoped in A (for example by shrinking a collection E of bridge disks very close to ∂A) so that the projection (called the height) $A - \Sigma_A \cong \partial A \times [0, 1) \to [0, 1)$ has a single maximum on each bridge α_i. For each α_i connect Σ_A to that maximum by an arc in A which is monotonic with respect to height. The union of Σ_A with that collection of arcs is called a spine $\Sigma_{(A,K)}$ of (A, K). Note that there is a homeomorphism $A - \Sigma_{(A,K)} \cong \partial A \times [0, 1)$ which carries $K - \Sigma_{(A,K)}$ to $(\partial A \cap K) \times [0, 1)$. Put another way, there is a map $(\partial A, \partial A \cap K) \times I \to (A, K)$ which is a homeomorphism except over $\Sigma_{(A,K)}$, and the map gives a neighborhood of $\Sigma_{(A,K)}$ a mapping cylinder structure.

Notation 2.4 *For the rest of this paper, unless otherwise specified, let (A, K) be a K–handlebody with $P = \partial A$ and spine $\Sigma_{(A,K)}$. We will always assume that if A is a ball, then K has at least 3 components. The surface $F \subset A$ will be properly embedded and transverse to K. We continue to denote by $N(K)$ a regular neighborhood of K.*

Definition 2.5 *Two embedded meridional surfaces S and T in (M, K) are called K–parallel if they cobound a region homeomorphic to $S_K \times I$ ie the region of parallelism contains only unknotted segments of K each with one endpoint in S and one endpoint in T.*

Two meridional surfaces S and T are K–isotopic if there exists an isotopy from S to T so that S remains transverse to K throughout the isotopy.

Lemma 2.6 *If $(E, \partial E) \subset (A_K, P_K)$ is a possibly punctured disk such that ∂E is an inessential curve in P_K, then E is parallel to a possibly punctured subdisk of P_K.*

Proof Let E' be the possibly punctured disk ∂E bounds in P_K. There are three cases to consider. If E and E' are both disks, then they cobound a ball as A_K is irreducible, and thus E is parallel to E'. If one of E and E' is a once punctured disk and the other one is a disk, then the sphere $E \cup E'$ intersects K only once. The manifold is irreducible and $E \cup E'$ is separating so this is not possible. Finally, if both E and E' are once punctured disks, then by irreducibility of A and the definition of a K–handlebody, E and E' cobound a product region in A_K. This product region intersects some bridge disk for K in a single arc, so the arc of K between E and E' is a product arc. It follows that E and E' are parallel as punctured disks. □

Definition 2.7 *A P–compressing disk for $F_K \subset A_K$ is a disk $D \subset A_K$ so that ∂D is the end-point union of two arcs, $\alpha = D \cap P_K$ and $\beta = D \cap F_K$, and β is an essential arc in F_K.*

The operation of compressing, cut-compressing and P–compressing the surface F_K have natural duals that we will refer to as *tubing* (possibly tubing along a subset of the knot) and *tunneling* along an arc dual to the c–disk or the P–compressing disk. The precise definitions of these operations were given by Scharlemann in [8]. Suppose $F \subset M$ is a properly embedded surface in a manifold containing a knot K. Let $\gamma \subset \text{interior}(M)$ be an embedded arc such that $\gamma \cap F = \partial\gamma$. There is a relative tubular neighborhood $\eta(\gamma) \cong \gamma \times D^2$ so that $\eta(\gamma)$ intersects F precisely in the two diskfibers at the ends of γ. Then the surface obtained from F by removing these two disks and attaching the cylinder $\gamma \times \partial D^2$ is said to be obtained by tubing along γ. We allow for the possibility that $\gamma \subset K$. Similarly if $\gamma \subset \partial M$, there is a relative neighborhood

$\eta(\gamma) \cong \gamma \times D^2$ so that $\eta(\gamma)$ intersects F precisely in the two disk fibers at the ends of γ and $\eta(\gamma)$ intersects ∂M in a rectangle. Then the surface obtained from F by removing the two half disks and attaching the rectangle $(\gamma \times \partial D^2) \cap M$ is said to be obtained by tunnelling along γ.

We will have many occasions to use P–compressions of surfaces so we note the following lemma.

Lemma 2.8 *Suppose $F_K \subset A_K$ is a properly embedded surface and F'_K is the result of P–compressing F_K along a P–compressing disk E_0. Then*

1. *if F'_K has a c–disk, F_K also has a c–disk of the same kind (cut or compressing),*
2. *if F_K intersects every spine $\Sigma_{(A,K)}$ then so does F'_K and*
3. *every curve of ∂F_K can be isotoped in P_K to be disjoint from any curve in $\partial F'_K$.*

Proof The original surface F_K can be recovered from F'_K by tunneling along an arc that is dual to the P–compressing disk. This operation is performed in a small neighborhood of P_K so if F'_K has compressing or cut-disks, they will be preserved in F_K. Also if F'_K is disjoint from some $\Sigma_{(A,K)}$, then adding a tunnel close to P_K will not introduce any intersections with this spine. For the last item consider the frontier of $N(F_K \cup E_0) \cap P_K$ where N denotes a regular neighborhood. This set of disjoint embedded curves in P_K contains both $F_K \cap P_K$ and $F'_K \cap P_K$. □

In the case of a handlebody it is also known that any essential surface must have boundary. The following lemma proves the corresponding result for a K–handlebody.

Lemma 2.9 *If F_K is an incompressible surface in A_K, then one of the following holds,*

1. *F_K is a sphere,*
2. *F_K is a twice punctured sphere, or*
3. *$F_K \cap P_K \neq \varnothing$.*

Proof Suppose F_K is an incompressible surface in A_K that is not a sphere or a twice-punctured sphere, such that $P_K \cap F_K = \varnothing$. Let Δ be the collection of a complete set of compressing disks for the handlebody A together with all bridge disks for K. Via an innermost disk argument, using the fact that F_K is incompressible, we may assume that $F_K \cap \Delta$ contains only arcs. Any arc of intersection between a disk $D \in \Delta$ and F_K must have both of its endpoints lying in $N(K)$ as $F_K \cap P_K = \varnothing$ and thus lies in one of the bridge disks. Consider an outermost such arc in D cutting a subdisk E of D. Doubling E along K produces a compressing disk for F_K which was assumed to be incompressible. Thus F_K must be disjoint from Δ and therefore F_K lies in the ball $A_K - \Delta$ contradicting the incompressibility of F_K. □

Finally it is well known that if F is a closed connected incompressible surface contained in $A - \Sigma_A \cong P \times I$, then F is isotopic to P. A similar result holds if we consider $F_K \subset (A_K - \Sigma_{(A,K)}) = P_K \times I$.

Lemma 2.10 *Suppose P is a closed connected surface, and $K \neq \varnothing$ is a 1–manifold properly embedded in $P \times I$ so that each component of K can be isotoped to be vertical with respect to the product structure. If $F_K \subset P_K \times I$ is a properly embedded connected incompressible surface such that $F_K \cap (P \times \{0\}) = F_K \cap (P \times \{1\}) = \varnothing$, then one of the following holds,*

(1) *F_K is a sphere disjoint from the knot,*

(2) *F_K is a twice punctured sphere, or*

(3) *F_K is K–isotopic to $P_K \times \{0\}$.*

Proof Suppose F_K is not a sphere or a twice punctured sphere. Consider the set S consisting of properly embedded arcs in P_K so that $P_K - S$ is a disk. This collection gives rise to a collection $\Delta = S \times I$ of disks in $P_K \times I$ so that $(P_K \times I) - \Delta$ is a ball. As F_K is not a sphere $F_K \cap \Delta \neq \varnothing$. As F_K is incompressible, by an innermost disk argument we may assume that it does not intersect Δ in any closed curves. If $F_K \cap \Delta$ contains an arc that has both of its endpoints in the same component of K, doubling the subdisk of Δ bounded by an outermost such arc would give a compressing disk for F_K. Consider the components of F_K lying in the ball $(P_K \times I) - \Delta$. As F_K is incompressible all of these components must be disks. In fact, as F_K is connected, there is a single disk component. This disk is isotopic to $(P_K - S) \times 0$ and the maps that glue $(P_K \times I) - \Delta$ to recover $P_K \times I$ do not affect the isotopy. □

3 The curve complex and distance of a knot

Suppose V is a compact, orientable, properly embedded surface in a 3–manifold M. The *curve complex* of V is a graph $\mathcal{C}(V)$, with vertices corresponding to isotopy classes of essential simple closed curves in V. Two vertices are adjacent if their corresponding isotopy classes of curves have disjoint representatives. If S and T are subsets of vertices of $\mathcal{C}(V)$, then $d(S, T)$ is the length of the shortest path in the graph connecting a vertex in S and a vertex in T.

Definition 3.1 *Let $(P, \partial P) \subset (M, \partial M)$ be a properly embedded surface in an orientable irreducible 3–manifold M. The surface P will be called a splitting surface if M is the union of two manifolds A and B along P. We will say P splits M into A and B. If P splits M into A and B and is compressible in both A and B, then P is bicompressible.*

If P is a closed embedded bicompressible surface with $\chi(P) < 0$ splitting M into submanifolds A and B, let $\mathcal{A}$ (resp $\mathcal{B}$) be the set of all simple closed curves in P that bound compressing disks for P in A (resp B). Then $d(P) = d(\mathcal{A}, \mathcal{B})$ ie, the length of the shortest path in the graph $\mathcal{C}(P)$ between a curve in $\mathcal{A}$ and a curve in $\mathcal{B}$. If $d(P) \leq 1$, ie there are compressing disks on opposite sides of P with disjoint boundaries, then the surface P is called *strongly compressible* in M. Otherwise P is *weakly incompressible*.

Much like bridge number and width, the distance of a knot measures its complexity. It was first introduced by Bachman and Schleimer in [2]. The definition we use in this paper is slightly different and corresponds more closely to the definition of the distance of a surface.

Definition 3.2 *Suppose M is a closed, orientable irreducible 3–manifold containing a knot K and suppose P is a bridge surface for K splitting M into handlebodies A and B. The curve complex $\mathcal{C}(P_K)$ is a graph with vertices corresponding to isotopy classes of essential simple closed curves in P_K. Two vertices are adjacent in $\mathcal{C}(P_K)$ if their corresponding classes of curves have disjoint representatives. Let $\mathcal{A}$ (resp $\mathcal{B}$) be the set of all essential simple closed curves in P_K that bound disks in A_K (resp B_K). Then $d(P, K) = d(\mathcal{A}, \mathcal{B})$ measured in $\mathcal{C}(P_K)$.*

The curve complex for a non-punctured torus and a 4 punctured sphere are not connected. However 2 bridge knots in S^3 cannot have multiple bridge surfaces, Scharlemann and Tomova [11], so these cases don't arise in our context.

4 Bounds on distance given by an incompressible surface

We will continue to assume that (A, K) is a K–handlebody, $P = \partial A$ and if A is a ball, then K has at least 3 components. For clarity we will refer to a properly embedded surface $E_K \subset A_K$ with zero Euler characteristic as an annulus only if it has 2 boundary components both lying in P_K and distinguish it from a punctured disk, a surface with one boundary component lying in P_K that intersects $cl(N(K))$ in a single meridional circle. Consider the curve complex $\mathcal{C}(P_K)$ of P_K and let $\mathcal{A}$ be the set of all essential curves in P_K that bound disks in A_K.

Proposition 4.1 *Suppose D^c is a cut-disk for P_K in A_K. Then there is a compressing disk D for P_K such that $d(\partial D^c, \partial D) \leq 1$.*

Proof Let κ be the arc of K that punctures D^c and B be its bridge disk. After perhaps an isotopy of B, $B \cap D^c$ is a single arc α that separates B into two subdisks B_1 and B_2. Consider a regular neighborhood of $D^c \cup B_1$ say. Its boundary contains a disk that intersects P_K in an essential curve and does not intersect ∂D^c as required. □

Proposition 4.2 *Consider $(F, \partial F) \subset (A, P)$, a properly embedded surface transverse to K.*

- *If the surface F_K contains a disk component whose boundary is essential in P_K, then $d(\mathcal{A}, f) \leq 1$ for every $f \in F_K \cap P_K$ that is essential in P_K.*
- *If F_K has a punctured disk component D^c whose boundary is essential in P_K, then $d(\partial D^c, \mathcal{A}) \leq 1$.*

Proof If F_K contains such a disk component D, then D is necessarily a compressing disk for P_K so $\partial D \in \mathcal{A}$ and $\partial D \cap f = \varnothing$ for every $f \in F_K \cap P_K$ as F_K is embedded thus $d(\mathcal{A}, f) \leq 1$.

The second claim follows immediately from Proposition 4.1. □

Proposition 4.3 *Consider $(F, \partial F) \subset (A, P)$, a properly embedded surface transverse to K and suppose it satisfies all of the following conditions:*

(1) *F_K has no disk components,*

(2) *F_K is c–incompressible,*

(3) *F_K intersects every spine $\Sigma_{(A,K)}$ and*

(4) *all curves of $F_K \cap P_K$ are either essential in P_K or bound punctured disks on both surfaces.*

Then there is at least one curve $f \in F_K \cap P_K$ that is essential in P_K and such that $d(\mathcal{A}, f) \leq 1 - \chi(F_K)$. Every $f \in F_K \cap P_K$ that is essential in P_K for which the inequality does not hold lies in the boundary of a P_K–parallel annulus component of F_K.

Proof If F_K is a counterexample to the proposition, the surface F_K^- obtained from F_K by deleting all P_K–parallel annuli and P_K–parallel punctured disk components would also be a counterexample with the same euler characteristic. Note that F_K^- is nonempty as otherwise F_K would be disjoint from a spine $\Sigma_{(A,K)}$ and is c–incompressible as any c–disk would also be a c–disk for F_K. Thus we assume F_K does no have any P_K–parallel annuli or punctured disk components.

Let E be a compressing disk for P_K in A_K (not punctured by the knot) so that $|E \cap F_K|$ is minimal among all such disks. If in fact $E \cap F_K = \emptyset$, then $d(\partial E, f) \leq 1$ for every $f \in \partial F_K$ as required so we may assume $E \cap F_K \neq \emptyset$. Circles of intersection between F_K and E and arcs that are inessential in F_K can be removed by innermost disk and outermost arc arguments. Thus we can assume F_K and E only intersect in arcs that are essential in F_K.

The proof now is by induction on $1 - \chi(F_K)$. As F_K has no disk components for the base case of the induction assume $1 - \chi(F_K) = 1$, ie all components of F_K are annuli or once punctured disks and no component is P_K–parallel. If E intersects a punctured disk component of F_K the arc of intersection would necessarily be inessential in F_K contradicting the minimality of $|F_K \cap E|$ so we may assume that if $F_K \cap E \neq \emptyset$, E only intersects annulus components of F_K. An outermost arc of intersection in E bounds a P–compressing disk E_0 for F_K. After the P–compression, the new surface F'_K contains a compressing disk D for P_K, the result of a P–compression of an essential annulus, and ∂D is disjoint from all $f \in \partial F_K$ by Lemma 2.8. As $\partial D \in \mathcal{A}$, $d(f, \mathcal{A}) \leq 1 = 1 - \chi(F_K)$ for every $f \in F_K \cap P_K$ as desired.

Now suppose $1 - \chi(F_K) > 1$. Again let E_0 be a subdisk of E cut off by an outermost arc of $E \cap F_K$ and F'_K be the surface obtained after the P–compression. By Lemma 2.8 F'_K also intersect every spine $\Sigma_{(A,K)}$ and is c–incompressible. By the definition of P–compression, F'_K cannot have any disk components as F_K did not have any. Thus F'_K satisfies the first 3 conditions of the proposition. There are two cases to consider.

Case 1 Any simple closed curves in $F'_K \cap P_K$ that are inessential in P_K bound punctured disks in both surface.

