ISSN 1364-0380 **243**

Geometry & Topology
Volume 4 (2000) 243–275
Published: 3 October 2000

Levelling an unknotting tunnel

Hiroshi Goda
Martin Scharlemann
Abigail Thompson

Graduate School of Science and Technology, Kobe University
Rokko, Kobe 657-8501, Japan

Mathematics Department, University of California
Santa Barbara, CA 93106, USA

Mathematics Department, University of California
Davis, CA 95616, USA

Email: goda@math.kobe-u.ac.jp, mgscharl@math.ucsb.edu
thompson@math.ucdavis.edu

Abstract

It is a consequence of theorems of Gordon–Reid [4] and Thompson [8] that a tunnel number one knot, if put in thin position, will also be in bridge position. We show that in such a thin presentation, the tunnel can be made level so that it lies in a level sphere. This settles a question raised by Morimoto [6], who showed that the (now known) classification of unknotting tunnels for 2–bridge knots would follow quickly if it were known that any unknotting tunnel can be made level.

AMS Classification numbers Primary: 57M25

Secondary: 57M27

Keywords: Tunnel, unknotting tunnel, bridge position, thin position, Heegaard splitting

Proposed: Cameron Gordon | Received: 17 January 2000
Seconded: Robion Kirby, David Gabai | Accepted: 18 September 2000

1 Background

In [2], Gabai introduced the notion of *thin position* for a knot in S^3. Choose a height function $h\colon S^3 - \{x, y\} = S^2 \times \mathbb{R} \to \mathbb{R}$, the projection to the second factor. Informally, a knot K is in thin position with respect to h if the number of points in which it intersects the level spheres of h has been minimized. More formally, let $S(t) = h^{-1}(t)$ and call it the sphere at *height* or *level* t. Put K in general position with respect to h and choose heights $k_1 < k_2 < \ldots < k_n$ between each successive pair of critical heights of K. Define the *width* $W(K)$ of this imbedding of K in S^3 to be the integer $\Sigma_i |S(k_i) \cap K|$.

Definition 1.1 $K \subset S^3$ is in *thin position* (with respect to the height function h) if $W(K)$ cannot be reduced by an isotopy of K in S^3.

Definition 1.2 $K \subset S^3$ is in *bridge position* (with respect to the height function h) if it is in general position with respect to h and all the minima of K occur below all the maxima. A *minimal bridge position* is a bridge position which minimizes the number of minima and maxima. A level sphere lying between the minima and maxima is called a *middle sphere*.

Two bridge positions of a knot K are regarded as equivalent if they are isotopic via an isotopy throughout which the knot remains in bridge position. A knot may have more than one minimal bridge position (see for example, Figure 13). It is a tedious but worthy exercise to show that if there is an isotopy between two knots that are in bridge position and have the same middle sphere, and during the isotopy the knot remains always transverse to the middle sphere, then the bridge positions are equivalent. In particular, a sphere that divides a knot K into two untangles, well defines a bridge position of the knot in which the sphere is a middle sphere.

Notation $\eta(X)$ means a regular neighborhood of X in Y for polyhedral spaces X and Y. For K a knot in the 3–sphere, put $E(K) = S^3 - \text{int}\eta(K)$.

Placing a knot in thin position can sometimes reveal essential meridional planar surfaces in the complement, ie, planar surfaces which are incompressible, have boundary a collection of meridians of K, and which are not just annuli parallel into the knot. For example, there is this result of Thompson:

Theorem 1.3 [8] *If $E(K)$ does not have any essential meridional planar surface, then any thin position of K is a minimal bridge position of K and vice versa.*

In this paper we will be interested only in knots K which have tunnel number 1. That is, knots K for which there is an arc $\gamma \subset S^3$ such that $K \cap \gamma = \partial\gamma$ and $S^3 - \mathrm{int}\eta(K \cup \gamma)$ is a genus 2 handlebody. The arc γ is called an *unknotting tunnel* or *tunnel* for K.

Gordon and Reid have shown:

Theorem 1.4 [4] *Let $K \subset S^3$ have tunnel number 1. Then $E(K)$ does not have an essential meridional planar surface.*

Combining these results gives the obvious corollary:

Corollary 1.5 *Suppose K has tunnel number 1. Then any thin position of K is a minimal bridge position of K and vice versa.*

It is easy to see that all 2–bridge knots have tunnel number one. Moreover, Kobayashi has shown that the only tunnels possible are the obvious ones:

Theorem 1.6 [5] *Any unknotting tunnel for a 2–bridge knot is one of 6 known types.*

These six types have the property that they can each be made level with respect to the height function that gives the knot its 2–bridge structure. That is, each can be put into a level sphere. Morimoto had earlier pointed out that this classification would follow rather quickly if we knew that each tunnel could be made level:

Theorem 1.7 [6] *If an unknotting tunnel for a 2–bridge knot can be put into a level sphere, then it is one of the 6 known types.*

So it is natural to turn Morimoto's theorem into a question: could it be true that any unknotting tunnel for any tunnel number 1 knot K can be put into a level sphere, once K is put in minimal bridge position? Here we allow the tunnel to be "slid" as well as isotoped. That is, we allow the ends of the tunnel to be moved around on K, eg, possibly past each other on K, or, indeed, one end of the tunnel may be moved up onto the other end of the tunnel, changing what was an edge into a 1–complex we will call an "eyeglass": an edge attached to a circle at one of its ends. Of course the reverse move is also allowed.

In this paper we show that the answer to Morimoto's question is yes:

Theorem 1.8 *If $K \subset S^3$ is a tunnel number one knot in minimal bridge position and γ is a tunnel for K, then γ may be slid and isotoped to lie entirely in a level sphere for K.*

Moreover, in Section 6, we characterize the position of γ on a minimal bridge sphere.

This conclusion allows us to say a bit about the bridge structure of tunnel number 1 knots.

For example, [3] gives a detailed description of tunnel number 1 links with the property that one component is the unknot. The authors of [3] ask [Problem 6.4] for an explicit description of tunnel number one links both of whose components are knotted. In [1] Berge constructed some examples by edge sliding, beginning with the case of an unknotted component.

The theorem here assures us that such links can all be obtained either in this way, or by beginning with a level unknotting tunnel for a tunnel number one knot and sliding the ends together by a (possibly complicated) path on the boundary of the knot neighborhood. As a kind of generalization of these ideas we get the following way of characterizing a tunnel number one knot as the result of a series of bridge-increasing slides:

Lemma 1.9 *Suppose γ is a tunnel for a tunnel number 1 knot K. Let m be a meridian of γ on $\partial\eta(K \cup \gamma)$. Suppose K' is a simple closed curve on $\partial\eta(K \cup \gamma)$ which intersects m transversely once. Then K' is a tunnel number 1 knot.*

Proof Slide the ends of γ around to "peel off" K'. □

Definition 1.10 We then say that K' is obtained by a *tunnel move* from K.

Once we have established that such tunnels can always be made level, we have:

Proposition 1.11 *Let K be a tunnel number 1 knot. Then there exists a sequence of knots beginning with the unknot $= K_0, K_1, K_2, \ldots, K_n = K$ of bridge number $b_0 = 1, b_1, \ldots, b_n$ such that*

(a) *K_i, $i = 1, \ldots, n$, is a tunnel number one knot,*

(b) *K_i is obtained from K_{i-1} by a tunnel move,*

(c) $b_i \geq 2b_{i-1} - 1$.

Proof By induction on bridge number, using the fact that tunnels can always be made level by slides. □

The core of these results extends easily to tunnel number one links and there is a brief discussion of this in the last section.

2 Thin position and unknotting tunnels

Let K be a tunnel number 1 knot and γ be an unknotting tunnel for K. It will be convenient to think about the tunnel γ in two different ways. Initially it is an arc with each of its ends on K. But the thinning process we will describe below may force the ends of the arc together on K until they are incident to the same point on K (so that γ becomes a loop) and then resolve the resulting 4–valent vertex into two 3–valent vertices by pinching the ends of the tunnel together into a single arc. (This process could also be described as sliding one end of the tunnel onto a neighborhood of the other end.) The result is to change γ into the union of an arc and a circle; as noted above, we will call this graph an *eyeglass* and continue to denote it by γ. No matter whether γ denotes an edge or an eyeglass, the neighborhood $\eta(K \cup \gamma)$ is a genus 2 handlebody V_1, and its closed exterior $E(K \cup \gamma)$ is also a genus 2 handlebody V_2. Let F denote the genus 2 surface that is their common boundary, so F is a Heegaard splitting surface for $E(K)$.

In this section we will extend Gabai's notion of thin position (see Definition 1.1) to the graph $K \cup \gamma$. As before, choose a height function $h\colon S^3 - \{x, y\} = S^2 \times \mathbb{R} \to \mathbb{R}$ and let $S(t) = h^{-1}(t)$.

Definition 2.1 $K \cup \gamma$ is in *Morse position* with respect to h if

(a) the critical points of $h|K$ or $h|\gamma$ are nondegenerate and those in γ lie in the interior of γ,

(b) the critical points of $h|K$, $h|\gamma$ and the two vertices in $K \cup \gamma$ all occur at different heights.

The heights at which there is a critical point of either $h|K$ or $h|\gamma$ or there is a vertex of $K \cup \gamma$ are called the *critical heights* for $K \cup \gamma$. Near its critical height, each vertex v can be classified into one of four types:

(1) all ends of incident edges lie below v,

(2) all ends of incident edges lie above v,

(3) exactly two ends of incident edges lie above v,

(4) exactly two ends of incident edges lie below v.