In this case F'_K satisfies all the hypothesis of the proposition so we can apply the induction hypothesis. Thus there exists a curve $f' \in F'_K \cap P_K$ that satisfies the distance inequality. Since, by Lemma 2.8, for every component f of $F_K \cap P_K$, $d(f, f') \leq 1$, we have the inequality $d(f, \mathcal{A}) \leq d(f', \mathcal{A}) + d(f, f') \leq 1 - \chi(F'_K) + 1 = 1 - \chi(F_K)$, as desired.

Case 2 Some curve of $F'_K \cap P_K$ is inessential in P_K but does not bound a punctured disk in F'_K.

Let c be this curve and let E^* be the possibly punctured disk c bounds in P_K. By our hypothesis, the tunnel dual to the P–compression must be adjacent to c as otherwise c would persist in $F_K \cap P_K$. Push a copy of E^* slightly into A_K. After the tunneling, E^* is no longer parallel to P_K. As F_K was assumed to be c–incompressible, $c = \partial E^*$

must be parallel to some component of ∂F_K. As c didn't bound a punctured disk in F'_K, ∂E^* must be parallel to some component $\tilde{c} \in F_K \cap P_K$ that is essential in P_K by hypothesis. Use this parallelism to extend E^* to a c–disk for P_K with boundary $\tilde{c}$, see Figure 1. Now for every $f \in F_K \cap P_K$, by Proposition 4.2 we have that $d(f, \mathcal{A}) \le d(f, \partial E^*) + d(\partial E^*, \mathcal{A}) \le 1 + 1 = 2 \le 1 - \chi(F_K)$. □

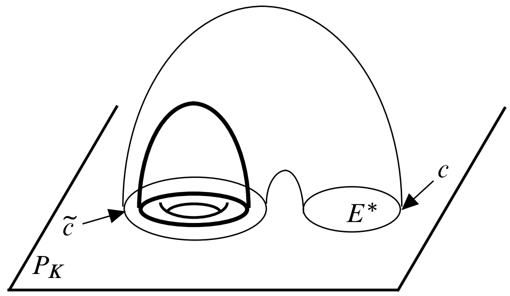


Figure 1

5 The genus of an essential surface bounds the distance of a knot

Notation 5.1 *For the rest of the paper we will assume that M is a closed irreducible orientable 3–manifold containing a knot K and P is a bridge surface for K such that $M = A \cup_P B$. Furthermore we assume that if P is a sphere, then P_K has at least 6 punctures.*

Let $Q \subset M$ be a properly embedded surface that is transverse to K. We will consider how the surfaces P_K and Q_K can intersect in M_K to obtain bounds on $d(P, K)$.

We import the next lemma directly from [12].

Lemma 5.2 *Let $Q \subset M$ be a properly embedded surface that is transverse to K and let $Q^A_K = Q_K \cap A_K$, $Q^B_K = Q_K \cap B_K$. Suppose Q_K satisfies the following conditions.*

- *All curves of $P_K \cap Q_K$ are essential in P_K and don't bound disks in Q_K.*
- *There is at least one curve $a \in Q^A_K \cap P_K$ such that $d(a, \mathcal{A}) \le 1 - \chi(Q^A_K)$ and any curve in $Q^A_K \cap P_K$ for which the inequality does not hold is the boundary of an annulus component of Q^A_K that is parallel into P_K.*

- *There is at least one curve $b \in Q_K^B \cap P_K$ such that $d(b, \mathcal{B}) \leq 1 - \chi(Q_K^B)$ and any curve in $Q_K^B \cap P_K$ for which the inequality does not hold is the boundary of an annulus component of Q_K^B that is parallel into P_K.*

Then $d(K, P) \leq 2 - \chi(Q_K)$.

Proof Call a component c of $P_K \cap Q_K$ *A–conforming* (resp *B–conforming*) if $d(c, \mathcal{A}) \leq 1 - \chi(Q_K^A)$ (resp $d(c, \mathcal{B}) \leq 1 - \chi(Q_K^B)$). By hypothesis there are both A–conforming components of $Q_K \cap P_K$ and B–conforming components. If there is a component c that is both A–conforming and B–conforming, then

$$d(K, P) = d(\mathcal{A}, \mathcal{B}) \leq d(\mathcal{A}, c) + d(c, \mathcal{B}) \leq 2 - \chi(Q_K^A) - \chi(Q_K^B) = 2 - \chi(Q_K)$$

as required.

If there is no such component, let γ be a path in Q_K from an A–conforming component to a B–conforming component, chosen to intersect P_K as few times as possible. In particular, any component of $P_K \cap Q_K$ incident to the interior of γ is neither A–conforming nor B–conforming, so each of these components of Q_K^A and Q_K^B is an annulus, parallel to an annulus in P_K. It follows that the components of $P_K \cap Q_K$ at the ends of γ are isotopic in P_K. Letting c be a simple closed curve in that isotopy class in P_K we have as above

$$d(K, P) = d(\mathcal{A}, \mathcal{B}) \leq d(\mathcal{A}, c) + d(c, \mathcal{B}) \leq 2 - \chi(Q_K^A) - \chi(Q_K^B) = 2 - \chi(Q_K)$$

as required. □

Corollary 5.3 *Suppose $Q_K \subset M_K$ is a properly embedded connected surface transverse to P_K so that all curves of $P_K \cap Q_K$ are essential in both surfaces. If Q_K^A and Q_K^B are c–incompressible and intersect every spine $\Sigma_{(A,K)}$ and $\Sigma_{(B,K)}$ respectively, then $d(K, P) \leq 2 - \chi(Q_K)$.*

Proof Proposition 4.3 shows that Q_K^A and Q_K^B satisfy respectively the second and third conditions of Lemma 5.2. □

The following definition was first used by Scharlemann and Tomova in [12].

Definition 5.4 *Suppose S and T are two properly embedded surfaces in a 3–manifold M containing a knot K and assume S and T intersect the knot transversely. Let $c \in S_K \cap T_K$ be a simple closed curve bounding possibly punctured disks $D \subset S_K$ and $E \subset T_K$. If D intersects T_K only in curves that are inessential in T_K and E intersects S_K only in curves that are inessential in S_K we say that c is* removable.

The term reflects the fact that all such curves can be removed by isotopies of S_K whose support lies away from any curves of intersection that are essential either in S_K or in T_K. Indeed, if c is removable, then any component of $D \cap E$ is clearly also removable.

The following definition was introduced by Bachman and Schleimer in [2].

Definition 5.5 *Suppose S and T are two properly embedded surfaces in a 3–manifold M. A simple closed curve $\alpha \in S \cap T$ is mutually essential if it is essential in both surfaces, it is mutually inessential if it is inessential in both surfaces and it is mutual if it is either mutually essential or mutually inessential.*

The following remark follows directly from the above two definitions.

Remark 5.6 *If every curve of intersection between S_K and T_K is mutual, then all inessential curves of $S_K \cap T_K$ are removable.*

Now we can recover the bound on distance obtained in [2] but using our definition of distance. Note that we only require the surface Q_K to have no compressing disks but allow it to have cut-disks.

Theorem 5.7 *Let M be a closed irreducible orientable manifold containing a knot K and let P be a bridge surface for K such that if P is a sphere, P_K has at least 6 punctures. Suppose $Q \subset M$ is a properly embedded essential (in M_K) meridional surface such that Q_K is neither a sphere nor an annulus. Then $d(K, P) \leq 2 - \chi(Q_K)$. If Q_K is an essential annulus, then $d(K, P) \leq 3$.*

Proof If Q_K has any cut-disks, cut-compress along them, ie if D^c is a cut-disk for Q_K, remove a neighborhood of ∂D^c from Q_K and then add two copies of D^c along the two newly created boundary components. Repeat this process until the resulting surface has no c–disks. Let Q'_K be the resulting surface and notice that $\chi(Q_K) = \chi(Q'_K)$. Suppose Q'_K has a compressing disk D. The original surface Q_K can be recovered from Q'_K by tubing along a collection of subarcs of K. Note that as $D \cap K = \varnothing$ none of these tubes can intersect D. Thus D is also a compressing disk for Q_K contrary to the hypothesis so Q'_K is also incompressible. Finally note that in this process no sphere, annulus or torus components are produced so at least one of the resulting components is not a sphere, annulus or torus, in particular Q'_K has at least one component that is not parallel to ∂M_K. By possibly replacing Q_K by Q'_K we may assume that Q_K is also cut-incompressible.

Recall that $\Sigma_{(A,K)}$ and $\Sigma_{(B,K)}$ are the spines for the K-handlebodies (A, K) and (B, K). Consider $H\colon P_K \times (I, \partial I) \to (M_K, \Sigma_{(A,K)} \cup \Sigma_{(B,K)})$, a sweep-out of P_K

between the two spines. For a fixed generic value of t, $H(P_K,t)$ will be denoted by P_K^t. By slightly abusing notation we will continue to denote by A_K and B_K the two components of $M_K - P_K^t$ and let $Q_K^A = Q_K \cap A_K$ and $Q_K^B = Q_K \cap B_K$. During the sweep-out, P_K^t and Q_K intersect generically except in a finite collection of values of t. Let $t_1,\dots,t_{n-1}$ be these critical values separating the unit interval into regions where P_K^t and Q_K intersect transversely. For a generic value t of H, the surfaces Q_K and P_K^t intersect in a collection of simple closed curves. After removing all removable curves, label a region $(t_i,t_{i+1}) \subset I$ with the letter A^* (resp B^*) if Q_K^A (resp Q_K^B) has a disk or punctured disk component in the region whose boundary is essential in P_K.

Suppose Q_K^A say, can be isotoped off some spine $\Sigma_{(A,K)}$. Then, using the product structure between the spines and the fact that all boundary components of Q_K lying on the knot are meridional, we can push Q_K to lie entirely in B_K contradicting Lemma 2.9. Therefore Q_K must intersect both spines $\Sigma_{(A,K)}$ and $\Sigma_{(B,K)}$ in meridional circles and so the subintervals adjacent to the two endpoints of the interval are labeled A^* and B^* respectively.

Case 1 Suppose there is an unlabeled region. If some curve of $Q_K \cap P_K$ is inessential in P_K in that region, it must also be inessential in Q_K as otherwise it would bound a c–disk for Q_K. Suppose some curve is essential in P_K but inessential in Q_K. This curve would give rise to one of the labels A^* or B^* contradicting our assumption. We conclude that all curves of $P_K \cap Q_K$ are mutual. In fact this implies that all curves $P_K \cap Q_K$ are essential in Q_K and in P_K as otherwise they would be removable by Lemma 2.6 and all removable curves have already been removed. Suppose Q_K^A say has a c–disk. The boundary of this c–disk would also be essential in Q_K contradicting the hypothesis thus we conclude that in this region Q_K^A and Q_K^B satisfy the hypothesis of Corollary 5.3 and thus $d(K,P) \le 2 - \chi(Q_K)$.

Case 2 Suppose there are two adjacent regions labeled A^* and B^*. (This includes the case when one or both of these regions actually have both labels)

The labels are coming from possibly punctured disk components of $Q_K - P_K$ that we will denote by D_A^* and D_B^* respectively. Using the triangle inequality we obtain

$$d(K,P) \le d(\mathcal{A}, \partial D_A^*) + d(\partial D_A^*, \partial D_B^*) + d(\partial D_B^*, \mathcal{B}). \tag{1}$$

The curves of intersection before and after going through the critical point separating the two regions can be made disjoint so $d(\partial D_A^*, \partial D_B^*) \le 1$ (the proof of this fact is similar to the proof of the last item of Lemma 2.8). By Proposition 4.2 $d(\mathcal{A}, \partial D_A^*), d(\mathcal{B}, \partial D_B^*) \le 1$ so the equation above gives us that $d(K,P) \le 3 \le 2 - \chi(Q_K)$ as long as $\chi(Q_K) < 0$.

If $\chi(Q_K)=0$ and Q_K is a torus, D_A^* and D_B^* must be disks, so $d(\mathcal{A},\partial D_A^*)=d(\mathcal{B},\partial D_B^*)=0$. Thus (1) gives us that $d(K,P)\le d(\partial D_A^*,\partial D_B^*)\le 1\le 2-\chi(Q_K)$. If Q_K is an essential annulus, we conclude that $d(K,P)\le 3$ □

Corollary 5.8 *Suppose $K=K_1\#K_2$, then any bridge surface for K has distance at most 3.*

Proof The sphere that decomposes K into its factors suggests an essential annulus in M_K. □

6 Edgeslides

This section is meant to provide a brief overview of edgeslides as described by Rubinstein and Scharlemann in [6]. Here we only give sketches of the relevant proofs and references for the complete proofs.

Suppose $(Q,\partial Q)\subset(M,P)$ is a bicompressible splitting surface in an irreducible 3–manifold with $P\subset\partial M$ a compact sub-surface, (in our context M will be a K–handlebody and P its punctured boundary). Let X,Y be the two components of $M-Q$ and let Q_X be the result of maximally compressing Q into X. The compressions can be undone by tubing along the edges of a graph Γ dual to the compressing disks, ie Q is contained in the boundary of a regular neighborhood of $Q_X\cup\Gamma$. We will denote by X^- and Y^+ the components of $M-Q_X$ with $X\supset X^-$ and $Y\subset Y^+$, in particular $\Gamma\subset Y^+$. Let $\Delta\subset Y$ be a set of compressing disks for Q. Using the fact that Q retracts to $Q_X\cup\Gamma$ we can extend these disks so that $\partial\Delta\subset Q_X\cup\Gamma$. Finally T will be a disk in Y^+ with $\partial T\subset(Q_X\cup P)$ that is not parallel to a subdisk of $Q_X\cup P$ and Λ will be the graph in T defined by the intersection of $\Delta\cup\Gamma$ and T. In other words Γ has vertices given by the points $T\cap\Gamma$ and edges given by the arcs $T\cap\Delta$.

The graph Γ described above is not unique; choosing a different graph is equivalent to an isotopy of Q. All graphs that are dual to the same set of compressing disks are related by *edge slides*, ie sliding the endpoint of some edge along other edges of Γ. The precise definition can be found in Saito et al [7] or Scharlemann and Thompson [10].

The following lemma is quite technical, a detailed proof of a very similar result can be found in [7, Proposition 3.2.2] or [10, Proposition 2.2]. We will only briefly sketch the proof here but we will provide detailed references to the corresponding results in [7] and note that there the letter P is used for the disk we call T but all other notation is identical.

Lemma 6.1 *Suppose T, Δ and Γ are as above. Suppose T', Δ' and Γ' is a second set of choices for a disk, a set of compressing disks and a graph as described above such that T' isotopic to T, rel. ∂T, $|\Delta'| = |\Delta|$ and Γ' is obtained from Γ via edge slides. Then either we can choose T', Δ' and Γ' so that the corresponding graph Λ' has an isolated vertex, or, we can choose them so that $\Gamma' \cup \Delta'$ is disjoint from T'.*

Proof Suppose every choice of T', Δ' and Γ' results in a graph Λ' with no isolated vertices. Pick an isotopy class of T rel. ∂T, an isotopy class of Δ and a representation of Γ such that $(|T \cap \Gamma|, |T \cap \Delta|)$ is minimal in the lexicographic order.

Claim 1 Each component of $T \cap \Delta$ is an arc [7, Lemma 3.2.3].

Suppose $T \cap \Delta$ contains a closed curve component. The innermost such in Δ, ω bounds a disk D_0 in Δ disjoint from T. Via an isotopy of the interior of T, using the fact that M is irreducible, the disk ω bounds in T can be replaced with D_0 thus eliminating at least ω from $T \cap \Delta$ contradicting minimality. As there are no simple closed curves, in this context a loop will mean an edge with both of its endpoints on the same vertex of Λ.