We will further simplify the local picture by isotoping a small neighborhood of a vertex of type 1 (respectively type 2), transforming it into a vertex of type 4 (respectively type 3) and a nearby maximum (respectively minimum). The end result is called *normal form* for $K \cup \gamma$.

Definition 2.2 A vertex of type 3 is called a *Y–vertex* and a vertex of type 4 is called a *λ–vertex*. When we refer to the *maxima* of $K \cup \gamma$ we will include all local maxima of K, all local maxima of $\mathrm{int}(\gamma)$, and all λ–vertices. Similarly, by the *minima* of $K \cup \gamma$ we mean all local minima of K, all local minima of $\mathrm{int}(\gamma)$, and all Y–vertices. A maximum (respectively minimum) that is not a λ–vertex (respectively Y–vertex) will be called a *regular* maximum (respectively minimum). The union of the maxima and minima (hence including the vertices) are called the *critical points* of $K \cup \gamma$ and their heights the *critical values* or *critical heights.*

Definition 2.3 Extending and modifying the idea of knot width $W(K)$ (see Definition 1.1), let $t_0 < \ldots < t_n$ be the successive critical heights of $K \cup \gamma$ and suppose t_j and t_k are the two levels at which the vertices occur. Let $s_i, 1 \leq i \leq n$ be generic levels chosen so that $t_{i-1} < s_i < t_i$. Define

$$W(K \cup \gamma) = 2(\Sigma_{i \neq j,k} |S(s_i) \cap (K \cup \gamma)|) + |S(s_j) \cap (K \cup \gamma)| + |S(s_k) \cap (K \cup \gamma)|.$$

Definition 2.4 For K in minimal bridge (hence thin) position, a *thin position* of a pair (K, γ) (rel K) is a position which minimizes the width of (K, γ) (without changing the bridge presentation of K).

Remark 2.5 Of course, we can combine the two types of thinness: Define the width $W(K, \gamma)$ of (K, γ) to be the pair of integers $(W(K), W(K \cup \gamma))$ in lexicographical order and define a thin presentation of (K, γ) to be one that minimizes $W(K, \gamma)$. This guarantees that the knot K will be thin. Note that there is no way of predicting, *prima facie*, whether in a thin presentation γ will be an edge or an eyeglass.

Remark 2.6 There is nothing magical about the description of width above; thinning with respect to it serves the ad hoc purpose of pushing maxima below minima, while remaining indifferent to pushing maxima past maxima and minima past minima. See the next section. Other definitions that serve the same purpose are possible. For example, here's an alternative which might strike some as more natural: make V_1 very thin and count components of intersection of level spheres with V_1 between successive critical points of h on ∂V_1. Thus two level spheres are used near a regular minimum or maximum of $K \cup \gamma$ but only one near a minimum or maximum that occurs at a vertex. This definition of width is not equivalent to ours, but has the same general properties that we need, as described in the next section.

Let D be a meridian disk of $V_2 = S^3 - \mathrm{int}(\eta(K \cup \gamma))$.

Definition 2.7 The disk $(D, \partial D) \subset (V_2, \partial V_2)$ is in *normal form* if

(a) ∂D intersects each meridian of each edge of the graph $K \cup \gamma$ minimally, up to isotopy in $\partial V_2 = \partial\eta(K \cup \gamma)$,
(b) each critical point of h on D is nondegenerate,
(c) no critical point of h on $\mathrm{int}(D)$ occurs at a critical height of $K \cup \gamma$,
(d) no two critical points of h on $\mathrm{int}(D)$ occur at the same height,
(e) the minima (respectively maxima) of $h|\partial D$ at the minima (respectively maxima) of $K \cup \gamma$ are also local extrema of h on D, ie, 'half-center' singularities,
(f) the maxima of $h|\partial D$ at Y–vertices and the minima of $h|\partial D$ at λ–vertices are, on the contrary, 'half-saddle' singularities of h on D.

Standard Morse theory ensures that, for $K \cup \gamma$ in normal form, any properly imbedded essential disk $(D, \partial D) \subset (V_2, \partial V_2)$ can be put in normal form.

Definition 2.8 $K \cup \gamma$ is in *bridge position* if there is no minimum of $K \cup \gamma$ at a height above a maximum.

Definition 2.9 Let S be a generic level 2–sphere and B_u and B_l denote the balls which are the closures of the region above S and below S respectively. An *upper disk* (respectively *lower disk*) for S is a disk $F \subset V_2$ transverse to S such that $\partial F = \alpha \cup \beta$, where $\alpha = \partial F \cap S$ is an arc properly imbedded in $S - \eta(K \cup \gamma)$, β is an arc imbedded on $\partial\eta(K \cup \gamma)$, $\partial\alpha = \partial\beta$ and a small product neighborhood of ∂F in F lies in B_u (respectively B_l) ie, it lies *above* (respectively *below*) S.

Note that $\mathrm{int}(F)$ may intersect S in simple closed curves.

3 Moves that thin and moves that don't

Width is defined as it is in order to ensure that certain types of moves from one normal form of $K \cup \gamma$ to another will decrease width, some will increase width, and others will have no effect. We list the possible ways in which adjacent minima and maxima can be moved: Suppose, as in the definition of $W(K \cup \gamma)$, s_i is a regular level, lying between adjacent critical levels or heights of vertices t_{i-1} and t_i, with $t_{i-1} < s_i < t_i$. For each $1 \leq i \leq n$, let $p_i = |S(s_i) \cap (K \cup \gamma)|$.

We will assume there is an isotopy of $K \cup \gamma$ whose only effect on the positioning of critical points is to exchange the critical points at two adjacent levels, say t_{i-1} and t_i. (We say that the critical point at t_{i-1} is *moved* past the critical point at t_i.) For this to make easy sense, we must assume that no subarc connecting the two critical points lies completely between the two levels.

Simple computations (see Figure 1) show the following effects:

Claim 3.1 *If the critical points at t_{i-1} and t_i are both maxima or both minima, moving one past the other has no effect on the width.*

Proof This is obvious if they are both regular maxima, both regular minima, both Y–vertices or both λ–vertices. The remaining representative case is this: at t_i there is a Y–vertex and at t_{i-1} there is a regular minimum.

Only p_i is changed (to $p'_i = p_i - 1$). But because the height of the Y–vertex changes, note that, before the change, the term $2p_{i-1} + p_i$ appears in the definition of $W(K \cup \gamma)$ and afterwards the term $p_{i-1} + 2p'_i$ appears. But since $p_{i-1} = p_i - 2 = p'_i - 1$ the change has no effect. □

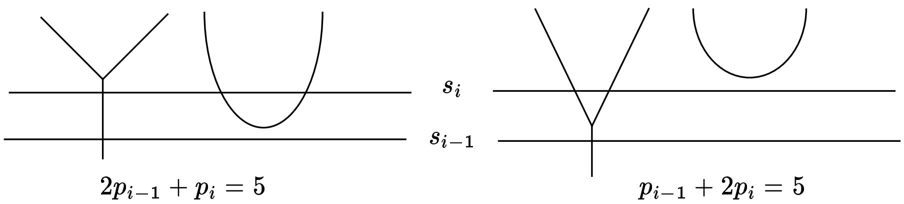

Figure 1

Claim 3.2 *If the critical point at t_{i-1} is a minimum and that at t_i is a maximum then $W(K \cup \gamma)$ is reduced.*

Proof Only p_i is affected. If both critical points are regular critical points, p_i is reduced by 4 so $W(K \cup \gamma)$ is reduced by 8.

If both are vertices, then p_i is reduced by 2, so $W(K \cup \gamma)$ is reduced by 2.

If the critical point at t_{i-1} is a Y–vertex and the one at t_i is a regular maximum or if the critical point at t_{i-1} is a regular minimum and the one at t_i is a λ–vertex, then p_i is reduced to $p_i' = p_i - 3$. The effect on the calculation of $W(K \cup \gamma)$ is, in the first case, to replace the term $p_{i-1} + 2p_i$ with $2p_{i-1} + p_i'$. Since $p_{i-1} = p_i - 1$ the net effect is to reduce $W(K \cup \gamma)$ by 4. (See Figure 2.) In the second case the term $2p_{i-1} + p_i$ is replaced by $p_{i-1} + 2p_i'$. Since here $p_{i-1} = p_i - 2$ the net effect is again to reduce $W(K \cup \gamma)$ by 4. □

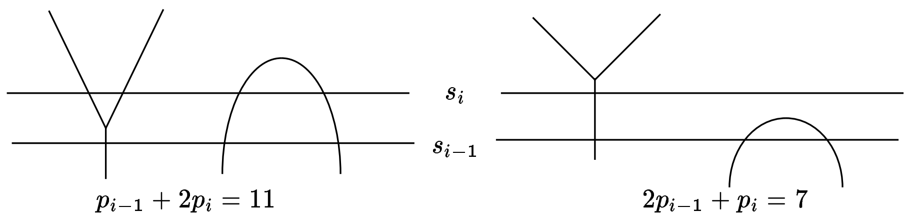

Figure 2

Now we consider the case in which the critical points at levels t_{i-1} and t_i are *adjacent*. That is, suppose there is a subarc of $K \cup \gamma$ that descends from one to the other. If one is a maximum and the other a minimum, it may be possible to

- cancel the pair (if neither is a vertex),
- cancel a regular maximum (respectively minimum) by changing a Y–vertex (respectively λ–vertex) to a λ–vertex (respectively Y–vertex), or
- replace a Y–vertex at t_{i-1} and λ–vertex at t_i with a λ–vertex at t_{i-1} and Y–vertex at t_i.

It's easy to check that in all these cases, $W(K \cup \gamma)$ is reduced (or, in the last case, it is possible that γ can be directly levelled). The upshot of all these comments is the following overarching principle:

Proposition 3.3 *Suppose K is in minimal bridge position and S is a generic level sphere lying just above a maximum of $K \cup \gamma$. Suppose a subarc $\alpha \subset (K \cup \gamma)$ doesn't contain the maximum, has ends incident to S from below, and α can be moved (by slide and/or isotopy) to lie on S, with its ends incident to the ends of $(K \cup \gamma) - \alpha$, as before. Then this move will reduce $W(K \cup \gamma)$. A symmetric statement holds for S just below a minimum. (See Figure 3.)*

Proof Conceptually, in terms of the height function on α, the arc α can be moved, leaving its complement $(K \cup \gamma) - \alpha$ fixed, in a process that only moves maxima down, minima up, with possible cancellations occurring. (There is no claim that the actual motion of α can be made to realize this conceptual process.) These moves can only thin and can never thicken. The arc α has at least one minimum, and it is eventually moved above the maximum that lies just below S. Hence, at least by this move, the presentation is thinned. $\square$

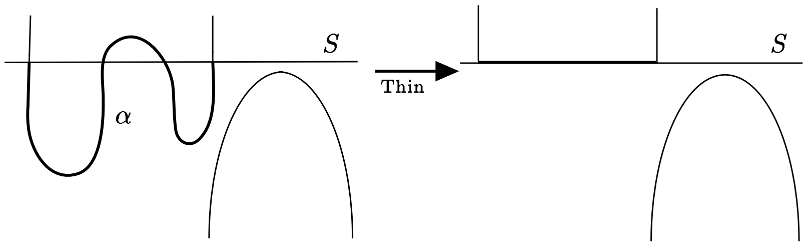

Figure 3

Note that the move described in Proposition 3.3 could change the bridge presentation of K. This can happen only if α has a single minimum (because otherwise K itself would have been thinned by the process) and, during the isotopy of α, there are times when α has more critical points than a single minimum.

4 A thin $K \cup \gamma$ is in bridge position

Just as Corollary 1.5 shows that if K is in thin position it is in bridge position, we will show in this section that if $K \cup \gamma$ is put in thin position rel K it is also in bridge position.

Recall that the neighborhood $V_1 = \eta(K \cup \gamma)$ is a genus 2 handlebody, as is its closed exterior $V_2 = E(K \cup \gamma)$. We begin by recalling some crucial facts about surfaces in genus two handlebodies.

Definition 4.1 Let M be a compact 3–manifold and P a compact 2–manifold properly imbedded in M. We say that P is *∂–parallel* if P is isotopic rel ∂P to a compact manifold in ∂M, and that P is *essential* if P is incompressible and has a component which is not ∂–parallel.

Proposition 4.2 [7] *Let M be an orientable closed 3–manifold with a genus 2 Heegaard splitting (V_1, V_2). If M contains a 2–sphere S such that each component of $S \cap V_1$ is a non-separating disk in V_1 and $S \cap V_2$ is an essential planar surface in V_2, then M has a lens space or $S^2 \times S^1$ summand.*

Proposition 4.3 [7] *Let V be a genus $g > 1$ handlebody, and P be a connected incompressible planar surface with $\ell > 1$ boundary components properly imbedded in V. If ∂P consists of mutually parallel separating loops in ∂V, then $\ell = 2$ and P is a ∂–parallel annulus.*

(In fact, Morimoto also shows that when $g = 2$ we can drop the hypothesis that the loops are separating.)

We proceed then to prove:

Proposition 4.4 *Suppose that K is in minimal bridge position and (K, γ) is in thin position rel K. Then $K \cup \gamma$ is in bridge position.*

Proof If not, there is a regular value t_0 such that $S(t_0)$ lies between adjacent critical values x and y of h, where x is a minimum of $K \cup \gamma$ lying above t_0 and y is a maximum of $K \cup \gamma$ lying below t_0. Let P be the planar surface $S(t_0) - \mathrm{int}(\eta(K \cup \gamma))$.

Case A γ is a single edge (ie, γ is not an eyeglass).

Compress P as much as possible in $E(K \cup \gamma)$. The compressions may take place to either side of P, and may have to be done in several steps. Let $\tilde{P}$ be the resulting collection of meridional planar surfaces. $\tilde{P}$ is incompressible in $E(K \cup \gamma)$. Since γ is an edge, each component of $\partial\tilde{P}$ is non-separating in $\partial\eta(K \cup \gamma)$. Thus, by Proposition 4.2, each component of $\tilde{P}$ is boundary parallel. Let P' be a component of $\tilde{P}$, so $S' = P' \cup (\text{meridian disks of } K \cup \gamma)$ bounds a 3–ball, one of 5 cases as in Figure 4. In each case, we can find a disk E as in Figure 4. That is, ∂E consists of two arcs α and $\beta, \partial\alpha = \partial\beta$, with α a properly imbedded essential arc in P', and β an arc imbedded on $\partial E(K \cup \gamma)$. Moreover,

(1) β is parallel to an arc of K, (P' is an annulus),
(2) β is parallel to an arc of γ, (P' is an annulus),
(3) β is parallel to an arc of $K \cup \gamma$ which contains a vertex, (P' is a three punctured sphere),

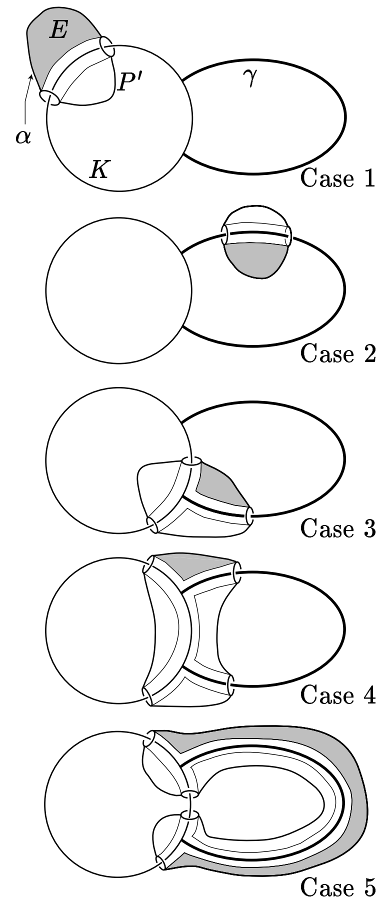

Figure 4

(4) β is parallel to an arc of $K \cup \gamma$ which contains a vertex, (P' is a four punctured sphere),

(5) β is parallel to whole $\gamma\cup$(two arcs) of K, (P' is a four punctured sphere).

Now reverse the compressions on $\tilde{P}$ to reassemble P. This means that the components of $\tilde{P}$ are reattached by tubes, which may run through each other, and which may intersect E. However, since the tubes can be chosen to be disjoint from a neighborhood of ∂E, α remains an arc entirely on $S(t_0)$ and the ends of β lie either both below or both above $S(t_0)$, say below. Then by a slide or isotopy E can be used to move β to α, and thus above level y. This contradicts thinness, via eg, Proposition 3.3. (In particular, in Case 5, γ becomes level and we are done.)

All but the first case can be viewed as slides and isotopies of γ and so do not affect the bridge presentation of K. So to finish Case A it suffices to show that not all components of $\tilde{P}$ are boundary parallel annuli. The first observation is that if any of these annuli were nested then the thinning move on E, defined for an annulus that is not innermost, would actually thin K, which we know to be impossible. Hence the annuli are not nested. For the same reason we know that the segments of K which are cut off by the annuli must all lie above (or all below) the sphere $S(t_0)$. So if these boundary parallel annuli constitute all of $\tilde{P}$ then the only part of $K \cup \gamma$ that lies above $S(t_0)$ are the arcs of K which the annuli cut off. In particular, the minimum of $K \cup \gamma$ that lies just above $S(t_0)$ must also belong to K. But this is a contradiction, for the isotopy of K which moves the segment containing the minimum down to $S(t_0)$ would thin K itself.

Case B γ is an eyeglass.

We denote by γ_c the circle part of γ, and by γ_a the arc part of γ as in Figure 5. If the link $K \cup \gamma_c$ is disjoint from $S(t_0)$, the proof proceeds much as in Case A: Compress P as much as possible in the complement of $K \cup \gamma$ and let P' denote a component of the result. Since P' intersects $K \cup \gamma$ only in γ_a, it follows from Proposition 4.3 that P' is a ∂–parallel annulus. This implies, as in Case A, that a subarc of γ_a can be isotoped to P, contradicting thinness via Proposition 3.3. K is left unchanged. An argument as in Case A-5 similarly applies if γ_a is disjoint from $S(t_0)$.

It remains to consider the case in which K or γ_c and also γ_a do intersect $S(t_0)$; a preliminary construction may be needed. By a slide, make γ into a simple arc γ'. (Of course, after the slide $K \cup \gamma'$ may no longer be in thin position.) The new arc γ' consists of three subarcs: γ_c (with a small subarc γ_ϵ removed near its vertex), γ_a, and a new arc γ'_a parallel to γ_a. Let G be a disk which gives the parallelism of γ_a and γ'_a (see Figure 5). The boundary of G consists of four arcs: γ_a, γ'_a, γ_ϵ, and a short subarc K_ϵ of K. We may take γ_ϵ and K_ϵ to be disjoint from $S(t_0)$.