Claim 2 Λ has no inessential loops, that is edges with both endpoints on the same vertex of Λ that bound disks in $T - \Gamma$ [7, Lemma 3.2.4].

Suppose μ is a loop in Λ and let $D \in \Delta$ be such that $\mu \subset D$. The loop μ cuts off a disk $E \subset T$. As a subset of D, μ is an arc dividing D into two subdisks D_1 and D_2. (The disk E resembles a boundary compressing disk for D if we think of $\eta(\Gamma)$ as a boundary component.) At least one of $D_1 \cup E$ and $D_2 \cup E$ must be a compressing disk. Replace D with this disk reducing $|T \cap \Delta|$.

Claim 3 Λ has no isolated vertices [7, Lemma 3.2.5].

This is true by hypothesis.

Claim 4 Every vertex of Λ is a base of a loop [7, Lemma 3.2.6].

Suppose w a vertex of Λ is not a base of any loop, we will show we can reduce $(|T \cap \Gamma|, |T \cap \Delta|)$.

Let σ be the edge of Γ such that $w \in \sigma \cap T$. As w is not isolated, there is a disk $D \in \Delta$ such that $w \in \partial D$. The collection of arcs $D \cap T$ is a subset of the edges of Λ. Let γ be an outermost arc in D of all arcs that have w as one endpoint. Let w' be the other end point of γ. Then γ cuts a subdisk D_γ from D the interior of which may intersect T but ∂D_γ only contains one copy of $w \in \partial\gamma$. Thus there cannot be an entire copy of the edge σ in ∂D_γ and so there are three possibilities.

Case 1 $(\partial D_\gamma - \gamma) \subset \sigma$. Then we can perform an edge slide of σ which removes γ from Λ, [7, Figure 23].

Case 2 $(\partial D_\gamma - \gamma)$ contains some subset of σ with only one copy of one of the endpoints of σ. By sliding σ along D_γ we can reduce this case to the first case, [7, Figure 24].

Case 3 $(\partial D_\gamma - \gamma)$ contains some subset of σ but it contains two copies of the same endpoint of σ. This is the most complicated case requiring *broken edge slides* and [7, Figure 25] has an excellent discussion on the topic.

By the above four claims we can conclude that $\Lambda = \varnothing$ as desired, for by claim 4 some loop must be inessential contradicting claim 2. □

Remark 6.2 *If Q is weakly incompressible, the hypothesis of the lemma are satisfied as a meridional circle of an isolated vertex of Λ will be a compressing disk for Q in X that is disjoint from the set of compressing disks $\Delta \in Y$.*

Corollary 6.3 *Let $(Q, \partial Q) \subset (M, \partial M)$ be a bicompressible weakly incompressible surface splitting M into component X and Y. Let Q_X be the result of maximally compressing Q into X. Then Q_X is incompressible in M.*

Proof The argument is virtually identical to the argument in Scharlemann [8]. Suppose Q_X is compressible with compressing disk D that necessarily lies in Y^+. Let E be a compressing disk for Q in Y. As Q is weakly incompressible, by the above remark we can apply Lemma 6.1, with D playing the role of T, and $\Delta = E$. By Lemma 6.1 we can arrange that $(E \cup \Gamma) \cap D = \varnothing$ so D is also a compressing disk for Q in Y and is disjoint from Γ and thus from all compressing disks for Q in X contradicting weak incompressibility of Q. □

Corollary 6.4 *Suppose (A, K) is a K–handlebody with $\partial A = P$ and F is a bicompressible surface splitting A into submanifolds X and Y. Let F_K^X be the result of maximally compressing F_K into X_K. Then there exists a compressing disk D for P_K that is disjoint from a complete collection of compressing disks for F_K in X_K and intersects F_K only in arcs that are essential in F_K^X.*

Proof Select a disk D and isotope F_K^X to minimize $|D \cap F_K^X|$ and then choose a representation of Γ that minimizes $|D \cap \Gamma|$. As $A - N(K)$ is irreducible, by an innermost disk and outermost arc arguments, D intersect F_K^X in essential arcs only. Applying Lemma 6.1 with the disk T playing the role of D, we conclude that Γ is disjoint from D. As the edges of Γ are dual to a complete collection of compressing disks for F_K in X_K, it follows that D is disjoint from this collection. □

7 Bounds on distance given by a c–weakly incompressible surface

Our ultimate goal in this paper is to extend Theorem 5.7 to allow for both P and Q to be bridge surfaces for the same knot. To do this, we need a theorem similar to Proposition 4.3 but allowing for F_K to have certain kinds of c–disks.

Notation 7.1 *In this section let (A, K) be a K–handlebody with boundary P such that if A is a ball, K has at least 3 components and let $F \subset A$ be a properly embedded surface transverse to K splitting A into submanifolds X and Y.*

Definition 7.2 *The surface F_K associated to F is called bicompressible if F_K has some compressing disks in both X_K and Y_K. The surface is called cut-bicompressible if it has cut-disks in both X_K and Y_K. Finally, the surface is called c–bicompressible if it has c–disks in both X_K and Y_K.*

The next definition is an adaptation of the idea of a weakly incompressible surface but taking into consideration not only compressing disks but also cut-disks.

Definition 7.3 *The surface F_K is called c–weakly incompressible if it is c–bicompressible and any pair D^*_X, D^*_Y of c–disks contained in X_K and Y_K respectively intersect along their boundary.*

Proposition 7.4 *If a splitting surface $F_K \subset A_K$ has a pair of two compressing disks or a compressing disk and a cut-disk that are on opposite sides of F_K and intersect in exactly one point, then F_K is c–strongly compressible.*

Proof Suppose $F \subset A$ splits A into manifolds X and Y and let $D_X \subset X$ and $D_Y \subset Y$ be a pair of disks that intersect in exactly one point. Then a neighborhood of $D_X \cup D_Y$ contains a pair of compressing disks on opposite sides of F_K with disjoint boundaries (in fact their boundaries are isotopic). If D_X say is a compressing disk and D_Y is a cut-disk, banding two copies D_X together along ∂D_Y produces a compressing disk disjoint from D_Y, see Figure 2. □

Proposition 7.5 *Let $F_K \subset A_K$ be a c–weakly incompressible splitting surface such that every component of $F_K \cap P_K$ is mutual and let F'_K be the surface obtained from F_K via a P–compression. If F'_K is also c–bicompressible, then every component of $F'_K \cap P_K$ is essential in P_K or is mutually inessential.*

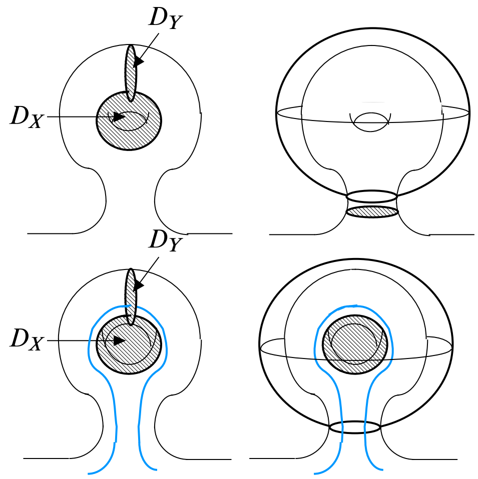


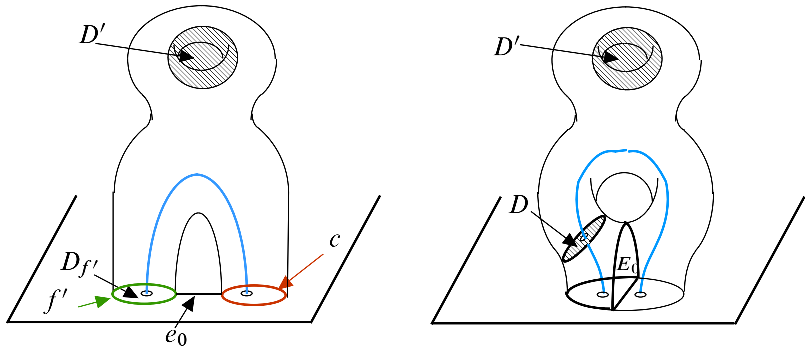
Figure 2

Proof Let X and Y be the two components of $A - F$. Without loss of generality, let $E_0 \subset X_K$ be the P–compressing disk for F_K. Suppose that there is some $f' \subset \partial F'_K$ that bounds a possibly punctured disk $D_{f'}$ in P_K but not in F'_K. The original surface F_K can be recovered by tunneling F'_K along an arc $e_0 \subset P_K$. As all curves of $F_K \cap P_K$ are mutual, $e_0 \cap f' \neq \varnothing$.

Case 1 e_0 has one boundary component in f' and the other in some other curve $c \in P_K \cap F_K$ (c may or may not be essential in P_K). If $c \subset D'_f$, then $F_K \cap P_K$ also has a curve that is inessential in P_K but essential in F_K contrary to the hypothesis. If c does not lie in D'_f then by slightly pushing the disk $D = D'_f \cup E_0$ away from P_K we obtain a c–disk for F_K contained in X_K, see Figure 3. By hypothesis F'_K is c–bicompressible, in particular there is a c–disk D' for F'_K that lies on the other side of F'_K than the side D'_f lies on. The c–disk D' is also a c–disk for F_K lying in Y that is disjoint from $D \subset X$ contradicting the c–weak incompressability of F_K.

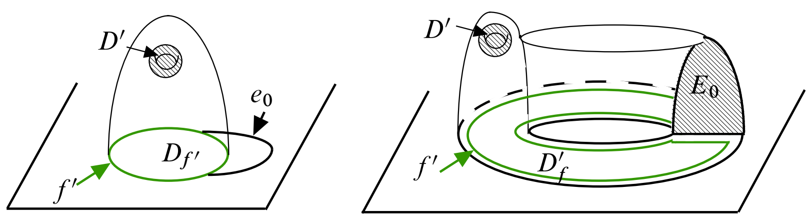
Case 2 e_0 has both boundary components in f'. If $e_0 \subset D'_f$ then again $F_K \cap P_K$ has a curve that is inessential in P_K but essential in F_K contrary to the hypothesis so assume $e_0 \cap D'_f = \partial e_0$, see Figure 4. Consider the possibly punctured disk D obtained by taking the union of $D_{f'}$ together with two copies of E_0. As in the previous case this is a c–disk for F_K lying in X_K that is disjoint from at least one c–disk for F_K lying in Y_K contradicting c–weak incompressibility of F_K. □

Proposition 7.6 *Suppose F'_K splitting A_K into X'_K and Y'_K satisfies one of the following two conditions:*



F'_K F_K

Figure 3



F'_K F_K

Figure 4

- *there is a spine* $\Sigma_{(A,K)}$ *entirely contained in* X'_K *say and* F'_K *has a c–disk in* X'_K *disjoint from that spine or*
- *there is at least one curve* $f' \subset F'_K \cap P_K$ *that is essential in* P_K *and* $d(f', \mathcal{A}) \leq 1 - \chi(F'_K)$.

If F_K *is obtained from* F'_K *by tunneling or tubing (possibly along subarcs of* K*) with all tubes lying in* Y'_K*, then* F_K *satisfies one of the following conditions:*

- *there is a spine* $\Sigma_{(A,K)}$ *entirely contained in* X_K*, and* F_K *has a c–disk in* X_K *disjoint from that spine or*
- *for every curve* f *in* $F_K \cap P_K$ *that is essential in* P_K *the inequality* $d(f, \mathcal{A}) \leq 1 - \chi(F_K)$ *holds.*

Proof Suppose first that F_K is obtained from F'_K via tunneling. If F'_K satisfies the first condition, then tunneling does not interfere with the c–disk and does not introduce intersections with the spine $\Sigma_{(A,K)}$. If F'_K satisfies the second condition, note that

$d(f, f') \leq 1$ for every $f \in F_K \cap P_K$ that is essential in P_K and $\chi(F'_K) \geq \chi(F_K) + 1$. The result follows by the triangle inequality.

If F_K is obtained from F'_K via tubing with all tubes contained in Y'_K, these tubes do not affect a c–disk for F'_K contained in X'_K and are disjoint from any spine $\Sigma_{(A,K)}$. Thus if F'_K satisfies the first condition, so does F_K. If F'_K satisfies the second condition, the curves of $P_K \cap F'_K$ are not altered by the tubing and $1 - \chi(F_K) \geq 1 - \chi(F'_K)$ so for any curve essential curve $f \in F_K \cap P_K$, $d(f, \mathcal{A}) \leq 1 - \chi(F'_K) \leq 1 - \chi(F_K)$ as desired. $\square$

The rest of this section will be dedicated to the proof of the following theorem.

Theorem 7.7 *Let A_K be a K–handlebody with $\partial A = P$ such that if P is a sphere, then P_K has at least six punctures. Suppose $F_K \subset A_K$ satisfies the following conditions:*

- *F_K has no closed components,*
- *F_K is c–bicompressible and c–weakly incompressible,*
- *F_K has no disk components and*
- *all curves of $P_K \cap F_K$ are mutually essential unless they bound punctured disks in both surfaces.*

Then at least one of the following holds:

- *There is a spine $\Sigma_{(A,K)}$ entirely contained on one side of F_K and F_K has a c–disk on the same side disjoint from the spine or*
- *$d(f, \mathcal{A}) \leq 1 - \chi(F_K)$ for every $f \in F_K \cap P_K$ that is essential in P_K unless f is the boundary of a P_K–parallel annulus component of F_K.*

Proof If c–disks for F_K were incident to two different components of F_K, then there would be a pair of such disks on opposite sides of F_K with disjoint boundaries violating c–weak incompressibility. So we deduce that all c–disks for F_K are incident to at most one component S_K of F_K. The surface S_K cannot be an annulus, else the boundaries of c–disks in X_K and Y_K would be parallel and so could be made disjoint. In particular S_K, and thus F_K, must have a strictly negative Euler characteristic. Suppose F_K is a counterexample to the theorem such that $1 - \chi(F_K)$ is minimal amongst all such counterexamples. As in Proposition 4.3 we may assume that F_K has no components that are P_K–parallel annuli or P_K–parallel punctured disk components. In particular this implies that all curves of $F_K \cap P_K$ are mutually essential. We will prove the theorem in a sequence of lemmas. We will use the following definition modelled after the definition of a strongly ∂–compressible surface first introduced by Scharlemann in [8].

Definition 7.8 *A splitting surface $F_K \subset A_K$ splitting A into submanifolds X and Y is called strongly P–compressible if there exist P–compressing disks $E_X \subset X_K$ and $E_Y \subset Y_K$ for F_K such that $\partial E_X \cap \partial E_Y = \varnothing$.*

Lemma 7.9 *Suppose F_K is the surface that provides a counterexample to Theorem 7.7 with maximal Euler characteristic. In other words F_K is the maximal Euler characteristic surface satisfying all of the following conditions:*

- *F_K has no closed components,*
- *F_K is c–bicompressible and c–weakly incompressible,*
- *F_K has no disk components,*
- *all curves of $P_K \cap F_K$ are mutually essential unless they bound punctured disks in both surfaces,*
- *if some spine $\Sigma_{(A,K)}$ is entirely contained in X_K say, then every c–disk for F_K contained in X_K intersects this spine and*
- *there is some curve $f \in F_K \cap P_K$ that is essential in P_K and not the boundary of a P_K–parallel annulus component of F_K such that $d(f, \mathcal{A}) > 1 - \chi(F_K)$.*

Then F_K is not strongly P–compressible.