As in Case A, compress P as much as possible in the complement of $K \cup \gamma'$ and let P' denote a component of the result, chosen to have at least one boundary component a meridian of either K or γ_c. From Proposition 4.2 P' is ∂–parallel in $V_2 = E(K \cup \gamma')$. After the compressions, it may be that P' intersects γ_ϵ, but it must remain disjoint from K_ϵ. We have 4 cases as illustrated in Figure 6.

In each case, we can find a disk E as in Figure 6. That is, ∂E consists of two arcs α and $\beta, \partial\alpha = \partial\beta$, with α a properly imbedded essential arc in P', and

β an arc imbedded on $\partial E(K \cup \gamma')$. Note that $E \cap G = \emptyset$. We write as in Case A:

(1) β is parallel to an arc of K, (P' is an annulus),

(2) β is parallel to an arc of $\gamma_a \cup \gamma_c$ or $\gamma_a' \cup \gamma_c$, (P' is an annulus),

(3) β is parallel to an arc of γ_c, (P' is an annulus),

(4) β is parallel to an arc of $K \cup \gamma_a$, (P' is a four punctured sphere).

Now reverse the compressions on $\tilde{P}$ to reassemble P. This means that the components of $\tilde{P}$ are reattached by tubes, which may run through each other, and which may intersect E and G. However, E and G remain disjoint, and ∂E remains unchanged since the tubes can be chosen to be disjoint from a neighborhood of ∂E. So E persists as a disk which can be used to move, by isotopies and edge slides, some subarc of the *original* $K \cup \gamma$ to the level sphere $S(t_0)$, reducing $W(K \cup \gamma)$ via Proposition 3.3.

As in Case A, this thinning can be done without affecting the bridge position of K unless all components of $\tilde{P}$ are boundary parallel annuli; but this would lead to the same contradiction as in Case A. □

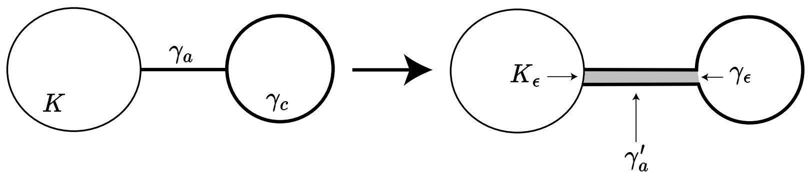

Figure 5

5 Proof of Theorem 1.8

Once $K \cup \gamma$ is in bridge position, one can push any Y–vertex below any distant minimum and any λ–vertex above any distant maximum without increasing the width or altering the bridge structure of K. So, typically, it is possible to choose a thin position in which neither the highest local minimum nor the lowest local maximum occurs at a vertex. The single possible exception (for which the classification of tunnels is anyway already known) is the case in which K is a 2–bridge knot and γ is an arc with exactly one critical point in its interior (see Figure 7).

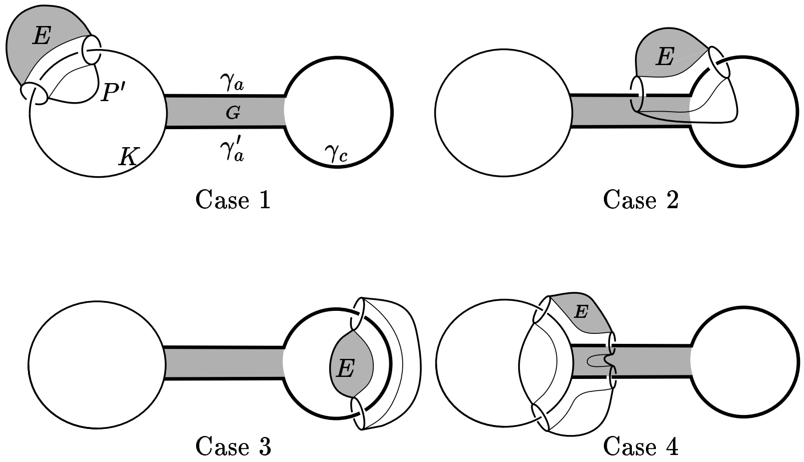

Figure 6

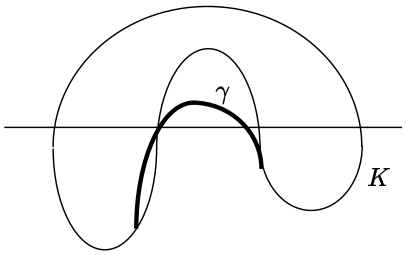

Figure 7

Lemma 5.1 *Suppose K is a tunnel number 1 knot in minimal bridge, hence thin, position. Then $K \cup \gamma$ can be put in thin position rel K so that, for x the level of the highest local minimum and y the level of the lowest local maximum, there is a $z \in (x, y)$ so that $S(z)$ simultaneously has an upper disk and a lower disk, and these disks are disjoint.*

Proof This proof is essentially that of Lemma 4.4 in [2]. As above, let D be a meridian disk for V_2. First observe that ∂D runs along every edge of the graph $K \cup \gamma$, for otherwise it would follow that K is trivial. As in Section 4 in [2], by an isotopy, we may assume that the local picture of D near regular maxima and regular minima is as illustrated in Figure 8 (see also Definition 2.7 (e)).

With the exception noted above (when K is a 2–bridge knot and γ is an arc with exactly one critical point in its interior), we may assume that the minimum at x and the maximum at y are regular critical points, not vertices.

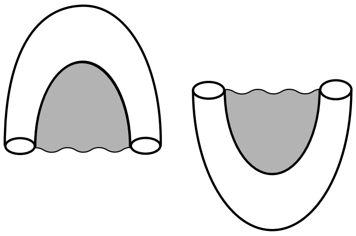

Figure 8

In particular, from Figure 8 we see that there is a small value ε such that $S(t)$ cuts off an upper outermost disk from D for $t \in [y-\varepsilon, y]$ and cuts off a lower outermost disk from D for $t \in [x, x+\varepsilon]$. At any regular level t between x and y we have $(K \cup \gamma) \cap S(t) \neq \emptyset$ so there is either an upper or a lower disk. Since at the lowest level there are lower disks and the upper level there are upper disks and at every regular value there is one or the other, it would seem to follow immediately that at some level between x and y there are both upper and lower disks, as required.

But there remains the possibility that passing upward through a critical level t_c for D (necessarily corresponding to a saddle intersection of $S(t_c)$ with the interior of D), two lower disks disappear while two upper disks are created. In this case, simply thicken D very slightly, creating a collar whose boundary consists of two disjoint copies of D, each of them transverse to $S(t_c)$. $S(t_c)$ will cut off from one an upper disk and from the other a lower disk.

This completes the proof for all but the exceptional case: Suppose there is a single critical point in the interior of γ, say a maximum, and the ends of γ are incident to K just above distinct minima of K, of which there are exactly two. (See Figure 7.) Then the highest minimum on $K \cup \gamma$ occurs at a Y–vertex, from which one end of γ ascends. The argument above still applies, unless the local picture of D near the Y–vertex looks as in Figure 9, that is, unless the height function on D has one or more half-saddles at x. This situation is characterized by the condition that ∂D intersects a meridian of the edge descending from the Y–vertex in more points than the sum of the number of points it intersects the meridians ascending from the Y–vertex.

There are two possible complications in this case: One is the possibility that there could be upper disks persisting just below level x. We leave as an exercise for the reader, most easily done after seeing the proof of Theorem 5.3, that such an upper disk could be used to thin the presentation further, a contradiction. A more subtle problem is that an upper disk D_u for t slightly above x could be destroyed by the half-saddle singularities at x and one of the pieces become a lower disk D_l for t slightly below x.

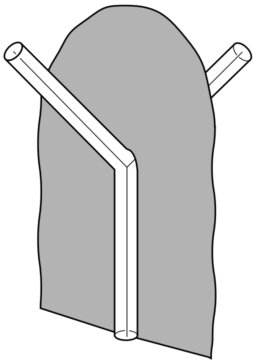

Figure 9

Note that the lower disk intersects $K \cup \gamma$ in a subarc of K, namely the arc around the minimum of K to which the Y–vertex is adjacent. This lower disk can be used to slide the end of γ at the Y–vertex across the minimum at K and back to height x so that afterwards no new critical points are introduced on γ. (See Figure 10.) In particular, the placement of $K \cup \gamma$ is still thin. But now the intersection of D with a neighborhood of the (new) Y–vertex is of a form that guarantees a lower disk for $t \in (x, x+\varepsilon]$, since it is now the meridian of an ascending arc from the Y–vertex that intersects ∂D most often. The argument then proceeds as above. □

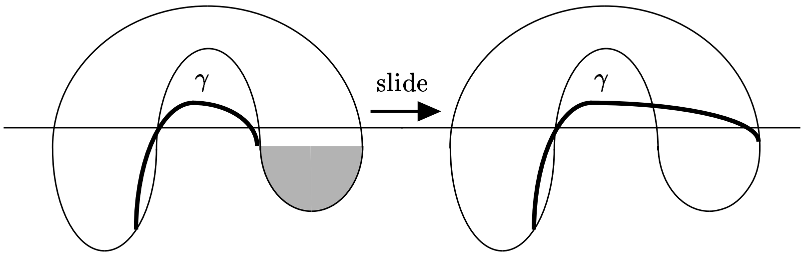

Figure 10

The previous proof shows that, even once $K \cup \gamma$ is in thin position, it may be useful to choose a particular thin position. In fact, we will need the following highly technical lemma.