Proof By way of contradiction suppose $E_X \subset X_K$ and $E_Y \subset Y_K$ is a pair of disjoint P–compressing disks for F_K. Let F_K^x, F_K^y denote the surfaces obtained from F_K by P–compressing F_K along E_X and E_Y respectively, and let F_K^- denote the surface obtained by P–compressing along both disks simultaneously. A standard innermost disk, outermost arc argument between E_X and a c–disk for F_K in X_K shows that F_K^x has a c–disk lying in X_K. Similarly, F_K^y has a c–disk lying in Y_K. If one of F_K^x or F_K^y has c–disks on both sides, say F_K^x, then all curves of $P_K \cap F_K^x$ must be mutually essential unless they bound punctured disks in both surfaces by Proposition 7.5. The surface F_K^x cannot be the union of punctured disks as it is bicompressible so at least one component of $F_K^x \cap P_K$ is essential in P_K. As $1-\chi(F_K^x) < 1-\chi(F_K)$ the surface F_K^x satisfies one of the conclusions of the theorem. By Proposition 7.6 tunneling to recover F_K from F_K^x preserves either of these properties so F_K is not a counterexample as we assumed.

If F_K^- has any c–disk, then one of F_K^x or F_K^y has c–disks on both sides as c–disks are preserved under tunneling and we are done as above. Suppose some curve of $F_K^- \cap P_K$ is inessential in P_K but essential in F_K^-. This curve must be adjacent to the dual arc to one of the P–compressing disks, say the dual arc to E_X. In this case, by an argument similar to the proof of Proposition 7.5, F_K^y is c–compressible in X_K. As we saw that

F_K^y is c–compressible in Y_K, it follows that F_K^y is c–bicompressible, a case we have already considered. Thus all curves essential in F_K^- are also essential in P_K, therefore if F_K^- has a component that is not P_K–parallel, the result follows from Proposition 4.3.

We have reduced the proof to the case that F_K^- is c–incompressible, each component of F_K^- is P_K–parallel and all curves of $P_K \cap F_K^-$ are essential in P_K or mutually inessential. It is clear that in this case we can isotope F_K^- to be disjoint from any spine $\Sigma_{(A,K)}$. The original surface F_K can be recovered from F_K^- by tunneling along two arcs on opposite sides of F_K^-. The tunnels can be made disjoint from $\Sigma_{(A,K)}$ and thus F_K can also be isotoped to be disjoint from $\Sigma_{(A,K)}$. Without loss of generality we will assume $\Sigma_{(A,K)} \subset X_K$, thus it suffices to show that F_K has a c–disk in X_K that is disjoint from $\Sigma_{(A,K)}$.

Consider how F_K^x can be recovered from F_K^-; the P–compression into Y_K must be undone via a tunnelling along an arc γ where the interior of γ is disjoint from F_K^-. Let γ connect components F_K^0 and F_K^1 (possibly $F_K^0 = F_K^1$) of F_K^- where F_K^i is parallel to a subsurface $\widetilde{F}_K^i \subset P_K$. There are three cases to consider. First assume that $F_K^0 \neq F_K^1$ and they are nested, ie $\widetilde{F}_K^0 \subset \widetilde{F}_K^1$. Consider the eyeglass curve $e = \eta(\gamma \cup \omega)$ where $\omega \subset F_K^0$ is parallel to the boundary component of F_K^0 that is adjacent to γ. Using the product structure between F_K^0 and F_K^1, a neighborhood of $e \times I$ contains the desired compressing disk for F_K that is disjoint from some spine $\Sigma_{(A,K)}$.

Next suppose $F_K^0 \neq F_K^1$ and they are not nested. Then each component of F_K^x is P_K–parallel. As we have already seen, F_K^x has a c–disk in X_K. The c–disk is either disjoint from some $\Sigma_{(A,K)}$, in which case we are done, or, via the parallelism to P_K, the c–disk represents a c–disk D^* for P_K in A_K whose boundary is disjoint from at least one curve in ∂F_K^x; the curve that is in the boundary of the c–compressible component of F_K^x. Call this particular curve f^x. If $\chi(F_K) < -1$ then $d(f^x, \partial D^*) \leq 1$ so $d(f, \mathcal{A}) \leq 3 \leq 1 - \chi(F_K)$. If $\chi(F_K) = -1$, then F_K^x consists only of P_K–parallel annuli and punctured disks components. Let N be the annulus component of F_K^x with boundary f^x parallel to a subannulus $\widetilde{N} \subset P_K$. Then f^x and ∂D^* both lie in $\widetilde{N}$ so $d(f^x, \partial D^*) = 0$. By Proposition 4.1 $d(\partial D^*, \mathcal{A}) \leq 1$. Thus for f any essential component of ∂F_K, $d(f, \mathcal{A}) \leq d(f, f^x) + d(f^x, \partial D^*) + d(\partial D^*, \mathcal{A}) \leq 1 + 1 = 1 - \chi(F_K)$.

The last case to consider is the case $F_K^0 = F_K^1$. If $\gamma \subset \widetilde{F}_K^0$ then $\gamma \times I$ is the desired compressing disk. If γ is disjoint from $\widetilde{F}_K^0$, then each component of F_K^x is P_K–parallel. Proceed as in the previous case to show that either F_K^x, and thus F_K, has a c–disk disjoint from $\Sigma_{(A,K)}$ or $d(f, \mathcal{A}) \leq 1 - \chi(F_K)$. □

Lemma 7.10 *If the surface F_K that provides a counterexample to Theorem 7.7 with maximal Euler characteristic is bicompressible, then the surfaces F_K^X and F_K^Y obtained from F_K by maximally compressing F_K into X_K and Y_K respectively have cut-disks.*

Remark 7.11 *Note that the hypothesis of this lemma holds when F_K does not have any cut-disks.*

Proof Suppose F_K^X say has no cut-disks. By Corollary 6.3 the surfaces F_K^X and F_K^Y are incompressible in A_K. If some component of F_K^X is not P_K–parallel, then the second conclusion of the theorem follows from Proposition 4.3. We may therefore assume that there is some spine $\Sigma_{(A,K)}$ that is disjoint from F_K^X.

Let X_K^- and Y_K^+ be the two sides of F_K^X and let $\Gamma \subset Y_K^+$ be the graph dual to the compressions we performed, ie F_K can be recovered from F_K^X by tunneling along the edges of Γ. Note that by general position we can always arrange that Γ is disjoint from any spine so in particular after an isotopy, $F_K \cap \Sigma_{(A,K)} = \varnothing$.

Claim *Recall that S_K is the component of F_K to which all c–disks for F_K are incident. To prove the lemma at hand it suffices to show that*

- *S_K has a c–disk D^* on the same side of S_K as the spine $\Sigma_{(A,K)}$ and disjoint from that spine or*
- *there is a compressing disk for P_K whose boundary is disjoint from at least one curve in ∂S_K or*
- *S_K is strongly P–compressible.*

Proof By an innermost disk argument we may isotope any c–disk for S_K to be disjoint from F_K.

In the first case we assume S_K has c–disk D^* on the same side of S_K as the spine $\Sigma_{(A,K)}$ and disjoint from that spine. Recall that $F_K \cap \Sigma_{(A,K)} = \varnothing$ so it is sufficient to show that F_K also has a c–disk on the same side as $\Sigma_{(A,K)}$ but disjoint from it. Note that D^* is not necessarily on the same side of F_K as the spine.

If there is a component of F_K that separates D^* and $\Sigma_{(A,K)}$ than this component also separates S_K and all its c–disks from the spine. As S_K is bicompressible, we can always find a c–disk for S_K on the same side as $\Sigma_{(A,K)}$ and all these c–disks will be disjoint from the spine. If there is no such separating component, then D^* is a c–disk for F_K on the same side as $\Sigma_{(A,K)}$ but disjoint from $\Sigma_{(A,K)}$.

In the second case, $d(s, \mathcal{A}) \leq 1$ where $s \in \partial S_K$ so $d(f, \mathcal{A}) \leq 2 \leq 1 - \chi(F_K)$.

In the third case, suppose first that all components of $F_K - S_K$ are annuli, necessarily not P_K–parallel. If one of these annuli is P–compressible, P–compressing it results in a compressing disk for P_K that is disjoint from F_K so $d(f, \mathcal{A}) \leq 1$. Thus we may assume that all other components of F_K are P–incompressible. By an innermost disk and outermost arc arguments, the pair of strongly P–compressing disks for S_K can be isotoped to be disjoint from all other components of F_K so F_K is also strongly P–compressible and by Lemma 7.9, F_K cannot be a counterexample to the theorem.

If some component of F_K other than S_K has a strictly negative Euler characteristic, then $1-\chi(S_K) < 1-\chi(F_K)$. This shows that S_K is not a counterexample to the theorem, so either $d(s, \mathcal{A}) \leq 1-\chi(S_K)$ in which case $d(f, \mathcal{A}) \leq d(f, s) + d(s, \mathcal{A}) \leq 1-\chi(F_K)$ or S_K has a c–disk on the same side of S_K as the spine $\Sigma_{(A,K)}$ but is disjoint from it. By repeating the argument from the first case, we conclude that F_K must also satisfy the second conclusion of the theorem. This concludes the proof of the claim. □

Note that S_K is itself a c–weakly incompressible surface as every c–disk for the surface S_K is also a c–disk for F_K. We will prove the lemma by showing that S_K satisfies one of the items in the claim above. Let S split A into submanifolds U and V and S_K^U be the surface obtained by maximally compressing S_K in U_K, S_K^U splits A_K into submanifolds U_K^- and V_K^+ and Γ is the graph dual to the compressing disk. We have already shown that for some spine $\Sigma_{(A,K)}$, $\Sigma_{(A,K)} \cap F_K = \varnothing$ so in particular $\Sigma_{(A,K)} \cap S_K = \varnothing$. As S_K^U is c–incompressible, we may assume each component is P_K–parallel as otherwise the result will follow by Proposition 4.3. We will show that S_K satisfies one of the conditions in the claim.

If $\Sigma_{(A,K)} \subset U_K^-$, then $\Sigma_{(A,K)}$ is also disjoint from every compressing disk for S_K lying in U_K as it is disjoint from the meridional circles for the edges of Γ and we have the desired result. Thus we may assume $\Sigma_{(A,K)} \subset V_K^+$. Let S_K^0 be an outermost component of S_K^U, ie a component cobounding a product region $R_K \cong S_K^0 \times I$ with P_K such that $R_K \cap S_K^U = \varnothing$.

Case 1 Suppose for some outermost component, $R_K \subset V_K^+$. As $\Gamma \subset V_K^+$ and S_K is connected, S_K^0 is the only component of S_K^U. This implies that $\Sigma_{(A,K)} \subset R_K$ so we can use the product structure to push $\Sigma_{(A,K)}$ into U_K^- and by the previous paragraph S_K satisfies the hypothesis of the claim.

Case 2 Suppose the components of S_K^U are nested and let S_K^1 be a second outermost component. The region between S_K^1 and the outermost components of S_K^U is a product region that must be contained in V_K^+ or we can apply Case 1. Again as S_K is connected, V_K^+ is also connected so in fact V_K^+ is a product region and $\Sigma_{(A,K)} \subset V_K^+$. Again we can push $\Sigma_{(A,K)}$ into U_K^- and complete the argument as in the previous case.

Case 3 Finally suppose that the components of S_K^U are all outermost and all outermost regions are contained in U_K^-. By Corollary 6.4, there is a compressing disk for P_K that is disjoint from a complete collection of compressing disks for S_K in U_K and intersects S_K only in arcs that are essential in S_K^U. Take such a disk D that intersects S_K minimally. Consider an outermost arc of $S_K^U \cap D$ cutting off a subdisk D_0 from D. If $D_0 \subset V_K$, P–compressing S_K along D_0 preserves the compressing disks of S_K lying in U_K (because D is disjoint from all compressing disks for S_K in U_K by hypothesis) and also preserves the c–disks lying in V_K (by an innermost disk argument) so the result follows by induction. If every outermost disks is contained in U_K^-, the argument of [8, Theorem 5.4, Case 3] now carries over to show that either S_K is strongly P–compressible or there is a compressing disk for P_K that is disjoint from S_K. We repeat the argument here for completeness.

If there is nesting among the arcs $D \cap S_K$ in D, consider a second outermost arc λ_0 in D and let D' be the disk this arc cuts from D, see Figure 5. If every arc of $S_K^U \cap D$ is outermost of D let $D = D'$. Let $\Lambda \subset D'$ denote the collection of arcs $D' \cap S_K$; one of these arcs (namely λ_0) will be in $\partial D'$. Consider how a c–disk E^* for S_K in V_K intersects D'. All closed curves in $D' \cap E^*$ can be removed by a standard innermost disk argument redefining E^*. Any arc in $D' \cap E^*$ must have its ends in Λ; a standard outermost arc argument can be used to remove any that have both ends in the same component of Λ. If any component of $\Lambda - \lambda_0$ is disjoint from all the arcs $D' \cap E^*$, then S_K could be P–compressed without affecting E^*. This reduces $1 - \chi(S_K)$ without affecting bicompressibility, so we would be done by induction. Hence we restrict to the case in which each arc component of $\Lambda - \lambda_0$ is incident to some arc components of $D' \cap E^*$.

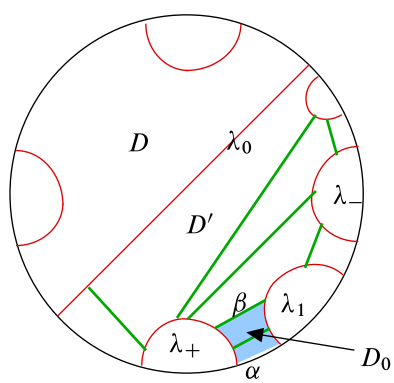


Figure 5

It follows that there is at least one component $\lambda_1 \neq \lambda_0$ of Λ with this property: any arc of $D' \cap E^*$ that has one end incident to λ_1 has its other end incident to one of the (at most two) neighboring components $\lambda_\pm$ of Λ along $\partial D'$. (Possibly one or both of $\lambda_\pm$ are λ_0.) Let β be the outermost arc in E^* among all arcs of $D' \cap E^*$ that are incidental to the special arc λ_1. We then know that the other end of β is incident to (say) λ_+ and that the disk $E_0 \subset E^*$ cut off by β from E^*, although it may be incident to D' in its interior, at least no arc of intersection $D' \cap \mathrm{interior}(E_0)$ is incident to λ_1. Notice that even if E^* is a cut-disk, we can always choose E_0 so that it does not contain a puncture.

Let D_0 be the rectangle in D' whose sides consist of subarcs of λ_1, λ_+, $\partial D'$ and all of β. Although E^* may intersect this rectangle, our choice of β as outermost among arcs of $D \cap E^*$ incident to λ_1 guarantees that E_0 is disjoint from the interior of D_0 and so is incident to it only in the arc β. The union of E_0 and D_0 along β is a disk $D_1 \subset V_K$ whose boundary consists of the arc $\alpha = P \cap \partial D_0$ and an arc $\beta' \subset S_K$. The latter arc is the union of the two arcs $D_0 \cap S_K$ and the arc $E_0 \cap S_K$. If β' is essential in F_K, then D_1 is a P–compressing disk for S_K in V_K that is disjoint from the P–compressing disk in U_K cut off by λ_1. So if β' is essential then S_K is strongly P–compressible. Suppose finally that β' is inessential in S_K so β' is parallel to an arc in ∂S_K. Let $D_2 \subset S_K$ be the disk of parallelism and consider the disk $D' = D_1 \cup D_2$. Note that $\partial D' \subset P_K$ and D' can be isotoped to be disjoint from S_K. Either D' is P_K–parallel or is itself a compressing disk for P_K. In the latter case $\partial D' \in \mathcal{A}$, $d(f, \mathcal{A}) \leq 1$ for every $f \in \partial S_K$ and we are done. On the other hand if D' cobounds a ball with P_K, then D_1 and D_2 are parallel and so we can isotope S_K replacing D_2 with D_1. The result of this isotopy is the curves λ_1 and λ_+ are replaced by a single curve containing β as a subarc lowering $|D \cap S_K|$. This contradicts our original assumption that S_K and D intersect minimally. We conclude that S_K satisfies the second or the third condition of the Claim completing the proof of Lemma 7.10. □

We return now to the proof of the theorem. By the above lemmas we may assume F_K is not strongly P–compressible, and if it is bicompressibleboth of F_K^X and F_K^Y have cut-disks.