Suppose, when $K \cup \gamma$ is put in thin position, some edge $e \neq \gamma$ of the graph $K \cup \gamma$ monotonically descends from one λ–vertex to another. This is equivalent to the condition that, for $S(z)$ a middle sphere, some edge of $K \cup \gamma$ lies entirely above $S(z)$ or, equivalently, some component X of $(K \cup \gamma) - S(z)$ lying above $S(z)$ is a tree with four ends and all other components of $(K \cup \gamma) - S(z)$ are arcs. Let P be the 4–punctured sphere $\eta(X) \cap \partial\eta(K \cup \gamma)$ and $C \subset P$ be a properly imbedded collection of essential arcs, with the property that the 4 components of ∂P can be grouped into two pairs, each with the same number of endpoints of ∂C. (In practice, C will be the intersection of the boundary of a meridian disk for V_2 with P.)

Lemma 5.2 *Under these conditions, after a series of edge slides which do not affect the width of $K \cup \gamma$ or the bridge presentation of K, a meridian for the (possibly different) monotonic edge e will intersect each component of C in at most one point.*

Proof Since $e \neq \gamma$, just two of the other 4 edges of X lie in K. By edge slides that introduce no new critical points, e can be shrunk to a point, so that the only vertex in X has one edge ascending from it and three edges descending. The mapping class group of the 4–punctured sphere is generated by twists around a pair of simple closed curves each separating a different pair of punctures and intersecting each other in exactly two points. In our context, that means that the entire mapping class group is generated by twists among the descending triple of edges. Such a twist is either an edge slide (if one of the strands is in γ) or just an isotopy (if both edges are in K).

It's easy to see that *some* simple closed curve in P separating one pair of punctures from another will intersect each component of C in at most one point. (To see this, note that the condition on $|\partial P \cap C|$ guarantees that all arcs of C that have both ends on a given boundary component of P will be parallel.) After an appropriate sequence of braid moves, we can arrange that this simple closed curve is the meridian of e. Afterwards, unshrink e, recovering a thin presentation of $K \cup \gamma$ with the required property. (Note that during this process the monotonic edge e may switch from being in K to being in γ or vice versa.) □

Of course a symmetric statement holds if there were instead an edge of $K \cup \gamma$ lying entirely below $S(z)$.

Theorem 5.3 *Let K be a tunnel number one knot in minimal bridge position and γ be a tunnel for K. Put $K \cup \gamma$ in thin position rel K. Then if γ is an arc, the height function is monotonic on γ and γ can be slid and isotoped* *rel K to be level. If γ is an eyeglass $\gamma_a \cup \gamma_c$, then the height function has a single maximum and a single minimum on the circuit γ_c, and γ can be slid and isotoped, rel K, so that γ_c is level.*

Proof Put $K \cup \gamma$ in thin position (following the procedure of Lemma 5.2 if applicable) so that there is a level z as in Lemma 5.1. That is, $S(z)$ simultaneously has an upper disk D_u and a lower disk D_l, and these disks are disjoint. Moreover, if there is an edge $e \neq \gamma$ of the graph $K \cup \gamma$ lying entirely above (below) $S(z)$, then ∂D_u (respectively ∂D_l) intersects the meridian of the edge at most once.

If γ itself is disjoint from $S(z)$ then, because $K \cup \gamma$ is in bridge position, it is easy to slide γ into $S(z)$, completing the proof. So we will henceforth assume that if an edge e of $K \cup \gamma$ is disjoint from $S(z)$, then $e \neq \gamma$ so, following the application of Lemma 5.2, a meridian of e will intersect an upper or lower disk in at most one point.

Let α_u and α_l denote the (disjoint) arcs $\partial D_u \cap S(z)$ and $\partial D_l \cap S(z)$ respectively. We will say that D_u (respectively D_l) is *bad* if α_u (respectively α_l) is a loop. Otherwise the disks are *good*. Notice that if an edge e of $K \cup \gamma$ is disjoint from $S(z)$, then it is necessarily also disjoint from any bad disk, since the boundary of a bad disk crosses each meridian an even number of times.

Definition 5.4 An *upper cap* (respectively *lower cap*) to $K \cup \gamma$ at $S(z)$ is an imbedded disk $C \subset S^3$ transverse to $S(z)$ so that

(1) $C \cap (K \cup \gamma) = \emptyset$,

(2) $C \cap S(z) = \partial C$,

(3) ∂C is essential in $S(z) - (K \cup \gamma)$, and

(4) the interior of C lies entirely above (respectively below) $S(z)$.

Claim 5.5 *There cannot be both an upper cap and a lower cap whose boundaries are disjoint.*

Proof The boundary of a cap divides $S(z)$ into two disks. If the cap C is an upper (respectively lower) cap, then, for each of the disks in $S(z) - \partial C$, an arc of $K - S(z)$ or $\gamma - S(z)$ lying above (respectively below) $S(z)$ can be isotoped to lie in the disk, indeed just below (respectively above) $S(z)$. For upper cap C_u and lower cap C_l with disjoint boundaries, there are disk components of $S(z) - C_u$ and $S(z) - C_l$ that are disjoint. Call these disks S_u and S_l. Then some arc of $K - S(z)$ or $\gamma - S(z)$ lying above (respectively below) $S(z)$ can be isotoped to lie in S_u (respectively S_l). If in both cases the arc were in K this would violate thin position of K, so at least one of the arcs (say the one in S_u) must be in γ. But even then the isotopy would violate thin position of $K \cup \gamma$ rel K, since pushing both arcs through $S(z)$ would reduce $W(K \cup \gamma)$ but would not change the bridge position of K since $S(z+\epsilon)$ would remain a middle sphere for K. □

Claim 5.6 *If there is an upper disk (respectively lower disk) and a disjoint lower cap (respectively upper cap) then we can find such a pair for which the interior of the upper (respectively lower) disk is disjoint from $S(z)$.*

Proof Let B_u and B_l denote the balls which are the closures of the region above $S(z)$ and below $S(z)$ respectively. It is easy to choose D_u so that each component of $D_u \cap B_u$ is incompressible in the complement of $K \cup \gamma$.

Since $K \cup \gamma$ is in bridge position, there is a disjoint collection of disks Δ (called *descending disks*) so that

(1) $\mathrm{int}(\Delta)$ lies above $S(z)$,

(2) each boundary component of Δ consists of an arc in $S(z)$ and an arc in $K \cup \gamma$, and

(3) the complement of an open neighborhood of $\Delta \cup K \cup \gamma$ in B_u is also a 3–ball.

For example, if no vertex of $K \cup \gamma$ lies above $S(z)$, then Δ is just a set of disks defining a parallelism between the set of arcs lying above $S(z)$ and arcs on $S(z)$. It is possible to choose Δ so that all intersection points of $\partial\Delta$ with ∂D_u lie on $S(z)$. This is obvious if the component of $(K \cup \gamma) - S(z)$ incident to D_u has three ends, since in a 3–punctured sphere, arcs connecting two of the punctures to the third can be isotoped off an arc with both ends on the third. If the component X of $(K \cup \gamma) - S(z)$ incident to D_u has four ends then we have carefully contrived, via Lemma 5.2, that ∂D_u runs at most once along the edge of $K \cup \gamma$ that lies in X. This observation allows us to revert to the case of the 3–punctured sphere as above.

The proof of the claim will be by induction on $|D_u \cap \Delta|$. If $|D_u \cap \Delta| = 0$ then each component of $D_u \cap B_u$ lies in a ball disjoint from $K \cup \gamma$ so, by incompressibility, each component is a disk. Since a neighborhood of ∂D_u lies in B_u, it follows that D_u must lie entirely inside B_u as required.

So suppose $D_u \cap \Delta \neq \emptyset$. A simple innermost disk argument could eliminate a closed curve of intersection, so we can take all components of intersection to be arcs. Surprisingly, we may also assume that the lower cap is a slight push-off of a disk component of $D_u \cap B_l$. Indeed, consider an innermost disk of $D_u - S(z)$. If it lies in B_u then it is an upper cap disjoint from the lower cap which we know by Claim 5.5 to be impossible. If it lies in B_l then we may as well take a slight push-off as our lower cap.

This surprising fact means that an outermost arc of $D_u \cap \Delta$ in Δ can be used to ∂–compress $D_u \cap B_u$ to an arc that is disjoint from the lower cap. This boundary compression defines an isotopy on the interior of D_u that reduces $|D_u \cap \Delta|$ without disturbing the disjoint lower cap. Further, this isotopy preserves the incompressibility of $D_u \cap B_u$ in the complement of $K \cap \gamma$. After the isotopy, the result follows by induction. □

Claim 5.7 *There cannot be both an upper disk and a disjoint lower cap (or a lower disk and a disjoint upper cap).*

Proof Following Claim 5.6 we can assume that the upper disk D_u lies entirely above $S(z)$. The boundary of the lower cap divides $S(z)$ into two disks; let S_l be the one disjoint from D_u.

If D_u is good, then D_u can be used to move an arc component of $K - S(z)$ or $\gamma - S(z)$ lying above $S(z)$ down to just below α_u. We have already noted that the lower cap ensures that an arc component of $K - S(z)$ or $\gamma - S(z)$ lying below $S(z)$ can be moved up just above S_l. This violates thin position just as in Claim 5.5.