Remark 7.12 *Some of the argument to follow here parallels the argument in [8, Theorem 5.4]. In fact it seems likely that the stronger result proven there still holds.*

If F_K has no compressing disks on some side (and necessarily has a cut-disk), pick that side to be X_K. If both sides have compressing disks, pick X_K to be the side that has a cut-disk if there is such. Thus if F_K has a cut-disk, then it has a cut-disk

$D^c \subset X_K$ and if F_K has a compressing disk lying in X_K, it also has a compressing disk lying in Y_K.

Suppose F_K has a cut-disk $D^c \subset X_K$. Let κ be the component of $K - P$ that pierces through D^c and B be a bridge disk of κ. We want to consider how F_K intersects B. After a standard innermost disk argument, we may assume that the cut-disk D^c intersects B in a single arc μ with one endpoint lying in κ and the other endpoint lying in a component of $F_K \cap B$. Label this component b (see Figure 6). The curve b is either a simple closed curve, has both of its endpoints in P_K or has at least one endpoint in κ.

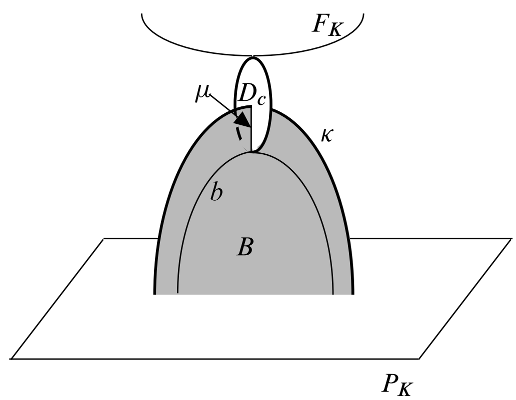


Figure 6

Assume $|B \cap F_K|$ is minimal. We will first show that if there are any simple closed curves of intersection, they cannot be nested in B. The argument is similar to the No Nesting Lemma in Scharlemann [9].

Suppose such nesting occurs and let δ be a second innermost curve cutting off a disk D_δ from B. The innermost curve of intersection contained in D_δ bound compressing disks for F_K disjoint from D^c and thus must lie in X_K, call these disks $D_1, \dots, D_n$. By our choice of labels this implies that F_K is in fact bicompressible, let E be a compressing disks for F_K lying in Y. By c–weak incompressability of F_K, $E \cap D_i \neq \varnothing$. By using edgeslides guided by E as in the proof of Lemma 6.1 $|B \cap F_K|$ can be reduced contradicting minimality.

We can in fact assume that there are no simple closed curves of intersection between F_K and the interior of B. Suppose $\sigma \neq b$ is an innermost simple closed curve of intersection bounding a subdisk $D_\sigma \subset B$. This disk is a compressing disk for F_K

disjoint from D^c so must lie in X_K by c–weak incompressibility of F_K. Thus F_K must also have a compressing disk in Y_K. Use this compressing disk and apply Lemma 6.1 with the subdisk of B bounded by b playing the role of T to isotope F_K so as to remove all such closed curves.

Suppose b is a simple closed curve. Let $D_b \subset B$ be the disk b bounds in B. Then by the above $D_b \cap F_K = b$ and thus D_b is a compressing disk for F_K lying in Y_K intersecting D^c in only one point contradicting Proposition 7.4. Thus we may assume b is an arc.

Case 1 There exists a cut-disk $D^c \subset X_K$ such that the arc b associated to it has both of its endpoints in P_K.

Again let $D_b \subset B$ be the disk b bounds in B. By the above discussion $D_b \cap F_K$ has no simple closed curves. Let σ now be an outermost in B arc of intersection between F_K and B cutting from B a subdisk E_0 that is a P–compressing disk for F_K.

Subcase 1A $b = \sigma$ and so necessarily $E_0 \subset Y_K$. This in fact implies that $F_K \cap B = b$. For suppose there is an arc in $F_K \cap (B - D_b)$. An outermost such arc γ bounds a P–compressing disk for F_K. If this disk is in X_K, then F_K would be strongly P–compressible, a possibility we have already eliminated. If the disk is in Y_K, note that we can P–compress F_K along this disk preserving all c–disks for F_K lying in Y_K and also preserving the disk D^c. The theorem then follows by Proposition 7.6.

Consider the surface F'_K obtained from F_K via P–compression along D_b and the disk D_B obtained by doubling B along κ, a compressing disk for P_K. In this case $F'_K \cap D_B = \varnothing$ so we can obtain the inequality $d(f, \mathcal{A}) \le d(f, \partial D_B) \le d(f, f') + d(f', \partial D_B) \le 2$ for every curve $f \in P_K \cap F_K$ as long as we can find at least one $f' \in P_K \cap F'_K$ that is essential in P_K.

If all curves in $P_K \cap F'_K$ are inessential in P_K, there are at most two of them. Suppose F'_K has two boundary components f'_1 and f'_2 bounding possibly punctured disks $D_{f'_1}, D_{f'_2} \subset P_K$ and $F_K \cap P_K$ can be recovered by tunneling between these two curves. As all curves of $F_K \cap P_K$ are essential in P_K, each of $D_{f'_1}$ and $D_{f'_2}$ must in fact be punctured and they cannot be nested. Consider the curve f_* that bounds a disk in P and this disk contains $D_{f'_1}, D_{f'_2}$ and the two points of $\kappa \cap P$, (see Figure 7). This curve is essential in P_K as it bounds a disk with four punctures on one side the other side either does not bound a disk in P if P is not a sphere, or contains at least two punctures of P_K if P is a sphere. As f_* is disjoint from both the curve $F_K \cap P_K$ and from at least one curve of $\mathcal{A}$, it follows that the unique curve $f \in F_K \cap P_K$ satisfies the equality $d(F_K \cap P_K, \mathcal{A}) \le 2 \le 1 - \chi(F_K)$.

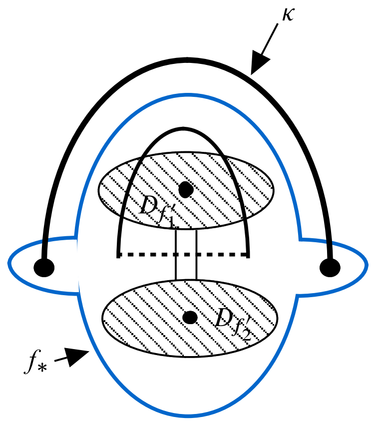


Figure 7

If F_K' has a unique boundary curve f' then F_K is recovered by tunneling along an arc e_0 with both of its endpoints in f'. Therefore F_K has exactly two boundary curves f_0, f_1 that cobound a possibly once punctured annulus in P_K (see Figure 8).

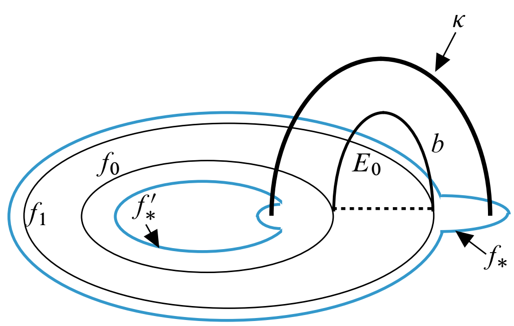


Figure 8

Let f_* and f_*' be the curves in P_K that cobound once punctured annuli with f_1 and f_0 respectively as in Figure 8. If both f_* and f_*' are inessential in P_K, then P_K is a sphere with at most four punctures contrary to the hypothesis. Thus we may assume that f_* say is essential in P_K. In this case $d(f_i, \mathcal{A}) \leq d(f_i, f_*) + d(f_*, \mathcal{A}) \leq 2 \leq 1 - \chi(F_K)$ for $i = 1, 2$ as desired.

Subcase 1B $b \neq \sigma$ and some disk $E_0 \subset D_b$ bound by an outermost arc of $F_K \cap D_b$ is contained in Y_K. (It can be shown that as in subcase 1A, $F_K \cap (B - D_b) = \varnothing$ but we won't need this observation). P–compressing via E_0 results in a surface F_K' with c–disks on both sides as E_0 is disjoint from D^c. By Proposition 7.5 F'_K satisfies the hypothesis and thus the conclusion of the theorem at hand and by Proposition 7.6 so does F_K contradicting our assumption that F_K is a counterexample.

Subcase 1C All outermost arcs of $F_K \cap D_b$ bound P–compressing disks contained in X_K. Consider a second outermost arc λ_0 in B (possibly b) and let D' be the disk this arc cuts from B. Let $\Lambda \subset D'$ denote the collection of arcs $D' \cap F_K$; one of these arcs (namely λ_0) will be in $\partial D'$. The argument is now identical to Case 3 of Lemma 7.10, and shows that F_K is strongly P–compressible, a possibility we have already eliminated, or $d(f, \mathcal{A}) \leq 1$. See Figure 9 for the pair of strongly P–compressing disks in this case.

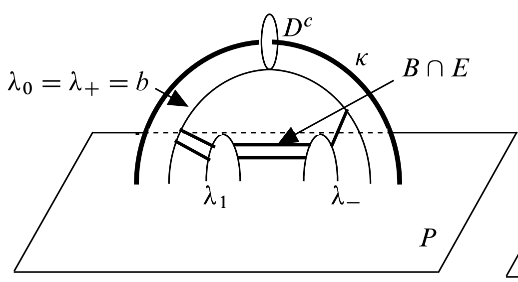


Figure 9

Case 2 No cut-disk for F_K has the property that the arc associated to it has both of its endpoints in P_K. In other words, every arc b associated to a cut-disk $D^c \subset X_K$ has at least one of its endpoints in κ. This also includes the case when F_K has no cut-disks at all.

First we will show that F_K actually has compressing disks on both sides. This is trivial if F_K has no cut-disks so suppose F_K has a cut-disk. Consider the triangle $R \subset B$ cobounded by μ, κ and b (See Figure 10). If R is disjoint from F_K, a neighborhood of $D^c \cup R$ contains a compressing disk D for F_K, necessarily contained in X_K. If $R \cap F_K \neq \varnothing$, there are only arcs of intersection as all simple closed curves have been removed. An outermost in R arc of intersection has both of its endpoint lying in κ and doubling the subdisk of R it cuts off results in a compressing disk D for F_K that also has to lie in X_K as its boundary is disjoint from D^c. These two types of disks will be called *compressing disks associated to* D^c. As F_K has a compressing disk in Y_K by our initial choice of labeling, F_K is bicompressible.

Compress F_K maximally in X_K to obtain a surface F_K^X. The original surface F_K can be recovered from F_K^X by tubing along a graph Γ whose edges are the cocores of the compressing disks for F_K on the X_K side. By Corollary 6.3 F_K^X does not have any compressing disks and by Lemma 7.10 it has cut-disks.

We will use X_K^- and Y_K^+ to denote the two sides of F_K^X and will show that in this case F_K^X doesn't have any cut-disks lying in X_K^-. Suppose $D'^c \subset X_K^-$ is a cut disk for F_K^X and B', b' are respectively the disk and the arc of $F_K^X \cap B'$ associated to it. Note that b' must have both of its endpoints in P_K as otherwise we can construct a compressing disk associated to D'^c and we have shown that F_K^X is incompressible. The original surface F_K can be recovered from F_K^X by tubing along the edges of a graph $\Gamma \subset Y_K^+$. This operation preserves the disk D'^c and b' so F_K also has a cut-disk whose associated arc has both of its endpoints of P_K contradicting the hypothesis of this case.

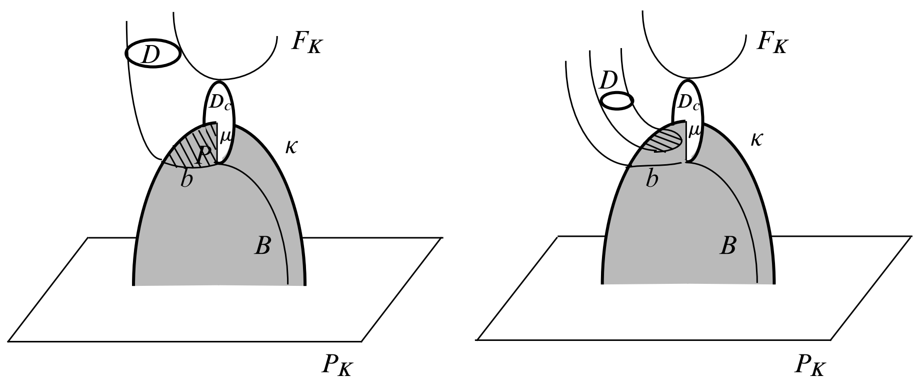


Figure 10

The remaining possibility is that F_K^X has a cut-disk in $D'^c \subset Y_K^+$. Let B' is its associated bridge disk, b' the arc of $F_K \cap B'$ adjacent to the cut-disk, D'_b is the disk b' cuts from B' and κ' the arc of the knot piercing D'^c. Assume $|F_K^X \cap B'|$ is minimal. There cannot be any circles of intersection for they would either be inessential in both surfaces or give rise to compressing disks for the incompressible surface F_K^X. Also the arc b' must have both of its endpoints in P, otherwise we can construct a compressing disk for F_K^X associated to D'^c, a similar situation is depicted in Figure 10. Consider an outermost arc of $D'_b \cap F_K^X$ cutting from D'_b a P–compressing disk E_0, possibly $D'_B = E_0$. We now repeat an argument similar to the argument in Case 1 but applied to F_K^X. There are again 3 cases to consider.

Subcase 2A $F_K^X \cap D'_b = b'$ so b' bounds a P–compressing disk for F_K^X lying in X_K^-. Let ${F'_K}^X$ be the surface obtained from F_K^X after this P–compression. The argument of Subcase 1A now shows $d(f, \mathcal{A}) \leq 2 \leq 2 - \chi(F_K)$ for every $f \in F_K^X \cap P_K = F_K \cap P_K$.

Subcase 2B Some E_0 lies in Y_K^+ (so b' is not an outermost arc). Pick a compressing disk D for F_K in Y_K as in Corollary 6.4. P–compressing F_K along E_0 does not affect c–disks lying in Y_K^+. It also preserves all compressing disks for F_K that lie in X_K as it is disjoint from the graph Γ and thus we are done by induction.

Subcase 2C All outermost arcs of $F_K^X \cap B'$ lie in X_K^-. Consider a second outermost arc component of $(F_K^X) \cap B'$ and let $E_1 \in D_b' - F_K^X$ be the disk it bounds, necessarily $E_1 \subset Y^+$. By Lemma 6.1 we may assume that Γ is disjoint from this disk. Let E be a compressing disk for F_K in Y_K. If $E \cap E_1 = \varnothing$ then P–compressing F_K along an outermost disk component preserves the compressing disk lying in Y_K and of course preserves all c–disks lying in X_K so we can finish the argument by induction. If there are arcs of intersection, we can repeat the argument of Subcase 1C to show that F_K is strongly boundary compressible, a case we have already eliminated. □

8 Distance and intersections of Heegaard splittings

For the remainder of this paper we will be considering the case of a closed orientable irreducible 3–manifold M containing a knot K with bridge surface P such that $M = A \cup_P B$. In this section we also assume that if P is a sphere then P has at least six punctures. The surface Q will be either a second bridge surface for K or a Heegaard surface for M_K. Let X and Y be the two components of $M - Q$. Thus if Q is a Heegaard splitting for the knot exterior, then one of X_K or Y_K is a compression body and the other component is a handlebody. If Q is a bridge surface, both X_K and Y_K are K-handlebodies.

Given a positioning of P_K and Q_K in M_K let Q_K^A and Q_K^B stand for $Q_K \cap A_K$ and $Q_K \cap B_K$ respectively. After removing all removable (see Definition 5.4) curves of intersection, proceed to associate to the configuration given by P_K and Q_K one or more of the following labels.

- Label A (resp B) if some component of $Q_K \cap P_K$ is the boundary of a compressing disk for P_K lying in A_K (resp B_K).
- Label A^c (resp B^c) if some component of $Q_K \cap P_K$ is the boundary of a cut-disk for P_K lying in A_K (resp B_K). (Notice that this labeling is slightly different from the labeling in Section 5 where the compressing disk was required to be a subdisk of Q_K.)
- X (resp Y) if there is a compressing disk for Q_K lying in X_K (resp Y_K) that is disjoint from P_K and the configuration does not already have labels A, A^c, B or B^c.

- X^c (resp Y^c) if there is a cut-disk for Q_K lying in X_K (resp Y_K) that is disjoint from P_K and the configuration does not already have labels A, A^c, B or B^c.
- x (resp y) if some spine $\Sigma_{(A,K)}$ or $\Sigma_{(B,K)}$ lies entirely in Y_K (resp X_K) and the configuration does not already have labels A, A^c, B or B^c.

We will use the superscript * to denote the possible presence of superscript ^{c}, for example we will use A^* if there is a label A, A^c or both.

Remark 8.1 *If all curves of $P_K \cap Q_K$ are mutually essential, then a curve is essential in Q_K^A say, only if it is essential in Q_K so any c–disk for Q_K^A or Q_K^B that is disjoint from Q_K is in fact a c–disk for Q_K.*

Lemma 8.2 *If the configuration of P_K and Q_K has no labels, then $d(K,P) \leq 2-\chi(Q_K)$.*

Proof If $P_K \cap Q_K = \varnothing$ then $Q_K \subset A_K$ say, so B_K is entirely contained in X_K or in Y_K, say in Y_K. But B_K contains all spines $\Sigma_{(B,K)}$ so there will be a label x contradicting the hypothesis. Thus $P_K \cap Q_K \neq \varnothing$.

Consider the curves $P_K \cap Q_K$ and suppose some are essential in P_K but inessential in Q_K. An innermost such curve in Q_K will bound a c–disk in A_K or B_K. Since there is no label, such curves can not exist. In particular, any intersection curve that is inessential in Q_K is inessential in P_K. Now suppose there is a curve of intersection that is inessential in P_K. An innermost such curve c bounds a possibly punctured disk $D^* \subset P_K$ that lies either in X_K or in Y_K but, because there is no label X^* or Y^*, this curve must be inessential in Q_K as well. Let E be the possibly punctured disk it bounds there. We have just seen that all intersections of E with P_K must be inessential in both surfaces, so c is removable and would have been removed at the onset. We conclude that all remaining curves of intersection are essential in both surfaces. As there are no labels X^* or Y^*, Q_K^A and Q_K^B are c–incompressible. We conclude that both surfaces satisfy the hypothesis of Proposition 4.3. The bound on the distance then follows by Corollary 5.3. □

Proposition 8.3 *If some configuration is labeled A^* and B^* then P_K is c–strongly compressible.*

Proof The labels imply the presence of c–disks for P_K that we will denote by D_A^* and D_B^* such that $\partial D_A^*, \partial D_B^* \in Q_K \cap P_K$. As Q_K is embedded, either $\partial D_A^* = \partial D_B^*$ or $\partial D_A^* \cap \partial D_B^* = \varnothing$. Thus P_K is c–strongly compressible. □

Lemma 8.4 *If $P_K \cap Q_K = \varnothing$ with say $P_K \subset X_K$ (recall that X_K may be a handlebody, a compression body or a K–handlebody) and $Q_K \subset A_K$, then either every compressing disk D for Q_K lying in X_K intersects P_K or at least one of P_K and Q_K is strongly compressible.*

Proof Suppose P_K and Q_K are both weakly incompressible and that there is a compressing disk for Q_K lying in $X_K \cap A_K$. As $Y_K \subset A_K$ this implies that Q_K is bicompressible in A_K. As Q_K is weakly incompressible in M_K, it must be weakly incompressible in A_K. Compress Q_K maximally in $A_K \cap X_K$ to obtain a surface Q_K^X incompressible in A_K by Corollary 6.3. Consider the compressing disks for P_K lying in A_K. Each of them can be made disjoint from Q_K^X by an innermost disk argument so the surface P_K^A obtained by maximally compressing P_K in A_K is disjoint from Q_K^X and so from Q_K (see Figure 11). As M has no boundary, P_K^A is a collection of spheres and of annuli parallel to $N(K)$. The surface P_K^A separates P_K and Q_K thus Q_K is entirely contained in a ball or in a ball punctured by the knot in one arc. This contradicts the assumption that if $M = S^3$, then K is at least a three bridge knot. □

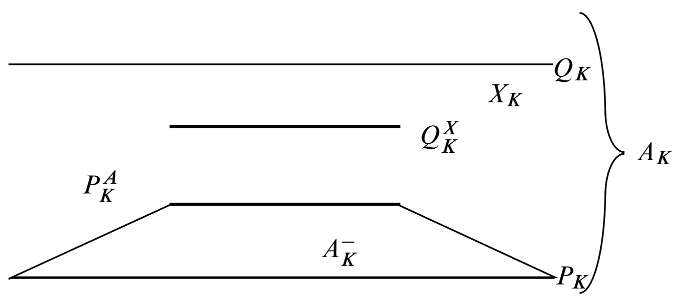


Figure 11

Lemma 8.5 *If there is a spine $\Sigma_{(A,K)} \subset Y_K$ (recall that Y_K may be a handlebody, a compression body or a K–handlebody) then either any c–disk for Q_K in Y_K that is disjoint from P_K intersects $\Sigma_{(A,K)}$ or at least one of P_K and Q_K is c–strongly compressible.*

Proof Suppose P_K and Q_K are both c–weakly incompressible and suppose E is a c–disk for Q_K in Y_K that is disjoint from P_K and from some spine $\Sigma_{(A,K)}$. Use the product structure between P_K and $\Sigma_{(A,K)}$ to push all of Q_K^A, as well as E, into

B_K. If E was a compressing disk, this gives a contradiction to Lemma 8.4 with the roles of X_K and Y_K reversed. We want to show that even if the initial disk E was a cut-disk, after the push we can find a compressing disk for Q_K lying in Y_K that is disjoint from P_K and contradict Lemma 8.4.

Suppose E is a cut-disk, let $\kappa \in B$ be the arc of $K - P$ that pierces E and let $D \subset B_K$ be its bridge disk with respect to P_K. Isotope Q_K and D so that $|Q_K \cap D|$ is minimal and consider $b \subset Q_K \cap D$, the arc of intersection adjacent to E (this situation is similar to Figure 10). If b is a closed curve, let D_b be the disk it bounds in D. If $D \cap Q_K = b$ then D_b is a compressing disk for Q_K that intersects E in exactly one point, contradicting c–weak incompressibility. Let δ be an innermost curve of intersection between D and Q_K bounding a subdisk $D_\delta \subset D$. If $D_\delta \subset X_K$, that would contradict c–weak incompressibility of Q_K so $D_\delta \subset Y_K$ and is the desired compressing disk. If b is not a closed curve, we can obtain a compressing disk for Q_K much as in Figure 10. Both endpoints of b lie in κ as $Q_K \cap P_K = \varnothing$. If b is outermost, let R be the disk b cuts from D. A neighborhood of $R \cup E$ consists of two compressing disks for Q_K in Y_K both disjoint from P_K as desired. If b is not outermost, let δ be an outermost arc. Doubling the disk D_δ that δ cuts from D gives a compressing disk for Q_K. If this compressing disk is in X_K that would contradict c–weak incompressibility of Q_K thus the disk must lie in Y_K as desired. □

Of course the symmetric statements hold if $\Sigma_{(A,K)} \subset X_K$, $\Sigma_{(B,K)} \subset Y_K$ or $\Sigma_{(B,K)} \subset X_K$.

Lemma 8.6 *Suppose P_K and Q_K are both c–weakly incompressible surfaces. If there is a configuration labeled both x and Y^* (or symmetrically X^* and y) then either P_K and Q_K are K–isotopic or $d(K,P) \leq 2 - \chi(Q_K)$.*

Proof From the label x we may assume, with no loss of generality, that there exists a spine $\Sigma_{(A,K)} \subset Y_K$. From the label Y^* we know that Q_K has a c–disk in $Y_K - P_K$, call this disk E. By Lemma 8.5, $E \cap \Sigma_{(A,K)} \neq \varnothing$ so in particular $E \subset Y_K$.

We first argue that we may as well assume that all components of $P_K \cap Q_K$ are essential in P_K. For suppose not; let c be the boundary of an innermost possibly punctured disk D^* in $P_K - Q_K$. If c were essential in Q_K then D^* cannot be in Y_K (by Lemma 8.5) and so it would have to lie in X_K. But then D^* is disjoint from E, contradicting the c–weak incompressibility of Q_K. We deduce that c is inessential in Q_K bounding a possibly punctured subdisk $D' \subset Q_K$. If D' intersects P_K in any curves that are essential, that would result in a label A^* or B^* contradicting our labeling scheme so c is removable and should be been removed at the onset. Suppose now that some curve of intersection bounds a possibly punctured disk in Q_K. By the above it must be essential in P_K but then an innermost such curve would give rise to a label A^* or B^* contradicting the labeling scheme. Thus all curves of $Q_K \cap P_K$ are mutually essential.

Consider first Q_K^B. It is incompressible in B_K because a compression into Y_K would violate Lemma 8.5 and a compression into X_K would provide a c–weak compression of Q_K. If Q_K^B is not essential in B_K then every component of Q_K^B is parallel into P_K so in particular Q_K^B is disjoint from some spine $\Sigma_{(B,K)}$ and thus $Q_K \subset P_K \times I$. If Q_K is incompressible in $P_K \times I$, then it is P_K–parallel by Lemma 2.10 as we know that Q_K is not a sphere or an annulus. A compression for Q_K in $P_K \times I$ would contradict Lemma 8.5 unless both $\Sigma_{(A,K)}$ and $\Sigma_{(B,K)}$ are contained in Y_K and Q_K has a compressing disk D^X contained in $(P_K \times I) \cap X_K$. In this case, as each component of Q_K^B is P_K–parallel, we can isotope Q_K to lie entirely in A_K so that $P_K \subset Y_K$ but then the disk E provides a contradiction to Lemma 8.4. We conclude that Q_K^B is essential in B_K so by Proposition 4.3 for each component q of $Q_K \cap P_K$ that is not the boundary of a P_K–parallel annulus in B_K, the inequality $d(q, \mathcal{B}) \leq 1 - \chi(Q_K^B)$ holds. Thus we can conclude that either P_K and Q_K are K–isotopic or Q_K^B satisfies the hypotheses of Lemma 5.2.

By Lemma 8.5 Q_K^A does not have c–disks in $Y_K \cap (A_K - \Sigma_{(A,K)})$ so it either has no c–disks in $A_K - \Sigma_{(A,K)}$ at all or has a c–disk lying in X_K. The latter would imply that Q_K^A is actually c–bicompressible in A_K. In either case we will show that Q_K^A also satisfies the hypotheses in Lemma 5.2 and the conclusion of that lemma completes the proof.

Case 1 Q_K^A is incompressible in $A_K - \Sigma_{(A,K)} \cong P_K \times I$. By Lemma 2.10 each component of Q_K^A must be P_K–parallel. The c–disk E of Q_K^A in $Y_K - P_K$ can be extended via this parallelism to give a c–disk for P_K that is disjoint from all $q \in Q_K \cap P_K$. Hence $d(q, \mathcal{A}) \leq 2 \leq 1 - \chi(Q_K^A)$ as long as Q_K^A is not a collection of P_K–parallel annuli. If that is the case, then $d(\partial E, q_0) = 0$ for at least one $q_0 \in (P_K \cap Q_K)$ so $d(q_0, \mathcal{A}) \leq 1 \leq 1 - \chi(Q_K^A)$ as desired.

Case 2 Q_K^A is c–bicompressible in A_K. Every c–disk for Q_K in Y_K intersects $\Sigma_{(A,K)}$, so we can deduce the desired distance bound by Theorem 7.7. □

Lemma 8.7 *Suppose P_K and Q_K are both c–weakly incompressible surfaces. If there is a configuration labeled both X^* and Y^* then either P_K and Q_K are K–isotopic or $d(K, P) \leq 2 - \chi(Q_K)$.*

Proof Since Q_K is c–weakly incompressible, any pair of c–disks, one in X_K and one in Y_K, must intersect in their boundaries and so cannot be separated by P_K. It follows that if both labels X^* and Y^* appear, the boundaries of the associated c–disks lie in one of Q_K^A or Q_K^B, say, Q_K^A.

Again we may as well assume that all components of $P_K \cap Q_K$ are essential in P_K. For suppose not; let c be the boundary of an innermost possibly punctured disk D^* in $P_K - Q_K$. If c were essential in Q_K then a c–disk in B_K parallel to D would be a c–disk for Q_K^B. From this contradiction we deduce that c is inessential in Q_K and proceed as in the proof of Lemma 8.6. As no labels A^* or B^* appear, all curves are also essential in Q_K.

If Q_K^A or Q_K^B could be made disjoint from some spine $\Sigma_{(A,K)}$ or $\Sigma_{(B,K)}$, then the result would follow by Lemma 8.6 so we can assume that is not the case. In particular Q_K^B is essential and so via Proposition 4.3 it satisfies the hypothesis of Lemma 5.2. The surface Q_K^A is c–bicompressible, c–weakly incompressible and there is no spine $\Sigma_{(A,K)}$ disjoint from Q_K^A. By Theorem 7.7, Q_K^A also satisfies the hypothesis of Lemma 5.2 so we have the desired distance bound. □

Lemma 8.8 *Suppose P_K and Q_K are both c–weakly incompressible surfaces. If there is a configuration labeled both x and y, then either P_K and Q_K are K–isotopic or $d(K, P) \leq 2 - \chi(Q_K)$.*

Proof As usual, we can assume that all curves in $P_K \cap Q_K$ are essential in both surfaces. Indeed, if there is a curve of intersection that is inessential in P_K then an innermost one either is inessential also in Q_K, and can be removed as described above, or is essential in Q_K and so would give rise to a label X^* or Y^*, a case done in Lemma 8.6. In fact we may assume that Q_K^A or Q_K^B are incompressible and c–incompressible as otherwise the result would follow by Lemma 8.6. As no labels A^* or B^* appear, we can again assume that all curves $P_K \cap Q_K$ are also essential in Q_K.

Both X_K and Y_K contain entire spines of A_K or B_K, though since we are not dealing with fixed spines the labels could arise if there are two distinct spines of A_K, say, one in X_K and one in Y_K. Indeed that is the case to focus on, since if spines $\Sigma_{(A,K)} \subset X_K$ and $\Sigma_{(B,K)} \subset Y_K$ then Q_K is an incompressible surface in $P_K \times I$ so by Lemma 2.10 Q_K is K–isotopic to P_K.

So suppose that $\Sigma_{(A,K)} \subset Y_K$ and there is another spine $\Sigma'_{(A,K)} \subset X_K$. The surface Q_K^A is incompressible in A_K so it is certainly incompressible in the product $A_K - \Sigma_{(A,K)}$ and so every component of Q_K^A is parallel in $A_K - \Sigma_{(A,K)}$ to a subsurface of P_K. Similarly every component of Q_K^A is parallel in $A_K - \Sigma'_{(A,K)}$ to a subsurface of P_K.