If D_u is bad then the loop α_u divides $S(z)$ into two disks, let S_u be the one that is disjoint from S_l. Now ∂D_u is incident to exactly one edge of $(K \cup \gamma) - S(z)$: this is obvious if each component of $(K \cup \gamma) - S(z)$ has at most three ends and has been previously arranged if a component of $(K \cup \gamma) - S(z)$ has four ends. The union of D_u and a neighborhood of that edge form a disk much like an upper cap. Since S_u contains some points of $S(z) \cap (K \cup \gamma)$, some edge of $K - S(z)$ or $\gamma - S(z)$ can be slid or isotoped to lie in S_u. Again this violates thin position as in Claim 5.5. □

Claim 5.8 *There cannot simultaneously be an upper bad disk and a lower good disk that are disjoint (or, symmetrically, a lower bad disk and an upper good disk that are disjoint).*

Proof We may assume that the interiors of both disks are disjoint from $S(z)$, for, if not, then an innermost curve of intersection in the disk would cut off a cap that would contradict Claim 5.7. Let D_u be the upper bad disk and D_l be the lower good disk. The loop α_u divides $S(z)$ into two disks S_l and S_u with $\alpha_l \subset S_l$. The disk D_l can be used to move an arc component of $K - S(z)$ or $\gamma - S(z)$ lying below $S(z)$ up to just below α_l. Since S_u contains some points of $S(z) \cap (K \cup \gamma)$, some edge of $K - S(z)$ or $\gamma - S(z)$ can be slid or isotoped to lie in S_u. Again this violates thin position as in Claim 5.5. □

By Lemma 5.1, there are an upper disk D_u and a lower disk D_l for $S(z)$ such that $D_u \cap D_l = \emptyset$. We may assume that the interiors of both disks are disjoint from $S(z)$, for, if not, then an innermost curve of intersection in the disk would cut off a cap that would contradict Claim 5.7. By Claim 5.8 we can assume that either both are good or both are bad.

If both are good, then they can be used to slide and isotope an arc component of $K - S(z)$ or $\gamma - S(z)$ lying below $S(z)$ to just above $\alpha_l \subset S(z)$ and an arc component of $K - S(z)$ or $\gamma - S(z)$ lying above $S(z)$ to be moved just below α_u. If α_l and α_u have no endpoints in common, then these moves could be done simultaneously and $K \cup \gamma$ would not have been thin. The same contradiction arises if they have one end point in common, unless it is in γ and the other two end points are in K. In that case the simultaneous isotopies and slide make γ level as required. If α_l and α_u have both end points in common, they can't both be in K (since K is knotted) and if both end points are in γ then there is a circuit $\gamma_c \subset \gamma$ that can be made level, as required. If one common end is in K and the other in γ then γ can be moved to a level loop and again we are done.

If both the upper and lower disks are bad, then the argument is much the same. That is, there are disjoint disks in $S(z)$ bounded by the loops α_u and α_l, and arc components of $K - S(z)$ or $\gamma - S(z)$ can be moved into those disks. If these loops were based at different points, or at a common point lying in K, this move would violate thin position. If the loops are based at a common point and it is in γ, then γ can be made level, as required. □

Definition 5.9 Suppose K is a tunnel number one knot and γ is a tunnel for K in the form $\gamma = \gamma_a \cup \gamma_c$, where γ_c is a circuit. Suppose K is in minimal bridge position with respect to a height function h, γ_c is level with respect to h, and $K \cup \gamma_a$ is in Morse position with respect to h with no critical point at the same height as γ_c. Then $K \cup \gamma$ is in *special position* with respect to h.

It follows from Theorem 5.3 that for K in thin position, either the entire tunnel γ can be made level, or $K \cup \gamma$ can be put in special position without affecting the bridge position of K.

We now mimic much of the previous argument:

Definition 5.10 Suppose $K \cup \gamma$ is in special position (so in particular, K is in minimal bridge position). Let $t_0 < \ldots < t_n$ be the list of critical heights of $K \cup \gamma_a$, plus the level of the cycle γ_c. Let $s_i, 1 \leq i \leq n$ be generic levels chosen so that $t_{i-1} < s_i < t_i$ and let t_j be the level of the vertex $K \cap \gamma_a$ (ie, the vertex *not* on γ_c). Define the width $W(K \cup \gamma)$ to be the integer

$$2(\Sigma_{i \neq j}|S(s_i) \cap (K \cup \gamma_a)|) + |S(s_j) \cap (K \cup \gamma_a)|.$$

For K in minimal bridge position *special thin position* of a pair (K, γ) is a special position presentation which minimizes $W(K \cup \gamma)$ rel K.

Much as in Section 3, maxima can be moved past maxima and minima past minima without affecting the width. Moving a maximum below a minimum will always reduce width. What is new is this: moving a maximum below the level of γ_c or moving a minimum above the level of γ_c will always thin the special presentation. This is easily checked. In particular, the overarching principle (Proposition 3.3) remains intact, and applies also to the level just above γ_c (and, symmetrically, to the level just below γ_c).

A *special bridge position* of $K \cup \gamma$ is a special position in which every minimum of $K \cup \mathrm{int}(\gamma_a)$ occurs at a lower level than every maximum. (Note that nothing is said about the level of γ_c.)

Proposition 5.11 *Suppose that (K, γ) is in special thin position. Then, $K \cup \gamma$ is in special bridge position.*

Proof The proof is essentially identical to that of Proposition 4.4, Case B. Only easier: the cases in which P' has a boundary component on γ_c don't arise. □

Lemma 5.12 *Suppose that (K, γ) is in special thin position, so in special bridge position. Let x be the level of the highest local minimum and y the level of the lowest local maximum of $K \cup \gamma$. Then the level of γ_c lies between x and y.*

Proof Suppose, on the contrary, that the level of γ_c is greater than y and choose a level z just below the level of γ_c (and so above y). Repeat the argument of Proposition 4.4, Case B, at level z. Either a maximum above the level of γ_c can be moved down to level z, thinning z, or a subarc below z can be moved up to the level of z. Since below z the next critical level is at a maximum, this necessarily thins the presentation. $\square$

Lemma 5.13 *Suppose that (K,γ) is in special thin position, x is the level of the highest local minimum, y is the level of the lowest local maximum and γ_c is at level z between x and y. Then there is not an upper disk for any $S(t), x < t < z$ or a lower disk for any $S(t), z < t < y$.*

Proof The proof mimics that of Lemma 5.1. As in that lemma, we can arrange that the levels x and y are not the levels of the other vertex. Then suppose, say, that $S(t), z < t < y$ has a lower disk. We know that $S(y-\varepsilon)$ has an upper disk, so for some level between t and y there is a level t_0 with both an upper and lower disk. The proof is now essentially identical to that of Theorem 5.3, except that the bad lower disk D_l might intersect $K \cup \gamma$ on a component which is the union of γ_c and a subarc of γ_a. But it's easy to see that the arc $\partial D_l \cap (K \cup \gamma)$ must then intersect a meridian of γ_c exactly once, and so can be used to slide γ_c up to level t_0 (and, of course, disjoint from the slide or isotopy given, as in the proof of Lemma 5.1, by D_u). The result would be to thin the presentation. $\square$

Theorem 5.14 *Suppose that (K,γ) is in special thin position. Then, γ may be isotoped and slid, rel K, until it is level.*

Proof Let z be the level of γ_c, necessarily lying between the level of the highest minimum and the level of the lowest maximum.

Thicken a neighborhood of γ_c slightly. Then $S(z) - \eta(\gamma_c)$ consists of two disks, S_1 and S_2, each punctured by K and γ_a. As above, let D be a meridian disk for V_2, with ∂D on the boundary of a regular neighborhood of $K \cup \gamma$. Isotope D to minimize the pair, in lexicographic order, $(|\partial D \cap S(z)|, |D \cap S(z)|)$. Suppose $D \cap S(z)$ contains closed components, and let C be an innermost disk cut off by $S(z)$ from D. C is a cap, lying either above or below $S(z)$. If, say, it lies above, then ∂C bounds a disk in, say, S_1. Then whatever maxima lie between $S(z)$ and the cap can be pushed down below S_1, thinning the presentation but not affecting the bridge presentation of K. We conclude that all components of $D \cap S(z)$ are properly imbedded arcs.

Let D_u be an outermost disk cut off from D by $S(z)$ and, say, D_u lies above $S(z)$. Let α_u be the arc $\partial D_u \cap S(z)$. If the ends of α_u lie at distinct points of $(K \cup \gamma_a) \cap S(z)$ then D_u could be used to isotope or slide a maximum below level z, thinning the presentation. Similarly, if α_u is a loop with ends at a point in $K \cup \gamma_a$, it bounds a disk in, say, S_1. Again, whatever maxima have ends in that disk can be pushed below S_1, thinning the presentation.

The possibility remains that one or both ends of α_u lie on ∂S_i. In this case, γ_a must be incident to $\eta(\gamma_c)$ from above. Consider the annulus half A of $\partial\eta(\gamma_c)$ lying above level z, punctured by the end of γ_a. No component of $\partial D \cap A$ can have both ends at the puncture, for otherwise there would be a lower disk cut off from D just above $\eta(\gamma_c)$, contradicting Lemma 5.13. (See Figure 11.) Hence there is a spanning arc of A which is disjoint from $\partial D \cap A$. This observation allows us to treat γ_c much like just another vertex in the previous argument. In particular, if α_u has its other end at a point in $\gamma_a \cap S_i$ then the end of γ_a cut off by that point could be isotoped just below S_i, thinning the presentation. If the other end is at a point in $K \cap S_i$ then γ_a can be made to lie in an S_i, completing the proof.

Finally, suppose that both ends of α_u lie in ∂S_i. Then ∂D_u must run up γ_a. For otherwise there is an arc of $\partial D \cap A$ which does not span A, namely $\partial D_u \cap A$. Any such non-spanning arc, either in A or in the annulus $\partial\eta(\gamma_c) - A$, must cut off a disk containing the end of γ_a, since $\partial D \cap S(z)$ has been minimized. So there can be no such arc in the annulus $\partial\eta(\gamma_c) - A$ and every such arc in A must have its ends on the same component of ∂A. The former fact implies that ∂D intersects both components of ∂A in the same number of points and the latter would imply that ∂D intersects one component more often than the other, a contradiction.