Let Q_0 be a component of Q_K^A that lies between $\Sigma_{(A,K)}$ and $\Sigma'_{(A,K)}$. This implies that Q_0 is parallel into P_K on both its sides, ie that $A_K \cong Q_0 \times I$.

As K is not a 2–bridge knot, then either $\chi(P_K) < -2$ (so in particular $\chi(Q_0) < -1$) or P_K a twice punctured torus. We will show that in either case $d(\mathcal{A}, q) \leq 2 \leq 1 - \chi(Q_K^A)$.

If P_K is a twice punctured torus, then Q_0 is a once punctured annulus so has Euler characteristic -1 and thus $\chi(Q_K^A) < 0$. Note that $d(\partial Q_0, \mathcal{A}) \leq 2$ (see Figure 12) and thus $d(\mathcal{A}, q) \leq 2 \leq 1 - \chi(Q_K^A)$.

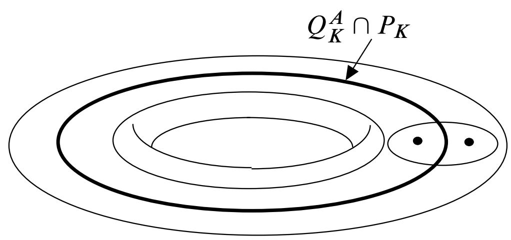


Figure 12

If $\chi(P_K) < -2$ let α be an essential arc in Q_0 with endpoints in $P_K \cap Q_K$. Then $\alpha \times I \subset Q_0 \times I \cong A_K$ is a meridian disk D for A_K that intersects Q_0 precisely in α. P–compressing Q_0 along one of the two disk components of $D - \alpha$ produces at most two surfaces at least one of which, Q_1 say, has a strictly negative Euler characteristic. In particular it is not a disk, punctured disk or an annulus. Every component of ∂Q_1 is essential in P_K and disjoint from both D and $Q_0 \cap P_K$. We can conclude that for every curve $q \in P_K \cap Q_K$, $d(\mathcal{A}, q) \leq d(\mathcal{A}, \partial Q_0) + d(\partial Q_0, q) \leq 2 \leq 1 - \chi(Q_K^A)$. Thus Q_K^A always satisfied the hypothesis of Lemma 5.2.

Now consider Q_K^B. If it is essential, then by Proposition 4.3 Q_K^B also satisfies the hypothesis of Lemma 5.2 and we are done by that lemma. If Q_K^B has c–disks in B_K, we have labels X^* and y (or x and Y^*) and we are done via Lemma 8.6. Finally, if each component of Q_K^B is parallel to a subsurface of P_K, then Q_K is disjoint from a spine $\Sigma_{(B,K)}$ as well, a case we have already considered. □

9 How labels change under isotopy

Suppose P and Q are as defined in the previous section and continue to assume that if P is a sphere, then P_K has at least six punctures. Consider how configurations and their labels change as P_K say is isotoped while keeping Q_K fixed. Clearly if there are no tangencies of P_K and Q_K during the isotopy then the curves $P_K \cap Q_K$ change only by isotopies and there is no change in labels. Similarly, if there is an index 0 tangency, $P_K \cap Q_K$ changes only by the addition or deletion of a removable curve. Since all such curves are removed before labels are defined, again there is no affect on the labeling. There are two cases to consider; P_K passing through a saddle tangency for Q_K and P_K passing through a puncture of Q_K. Consider first what can happen to the labeling when passing through a saddle tangency of P_K with Q_K.

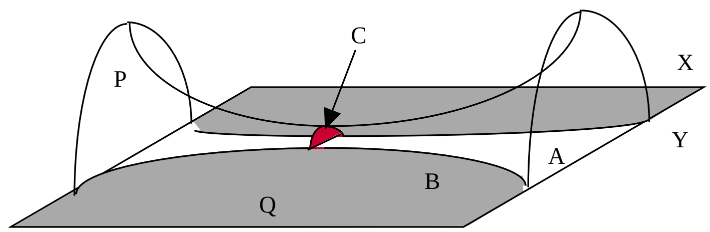


Figure 13

Lemma 9.1 *Suppose P_K and Q_K are c–weakly incompressible surfaces and P_K is isotoped to pass through a single saddle tangency for Q_K. Suppose farther that the bigon C defining the saddle tangency (see Figure 13) lies in $X_K \cap A_K$. Then*

- *no label x or X^* is removed,*
- *no label y or Y^* is created,*
- *if there is no label x or X^* before the move, but one is created after and if there is a label y or Y^* before the move and none after, then either P_K and Q_K are isotopic or $d(K,P) \leq 2-\chi(Q_K)$.*

Proof Much of the argument here parallels the argument in the proof of [12, Lemma 4.1]. The main difference is in Claim 2.

We first show that no label x or X^* is removed. If there is a c–disk for Q_K in $X_K \cap A_K$, a standard innermost disk, outermost arc argument on its intersection with C shows that there is a c–disk for Q_K in $X_K \cap A_K$ that is disjoint from C. The saddle move has no effect on such a disk. It is clear that the move doesn't have an effect on a c–disk for Q_K lying in $X_K \cap B_K$ so a label X^* will not be removed. If there is a spine of (A,K) or (B,K) lying entirely in Y_K then that spine too is unaffected by the saddle move.

Dually, no label y or Y^* is created: the inverse saddle move, restoring the original configuration, is via a bigon that lies in $B_K \cap Y_K$.

To prove the third item position Q_K so that it is exactly tangent to P_K at the saddle. A bicollar of Q_K then has ends that correspond to the position of Q_K just before the move and just after. Let Q_K^a denote $Q_K \cap A_K$ after the move and Q_K^b denote $Q_K \cap B_K$ before the move. The bicollar description shows that Q_K^a and Q_K^b have disjoint boundaries in P_K. Moreover the complement of $Q_K^a \cup Q_K^b$ in Q_K is a regular neighborhood of the singular component of $P_K \cap Q_K$, with Euler characteristic -1. It follows that $\chi(Q_K^a) + \chi(Q_K^b) = \chi(Q_K) + 1$.

With Q_K positioned as described, tangent to P_K at the saddle point but otherwise in general position, consider the closed (non-singular) curves of intersection.

Claim 1 It suffices to consider the case in which all non-singular curves of intersection are essential in P_K.

To prove the claim, suppose a non-singular curve is inessential and consider an innermost one. Assume first that the possibly punctured disk D^* that it bounds in P_K does not contain the singular curve s (ie the component of $P_K \cap Q_K$, homeomorphic to a figure 8, that contains the saddle point). If ∂D^* is essential in Q_K, then it would give rise to a label X^* or a label Y^* that persists from before the move until after the move, contradicting the hypothesis. Suppose on the other hand that ∂D^* is inessential in Q_K and so bounds a possibly punctured disk $E^* \subset Q_K$. All curves of intersection of E^* with P_K must be inessential in P_K, since there is no label A^* or B^*. It follows that $\partial D^* = \partial E^*$ is a removable component of intersection so the disk swap that replaces E^* with a copy of D^*, removing the curve of intersection (and perhaps more such curves) has no effect on the labeling of the configuration before or after the isotopy. So the original hypotheses are still satisfied for this new configuration of P_K and Q_K.

Suppose, on the other hand, that an innermost non-singular inessential curve in P_K bounds a possibly punctured disk D^* containing the singular component s. When the saddle is pushed through, the number of components in s switches from one s_0 to two $s_\pm$ or vice versa. All three curves are inessential in P_K since they lie in the punctured disk D^*. Two of them actually bound possibly punctured subdisks of D^* whose interiors are disjoint from Q_K. Neither of these curves can be essential in Q_K otherwise they determine a label X^* or Y^* that persist throughout the isotopy. At least one of these curves must bound a nonpunctured disk in P_K (as D^* has at most one puncture) and thus it also bounds a nonpunctured disk in Q_K. We conclude that at least two of the curves are inessential in Q_K and at least one of them bounds a disk in Q_K. As the three curves cobound a pair of pants of Q_K the third curve is also inessential in Q_K. This implies that all the curves are removable so passing through the singularity has no effect on the labeling. This proves the claim.

Claim 2 We may assume that if any of the curves $s_0, s_\pm$ are inessential in P_K they bound punctured disks in both surfaces.

The case in which all three curves are inessential in P_K is covered in the proof of Claim 1. If two are inessential in P_K and at least one of them bounds a disk with no punctures then the third curve is also inessential. Thus if exactly two curves are inessential in P_K, they both bound punctured disks in P_K and as no capital labels are preserved during the tangency move, they also bound punctured disks in Q_K which are parallel into P_K.

We are left to consider the case in which exactly one of $s_0, s_\pm$ is inessential in P_K, bounds a disk there and, following Claim 1, the disk it bounds in P_K is disjoint from Q_K. If the curve were essential in Q_K then there would have to be a label X or Y that occurs both before and after the saddle move, a contradiction. If the curve is inessential in Q_K then it is removable. If this removable curve is $s_\pm$ then passing through the saddle can have no effect on the labeling. If this removable curve is s_0 then the curves $s_\pm$ are parallel in both P_K and Q_K. In the latter case, passing through the saddle has the same effect on the labeling as passing an annulus component of $P_K - Q_K$ across a parallel annulus component Q_K^0 of Q_K^A. This move can have no effect on labels x or y. A meridian, possibly punctured disk E^* for Y_K that is disjoint from P_K would persist after this move, unless ∂E^* is in fact the core curve of the annulus Q_K^0. But then the union of E^* and half of Q_K^0 would be a possibly punctured meridian disk of A_K bounded by a component of $\partial Q_K^0 \subset P_K$. In other words, there would have to have been a label A^* before the move, a final contradiction establishing Claim 2.

Claims 1 and 2, together with the fact that neither labels A^* nor B^* appear, reduce us to the case in which all curves of intersection are essential in both surfaces both before and after the saddle move except perhaps some curves which bounds punctured disks in Q_K and in P_K. Let $\widetilde{Q_K^a}$ and $\widetilde{Q_K^b}$ be the surfaces left over after deleting from Q_K^a and Q_K^b any P_K–parallel punctured disks. As Q_K^a and Q_K^b cannot be made disjoint from any spine $\Sigma_{(A,K)}$ or $\Sigma_{(B,K)}$, $\widetilde{Q_K^a}$ and $\widetilde{Q_K^b}$ are not empty and, as we are removing only punctured disks, $\chi(Q_K^a) = \chi(\widetilde{Q_K^a})$ and $\chi(Q_K^b) = \chi(\widetilde{Q_K^b})$. Note then that $\widetilde{Q_K^a}$ and $\widetilde{Q_K^b}$ are c–incompressible in A_K and B_K respectively. For example, if the latter has a c–disk in B_K, then so does Q_K^a. Since no label X^* exists before the move, the c–disk must be in Y_K and such a c–compression would persist after the move and so then would the label Y^*. Similarly neither $\widetilde{Q_K^a}$ nor $\widetilde{Q_K^b}$ can consist only of P_K parallel components. For example, if all components of $\widetilde{Q_K^b}$ are parallel into P_K then Q_K^b is also disjoint from some spine of B_K and such a spine will be unaffected by the move, resulting in the same label (x or y) arising before and after the move. We deduce that $\widetilde{Q^a_K}$ and $\widetilde{Q^b_K}$ are essential surfaces in A_K and B_K respectively.

Now apply Proposition 4.3 to both sides. Let q_a (resp q_b) be a boundary component of an essential component of $\widetilde{Q^a_K}$ (resp $\widetilde{Q^b_K}$). Then

$$\begin{aligned} d(K,P) = d(\mathcal{A},\mathcal{B}) &\le d(q_a,\mathcal{A}) + d(q_a,q_b) + d(q_b,\mathcal{B}) \\ &\le 3 - \chi(\widetilde{Q^a_K}) - \chi(\widetilde{Q^b_K}) = \le 3 - \chi(Q^a_K) - \chi(Q^b_K) = 2 - \chi(Q_K) \end{aligned}$$

as required. □

It remains to consider the case when P_K passes through a puncture of Q_K as in Figure 14. This puncture defines a bigon C very similar to the tangency bigon in the previous lemma: let Q^a_K and Q^b_K be as before, then $Q_K - (Q^a_K \cup Q^b_K)$ is a punctured annulus. The knot strand that pierces it is parallel to this annulus, let C be the double of the parallelism rectangle so that $C \subset X_K \cap A_K$.

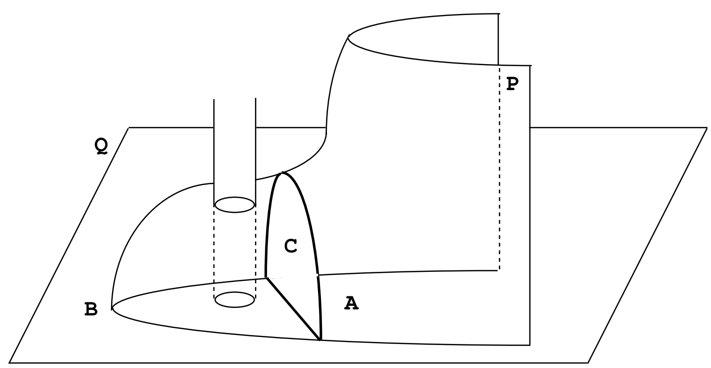


Figure 14

Lemma 9.2 *Suppose P_K and Q_K are c–weakly incompressible bridge surfaces for a knot K and P_K is isotoped to pass through a single puncture for Q_K. Suppose further that the bigon C defined by the puncture (see Figure 14) lies in $X_K \cap A_K$.*

- *No label x or X^* is removed.*
- *No label y or Y^* is created.*
- *Suppose that, among the labels both before and after the move, neither A^* nor B^* occur. If there is no label x or X^* before the move, but one is created after and if there is a label y or Y^* before the move and none after, then either P_K and Q_K are K–isotopic or $d(P,K) \le 2 - \chi(Q_K)$.*

Proof The proof is very similar to the proof of the previous lemma. It is clear that if there is a c–disk for X_K that lies in A_K, there is a c–disk that is disjoint from C and thus the label survives the move. If there is a spine of A_K or B_K lying entirely in Y_K then that spine, too, is unaffected by the saddle move. The proof of the third item is identical to the proof in the above lemma in the case when at least one of the curves $s_0, s_\pm$ bounds a punctured disk in Q_K. □

We will use X (resp Y) to denote any subset of the labels x, X, X^c (resp y, Y, Y^c). The results of the last two sections then can be summarized as follows

Corollary 9.3 *If two configurations are related by a single saddle move or going through a puncture and the union of all labels for both configurations contains both* X *and* Y *then either* P_K *and* Q_K *are* K*–isotopic or* $d(K, P_K) \leq 2 - \chi(Q_K)$

Proof With no loss of generality, the move is as described in Lemma 9.1 or Lemma 9.2. These lemmas show that either we have the desired bound or there is a single configuration for which both X and Y appear. The result then follows from one of Lemma 8.7, Lemma 8.8 or Lemma 8.6. □

We will also need the following easy Lemma.

Lemma 9.4 *If a configuration carries a label* A^* *before a saddle move or going through a puncture and a label* B^* *after then* P_K *is c–strongly compressible.*

Proof As already discussed the curves before and after the saddle move are distance at most one in the curve complex of P_K. □

10 Main result

Given a bridge surface for a link K there are three ways to create new, more complex, bridge surfaces for the link: adding dual one-handles disjoint from the knot (stabilizing), adding dual one-handles where one of them has an arc of K as its core (meridionally stabilizing), and introducing a pair of a canceling minimum and maximum for K (perturbing). These are depicted in Figure 15, the precise definitions can be found in Tomova [13] .

Definition 10.1 *Let* P *and* Q *be two bridge surfaces for a knot* $K \subset M$. *We say that* Q *is equivalent to* P *if* Q *is* K*–isotopic to a copy of* P *which may have been stabilized, meridionally stabilized and perturbed.*

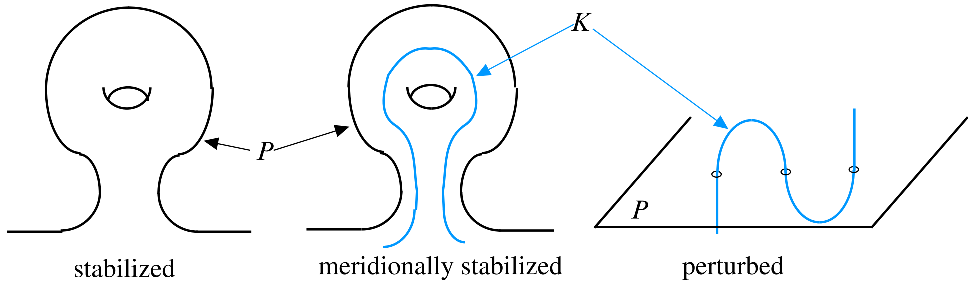


Figure 15

There is a fourth way to construct a bridge surface for a knot K. Suppose Q is a Heegaard splitting for M splitting it into handlebodies X and Y and suppose K is isotopic to a subset of the spine for X. Then by introducing a single minimum, K can be placed in bridge position with respect to Q. In this case K is said to be removable as Q is also a Heegaard surface for M_K after an isotopy of K. Scharlemann and Tomova discuss all four of these operations in detail in [11].

Casson and Gordon have demonstrated that if a 3–manifold has a Heegaard splitting which is irreducible but strongly compressible then the manifold contains an essential surface. In [13], Tomova extended this result to prove the following Theorem.

Theorem 10.2 *Suppose M is a closed orientable irreducible 3–manifold containing a link K. If Q is a c–strongly compressible bridge surface for K then either*

- *Q is stabilized,*
- *Q is meridionally stabilized,*
- *Q is perturbed,*
- *K is removable with respect to Q or*
- *M_K contains a meridional essential surface F_K such that $2-\chi(F_K) \leq 2-\chi(Q_K)$.*

We can now prove the main result of this paper.

Theorem 10.3 *Suppose K is a nontrivial knot in a closed, irreducible and orientable 3–manifold M and P is a bridge surface for K. If P is a sphere assume that $|P\cap K| \geq 6$. If Q is also a bridge surface for K that is not equivalent to P, or if Q is a Heegaard surface for $M-\eta(K)$ then $d(K,P) \leq 2-\chi(Q-K)$.*

Proof If Q_K is stabilized, meridionally stabilized or perturbed we can perform the necessary compressions to undo these operations as described by Scharlemann and Tomova in [11]. Note that these operations increase $\chi(Q_K)$ so we may assume that Q_K is not stabilized, meridionally stabilized or perturbed. If K is removable with respect to Q, we may assume that K has been isotoped to lie in the spine of one of the handlebodies $M - Q$ so Q is a Heegaard splitting for M_K. This operation decreases $|Q \cap K|$ and thus also increases $\chi(Q_K)$.

Suppose first that Q_K is c–strongly compressible. If K is not removable with respect to Q, by Theorem 10.2, there is an essential surface F_K such that $2-\chi(F_K) < 2-\chi(Q_K)$. If Q is a Heegaard surface for M_K, the existence of such an essential surface follows by Casson and Gordon, [3]. Then the result follows from Theorem 5.7. If P_K is c–strongly compressible, then $d(P, K) \leq 3$ by applying Proposition 4.1 twice. Thus we may assume that both P_K and Q_K are c–weakly incompressible.

The proof now is almost identical to the proof of the main result in [12] so we will only give a brief summary.

Recall that if $\Sigma_{(A,K)}$ is a spine for the K–handlebody A_K, then $A - \Sigma_{(A,K)} \cong P_K \times I$. Thus if P is a bridge surface for K, there is a map $H\colon (P, P \cap K) \times I \to (M, K)$ that is a homeomorphism except over $\Sigma_{(A,K)} \cup \Sigma_{(B,K)}$ and near $P \times \partial I$ the map gives a mapping cylinder structure to $\Sigma_{(A,K)} \cup \Sigma_{(B,K)}$. If we restrict H to $P_K \times (I, \partial I) \to (M, \Sigma_{(A,K)} \cup \Sigma_{(B,K)})$, H is called a sweep-out associated to P.

If Q is a Heegaard surface for M_K, splitting M_K into a compression body and a handlebody, then a similar sweep-out is associated to Q between the two spines. We will denote these spines by Σ_X and Σ_Y.

Consider a square $I \times I$ that describes generic sweep-outs of P_K and Q_K from $\Sigma_{(A,K)}$ to $\Sigma_{(B,K)}$ and from $\Sigma_{(X,K)}$ to $\Sigma_{(Y,K)}$ if Q is a bridge surface for K or from Σ_X to Σ_Y if K is removable with respect to Q. See Figure 16. Each point in the square represents a positioning of P_K and Q_K. Inside the square is a graph Γ, called the *graphic* that represents points at which the intersection is not generic. At each point in an edge in the graphic there is a single point of tangency between P_K and Q_K or one of the surfaces is passing through a puncture of the other. At each (valence four) vertex of Γ there are two points of tangency or puncture crossings. By general position of, say, the spine $\Sigma_{(A,K)}$ with the surface Q_K the graphic Γ is incident to $\partial I \times I$ in only a finite number of points (corresponding to tangencies between $\Sigma_{(A,K)}$ and Q_K). Each such point in $\partial I \times I$ is incident to at most one edge of Γ.

Any point in the complement of Γ represents a generic intersection of P_K and Q_K. Each component of the graphic complement will be called a *region*; any two points in the same region represent isotopic configurations. Label each region with labels A, B, X and Y as described previously where a region is labeled X (resp Y) if any of the labels x, X, X^c (resp y, Y, Y^c) appear and A (resp B) if the labels A or A^c (resp B or B^*) appear. See Figure 16. If any region is unlabeled we are done by Lemma 8.2. Also if a region is labeled X and Y we are done by one of Lemma 8.7, Lemma 8.8 or Lemma 8.6. Finally by Proposition 8.3 no region is labeled both A and B so we can assume that each region of the square has a unique label.

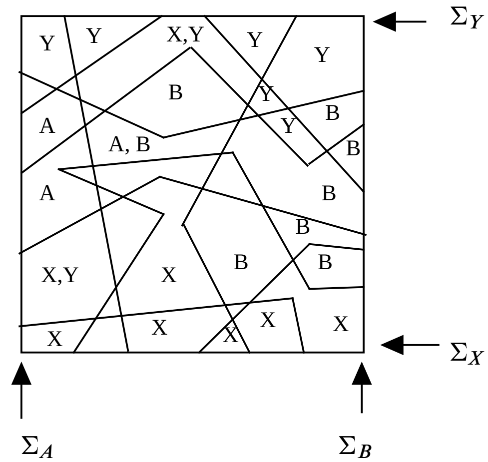


Figure 16

Let Λ be the dual complex of Γ in $I \times I$; Λ has one vertex in each face of Γ and one vertex in each component of $\partial I \times I - \Gamma$. Each edge of Λ not incident to $\partial I \times I$ crosses exactly one interior edge of Γ. Each face of Λ is a quadrilateral and each vertex inherits the label of the corresponding region of Γ. Consider the labeling of two adjacent vertices of Λ. Corollary 9.3 says that if they are labeled X and Y we have the desired result and Lemma 9.4 says they cannot be labeled A and B. Finally, a discussion identical to the one in [12] about labeling along the edges of $I \times I$ shows that no label B appears along the $\Sigma_{(A,K)}$ side of $I \times I$ (the left side in the figure), no label A appears along the $\Sigma_{(B,K)}$ side (the right side), no label Y appears along the $\Sigma_{(X,K)}$ side (Σ_X side if Q is a bridge surface for M_K) (the bottom) and no label X appears along the $\Sigma_{(Y,K)}$ side (Σ_Y side if Q is a bridge surface for M_K) (the top).

We now appeal to the following quadrilateral variant of Sperner's Lemma proven in the appendix of [12].

Lemma 10.4 *Suppose a square $I \times I$ is tiled by quadrilaterals so that any two that are incident meet either in a corner of each or in an entire side of each. Let Λ denote the graph in $I \times I$ that is the union of all edges of the quadrilaterals. Suppose each vertex of Λ is labeled N, E, S, or W in such a way that*

- *no vertex on the East side of $I \times I$ is labeled W, no vertex on the West side is labeled E, no vertex on the South side is labeled N and no vertex on the North side is labeled S and*
- *no edge in Λ has ends labeled E and W nor ends labeled N and S,*

then some quadrilateral contains all four labels.

In our context the lemma says that there are four regions in the graphic incident to the same vertex of Γ labeled $\mathrm{A}, \mathrm{B}, \mathrm{X}$ and Y. Note then that only two saddle or puncture moves are needed to move from a configuration labeled A to one labeled B. The former configuration includes a c–disk for P_K in A and the latter a c–disk for P_K in B. Note that as K is nontrivial $\chi(Q_K) \le -2$. Using Proposition 4.1 it follows that $d(K, P) \le 4 \le 2 - \chi(Q_K)$, as long as at least one of the regions labeled X and Y contains at least one essential curve.

Suppose all curves of $P \cap Q$ in the regions X and Y are inessential. Consider the region labeled X. Crossing the edge in the graphic from this region to the region labeled A corresponds to attaching a band b_A with both endpoints in an inessential curve $c \in P \cap Q$ or with endpoint in two distinct curves c_1 and c_2 where c_1 and c_2 both bound once punctured disks in P_K. Note that attaching this band must produce an essential curve that gives rise to the label A, call this curve c_A. Similarly crossing the edge from the region X into the region B corresponds to attaching a band b_B to give a curve c_B. The two bands have disjoint interiors and must have at least one endpoint in a common curve otherwise c_A and c_B would be disjoint curves giving rise to labels A and B. By our hypothesis attaching both bands simultaneously results in an inessential curve c_{AB}. We will show that in all cases we can construct an essential curve γ in P_K that is disjoint from c_A and c_B. After possibly applying Proposition 4.1, this implies that $d(K, P) \le 4$.

Case 1 Both bands have both of their endpoints in the same curve c.

Attaching b_A to c produces two curves that cobound a possibly once punctured annulus, one of these curves is c_A. We will say that the band is *essential* if c_A is essential in the closed surface P and *inessential* otherwise. If b_A and b_B are both essential but c_{AB} is inessential in P, then P is a torus so P_K is a torus with at least two punctures.

In this case $c_A \cup c_B$ doesn't separate the torus so we can consider the curve γ that bounds a disk in P containing at least two punctures of P_K.

If b_A is essential but b_B isn't, then c_{AB} is parallel to c_A in P and thus must be essential also so this case cannot occur.

Finally if both b_A and b_B are inessential in P and P is not a sphere, then let γ be an essential curve in P that is disjoint from $c_A \cup c_B$. If P is a sphere, it must have at least six punctures. Note that $c \cup b_A \cup b_B$ separates P into four regions that may contain punctures. As P has at least six punctures, one of these regions contains at least two punctures. Take γ to be a curve that bounds a disk containing two punctures and that is disjoint from $c \cup b_A \cup b_B$.

Case 2 One band, say b_A has endpoint lying in two different curves c_1 and c_2 and the other band, b_B has both endpoints lying in c_1.

If b_B is essential in P, then adding both bands simultaneously results in a curve that is parallel to c_B in P and therefore is essential contradicting the hypothesis. If b_B is inessential in P, then $c_1 \cup c_2 \cup b_A \cup b_B$ separates P into four regions that may contain punctures. As in the previous case we can construct an essential curve γ in P_K that is disjoint from c_A and c_B either by taking a curve essential in P or, if P is a sphere, by taking a curve that lies in one of the four regions and bounds two punctures on one side.

Case 3 The band b_A has endpoint lying in two different curves c_1 and c_2 and b_B has endpoint lying in c_1 and c_2', possibly $c_2 = c_2'$.

In this case c_A and c_B are both inessential in P so if P is not a sphere we can again find a curve γ disjoint from both that is essential in P. If P is a sphere, then $c_1 \cup c_2 \cup c_2' \cup b_A \cup b_B$ separates P into four regions that may contain punctures and so we can find a curve γ that is essential in P_K and disjoint from c_A and c_B as above. $\square$

The curve complex for a 4–times punctured sphere is not connected so a bound on the distance of a minimal bridge surface for a 2–bridge knot cannot be obtained. However Scharlemann and Tomova have proven the following uniqueness result.

Theorem 10.5 [11, Corollary 4.4] *Suppose K is a knot in S^3, 2–bridge with respect to the bridge surface $P \cong S^2$, and K is not the unknot. Suppose Q is any other bridge surface for K. Then either*

- *Q is stabilized,*
- *Q is meridionally stabilized,*

- *Q is perturbed or*
- *Q is properly isotopic to P.*

Corollary 10.6 *Suppose P and Q are two bridge surfaces for a knot K and K is not removable with respect to Q. Then either Q is equivalent to P or $d(P) \leq 2-\chi(Q_K)$.*

Proof If K is a two bridge knot with respect to a sphere P, then by Theorem 10.5, Q is equivalent to P. If P is not a four times punctured sphere, the result follows from Theorem 10.3. □

Corollary 10.7 *If $K \subset M^3$ is in bridge position with respect to a Heegaard surface P such that $d(K,P) > 2-\chi(P_K)$ then K has a unique minimal bridge position.*

Proof Suppose K can also be placed in bridge position with respect to a second Heegaard surface Q such that Q is not equivalent to P. By Theorem 10.3, $d(K,P) \leq 2-\chi(Q_K) = 2-\chi(P_K)$ contradicting the hypothesis. □

Acknowledgments

I would like to thank Martin Scharlemann for many helpful conversations. This research was partially supported by an NSF grant.

References

[1] **D Bachman**, *Thin position with respect to a heegaard surface*

[2] **D Bachman**, **S Schleimer**, *Distance and bridge position*, Pacific J. Math. 219 (2005) 221–235 MR2175113

[3] **A J Casson**, **C M Gordon**, *Reducing Heegaard splittings*, Topology Appl. 27 (1987) 275–283 MR918537

[4] **K Hartshorn**, *Heegaard splittings of Haken manifolds have bounded distance*, Pacific J. Math. 204 (2002) 61–75 MR1905192

[5] **J Hempel**, *3–manifolds as viewed from the curve complex*, Topology 40 (2001) 631–657 MR1838999

[6] **H Rubinstein**, **M Scharlemann**, *Comparing Heegaard splittings of non-Haken 3–manifolds*, Topology 35 (1996) 1005–1026 MR1404921

[7] **T Saito**, **M Scharlemann**, **J Schultens**, *Lecture notes on generalized heegaard splittings* `arXiv:math.GT/0504167`

[8] **M Scharlemann**, *Proximity in the curve complex* arXiv:math.GT/0410278

[9] **M Scharlemann**, *Local detection of strongly irreducible Heegaard splittings*, Topology Appl. 90 (1998) 135–147 MR1648310

[10] **M Scharlemann**, **A Thompson**, *Thinning genus two Heegaard spines in* S^3, J. Knot Theory Ramifications 12 (2003) 683–708 MR1999638

[11] **M Scharlemann**, **M Tomova**, *Uniqueness of bridge surfaces for 2–bridge knots* arXiv:math.GT/0609567

[12] **M Scharlemann**, **M Tomova**, *Alternate Heegaard genus bounds distance*, Geom. Topol. 10 (2006) 593–617

[13] **M Tomova**, *Thin position for knots in a 3–manifold* arXiv:math.GT/0609674

Mathematics Department, Rice University
6100 S Main Street, Houston TX 77005-1892, USA

mt2@rice.edu

Received: 5 April 2007