Having established that ∂D_u runs up γ_a note that, just as in the case of a bad upper disk, $D_u \cup \gamma_a$ defines a disk above $S(z)$ which, like an upper cap, can be used to push maxima down to level $S(z)$. One of these contains γ_a; when it is pushed down to level $S(z)$ the entire eyeglass will be level, as required. □

6 A characterization of unknotting tunnels

Let K be an tunnel number one knot of bridge number n and suppose $\{B_1, B_2\}$ a pair of 3–balls which gives a minimal bridge decomposition of K. That is, $S^3 = B_1 \cup B_2, B_1 \cap B_2 = \partial B_1 = \partial B_2$ and $(B_i, B_i \cap K)$ is a trivial tangle of n components $(i = 1, 2)$. Let $S = \partial B_1 = \partial B_2$ be the middle sphere and for

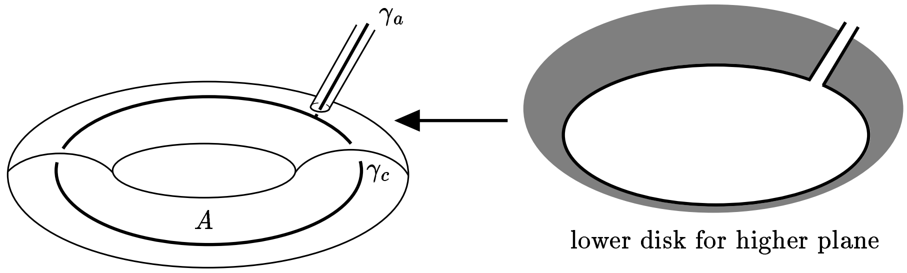

Figure 11

$i = 1, 2$, let K_i denote the collection of arcs $B_i \cap K$, parallel to a collection of arcs in S. We have shown that any tunnel γ for K can be slid and isotoped to lie in S. In this section we glean a bit more information about how γ lies in S. We show:

Theorem 6.1 *For one of (B_i, K_i, γ), $i = 1, 2$, either:*

(1) *γ is an arc with its ends on different components of K_i and K_i is parallel to a collection of arcs in $S - \gamma$,*

(2) *γ is an arc with both ends on the same component of K_i. In this case, γ can be slid and isotoped in B_i so that it lies in S as a loop with its ends at the same point of ∂K_i, or*

(3) *γ is an eyeglass and a disk that γ bounds in S contains exactly one end of each of $n - 1$ components of K_i.*

See Figure 12 for typical examples. In the last case (hence also in the second case), it is shown in [3] that there is a possibly different minimal bridge position for K with middle sphere S' for which γ appears as part of a 'quasi-Hopf tangle' (see Figure 12 (3) or [3, Figure 2]). In words, $S' \supset \gamma$ bounds a ball B' so that the collection of arcs $K \cap B'$ is parallel to a collection of arcs $\kappa \subset S'$ and the interior of exactly $n - 1$ of the arcs in κ intersects γ, each in precisely one point.

Figure 13 shows how to construct many examples where it is necessary to switch minimal bridge positions of K in order to see the quasi-Hopf tangle.

Thus we have:

Corollary 6.2 *By sliding and isotopy of γ, we can find a minimal bridge sphere of K that bounds a ball in which $K \cup \gamma$ is as seen in Figure 12 (1) or (3).*

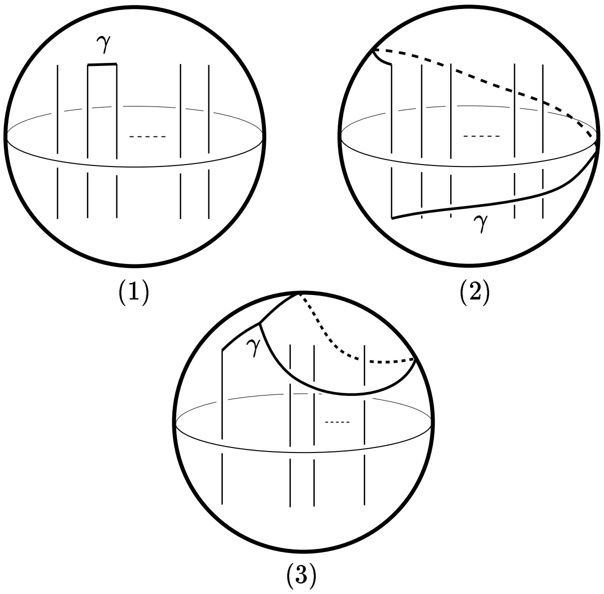

Figure 12

Besides Figure 3, here is another way in which the difference between the corollary and the theorem can be illustrated. Imagine replacing the right-most two vertical arcs in Figure 12 (3) by a rational tangle, chosen to retain the property that each of the two arc has one end on each side of the loop formed by γ. The resulting picture would satisfy the conclusions of the theorem, but not of the corollary.

Proof The initial work is done above. Recall the definitions of V_1, V_2 and their common boundary F, a genus two surface. With K in minimal bridge position we proceed to put $K \cup \gamma$ in thin position rel K and then put γ into a level sphere that lies above all the local minima for K and below all the local maxima. The end(s) of γ are incident to one or two maximal arcs of K lying above the level of γ and one or two minimal arcs lying below. We call these arcs of K *contiguous* to γ. We divide into three cases.

Case 1 Some maximal arc (minimal arc) of K not contiguous to γ can be pushed below (above) the level of γ without changing the bridge presentation of K.

Push as many maximal arcs as possible below the level of γ. Let S_- be a level sphere lying just below γ but above some maxima of K. Maximally compress

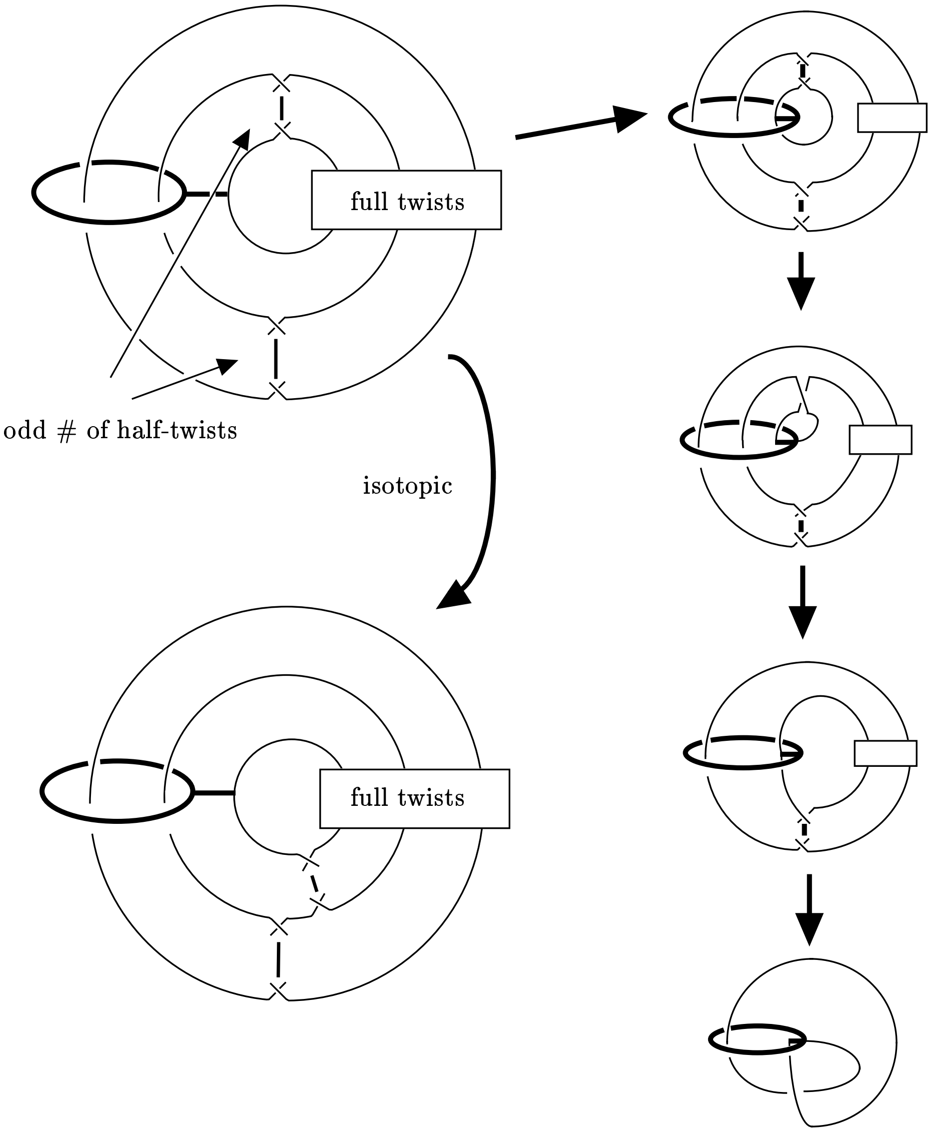

Figure 13

the planar surface $S_- - \eta(K)$ to get a planar surface P which is incompressible in $V_2 = S^3 - \eta(K \cup \gamma)$. Then, as in Section 4, P is a collection of boundary parallel planar surfaces in V_2 by Proposition 4.2. Much as in Section 4 we observe this contradiction: if an *arc* component of $(K \cup \gamma) - S_-$ is parallel to an arc of S_- this either violates thin position (if the arc lies below S_-)

or allows us to push a maximal arc below γ (if the arc lies above S_-). The latter move might alter the bridge presentation of K, but can be refined using the techniques of Claim 5.6 so that it does not, which is a contradiction to assumption. Here is a sketch of this refinement: if there is an upper cap for S_- then an arc can be pushed down without altering the bridge position of K. Since there can be no upper cap, any upper disk either has interior disjoint from S_- or is disjoint from some lower cap. But then it follows from the argument of Claim 5.6 that an upper disk can be found that is disjoint from S_-. So the isotopy it describes cannot alter the bridge presentation of K.

Thus we conclude that there is no such arc component of $(K \cup \gamma) - S_-$, so the unique component γ_+ of $(K \cup \gamma) - S_-$ lying above S_- is the one containing γ. Moreover, the part $F \cap \eta(\gamma_+)$ of the splitting surface adjacent to γ_+ must be a planar surface parallel to a component of P. In particular γ_+ cannot contain a circuit and must be parallel to a subgraph of S_-. So γ_+ is a tree with four ends and has a single interior edge, necessarily γ. This means that γ is a level edge connecting the highest two maxima of K, matching the first possibility described in the theorem.

Case 2 The maximum (minimum) of a contiguous maximal (respectively minimal) arc of K can be pushed to the level of γ without altering the bridge presentation of K.

Suppose γ is an arc with ends on two different maximal arcs. Once one of these arcs is pushed down below S, the descending disks from all the other maximal arcs can be made disjoint from γ, just by sliding any intersection off the end of γ. In particular, a maximal arc not contiguous to γ (if one exists) can be pushed below S, so we conclude as in Case 1. If there are no maximal arcs other than the two contiguous to γ, that is in the 2–bridge case, the argument still allows us to push both these maximal arcs to the level of γ, matching the first possibility described in the theorem.

Suppose γ is an arc with ends on the same maximal arc k_0. Since the assumption is that k_0 can be pushed to the same level as γ, there is a disk D_0 whose boundary consists of the union of k_0 and a subarc of $S - \gamma$. Slide one end of γ over k_0 so that γ becomes a loop, then use D_0 to push the new subarc of γ down onto S. This move demonstrates that γ satisfies the second possibility described in the theorem.

Finally, suppose that γ is an eyeglass, which we may as well take to be a loop based at a point of $K \cap S$ (ie, set $\gamma_a = \emptyset$). Once the maximum of its contiguous maximal arc is pushed down to the level of γ, we can assume that the interior of one of the two disks bounded by γ intersects K in $k \leq n-1$ points, where n is the bridge number of K, ie, the number of arcs in K_i. It then follows from [3] that a regular neighborhood of that closed disk has boundary a middle sphere S' for a possibly different bridge presentation of K. But $|S' \cap K| = 2k+2$ and since n is the minimal bridge number of K we have $2k+2 \geq 2n$. Hence in fact $k = n-1$, matching description 3 of the theorem.

Case 3 The general case.

$V_1 \cap S$ consists of some meridian disks E_i of K together with either a disk component E_0 or an annulus component A_0 that cuts γ lengthwise. It is a disk when γ is an arc and an annulus when γ is an eyeglass.

Let D be a meridian disk for V_2 and P be the planar surface $S - V_1$. If $D \cap P$ has a closed component, consider an innermost disk component of $D - P$; if not, consider an outermost disk of $D - P$, cut off by an outermost arc. The argument is similar in both cases, so we focus on a disk D_0, an outermost disk of $D - P$. With no loss of generality, assume D_0 lies entirely above S. If the arc $\beta = F \cap \partial D_0$ lies on an annulus component of $F - S$ then D_0 provides an isotopy of a maximal arc of K to the level of S, hence to a level below S and we conclude as in Case 1. Otherwise β lies on a component F_γ of $F - S$ neighboring γ. There are three possibilities:

Subcase 3a F_γ is a once-punctured torus, with $\partial F_\gamma = \partial E_0$.

This case arises when γ is an arc with ends coincident to the same maximal arc k_0 of K. The torus $F_\gamma \cup (S - E_0)$ bounds a solid torus, namely the union of a neighborhood of k_0 with the ball below S. ∂D_0 is an essential circle on the boundary of this torus bounding a disk in its complement. Hence it is a longitude of the solid torus. It follows that β runs exactly once over the arc k_0, so in fact D_0 provides an isotopy of k_0 to a subarc of S disjoint from γ. Then we conclude as in Case 2.

Subcase 3b F_γ is a three-punctured sphere, whose boundary consists of ∂E_0 and the boundary of two meridia of K.

Suppose first that both ends of β lie on ∂E_0. Then one way to think of D_0 is as a disk disjoint from K, lying above S, and crossing γ exactly once. But this implies that the descending disks Δ from the maxima of K can be made disjoint from γ: any intersection with γ can be piped to D_0 and thereby removed. In particular, all maximal arcs of K can be isotoped to lie on $S - \gamma$ so we conclude as in Case 2.

If exactly one end of β lies on ∂E_0, then D_0 can be used to push a maximal arc contiguous to γ to the level of γ and again we conclude as in Case 2.

Finally, if neither end of β lies on ∂E_0 then a subdisk of D_0 is a lower disk for a level sphere S_+ lying a bit above S. So at some level $S(z)$ above S and below a contiguous maximum there is both an upper disk D_u and a lower disk D_l cut off by $S(z)$; these disks are parallel to subdisks of D_0. Normally, the existence of an upper and lower disk would not be particularly useful information, since the component γ_+ of $K \cup \gamma$ lying below S_+ and containing γ has four ends, so the arc of D_l incident to $K \cup \gamma$ might be a complicated arc on a neighborhood of γ_+. But here we know that $\partial\eta(\gamma_+) \cap \partial D_l \subset \beta$, so ∂D_l is disjoint from ∂E_0. We note that D_u is good.

In particular, if D_l is good, then β runs once across a meridian of γ and can be used to isotope γ up at the same time that we use D_u to bring a contiguous maximum down. Once this is done, we conclude as in Case 2. Suppose D_l is bad so the loop ∂D_l divides $S(z)$ into two disk components S_l and S_u, with $(\partial D_u \cap S(z)) \subset S_u$. Then it is not hard to use D_l to isotope γ up to a subarc of S_l and use D_u to isotope a maximum down to a subarc of S_u. Again we conclude as in Case 1 or 2.

Subcase 3c F_γ is a three-punctured sphere, whose boundary consists of ∂A_0 and the boundary of a single meridian of K.

Consider an outermost disk D_0 as above, with $\beta = \partial D_0 \cap F_\gamma$. Its ends can't lie on different components of ∂A_0 since these bound disjoint disk components of $S - A_0$. If one end lies on a component of ∂A_0 and the other on a meridian of K, then D_0 defines an isotopy of a contiguous maximum down to S, and we conclude as in Case 2. If both ends of β lie on a meridian of K, then β must be an arc that crosses a meridian of γ exactly once. This situation was dealt with in Lemma 5.13 above: A level sphere just above S cuts off a lower disk from D_0, so at some level there will be an upper disk and a lower disk that allow γ and its contiguous maximal arc to be pushed to the same level. Again, the conclusion follows from Case 2.

Finally, suppose both ends of β lie on the same component of ∂A_0. The argument is much as in that part of Subcase 3b in which both ends of β lie on ∂E_0: One can think of D_0 as a disk disjoint from K, lying above S, and crossing γ in exactly one point, a point in γ_a. (Here we drop the conceptual convention that $\gamma_a = \emptyset$.) Then ∂D_0 divides S into two disks; let S_0 be the one that does not contain γ_c. Then, as in Subcase 3b, all descending disks of maxima with ends in S_0 can be made disjoint from γ, first by removing all intersections of the descending disks with D_0 (thereby ensuring that the descending disks never intersect γ_c) then by using D_0 to pipe away any intersections with γ_a. Once the descending disks are disjoint from γ, the descending disk for the maximum of K contiguous to γ can be used to push the maximum to S, and the conclusion again follows from Case 2. □

7 A brief remark on tunnel number 1 links.

The central results here can easily be extended to tunnel number one links, though the trajectory of proof is a bit different. It is not true that the complement of a tunnel number link has no incompressible meridional planar surfaces. But if there is such a surface, then Gordon and Reid show there is one that cuts off a *Hopf tangle* (see [4, Figure 12]) and which has no more boundary components than the original one. Unknotting tunnels for such links have been extensively studied in [3].

In our context, the results of [4] and [3] imply that if L is a tunnel number one link and a minimal bridge position is not thin, then one way to put L into minimal bridge position is via decomposing it into the union of the n–string quasi Hopf tangle and the n–string trivial tangle. (The minimal bridge number of L is then $n+1$.) In particular, the link can be described as in (3) of Theorem 6.1, except that γ_c is here the unknotted component of the link.

If a minimal bridge position for the link L is thin, then we can proceed (as in the case of a tunnel number one knot) to deduce that not only is the link L in bridge position, but so is the graph $L \cup \gamma$; indeed, γ must lie in a level sphere.

Acknowledgements Part of this work was carried out while the first author was visiting at University of California, Davis. He would like to express sincere thanks to the department for its hospitality. The second author was partially supported by an NSF grant and the Miller Foundation. The third author was partially supported by an NSF grant.

References

[1] **J Berge**, *Embedding the exteriors of one-tunnel knots and links in the 3–Sphere*, unpublished preprint

[2] **D Gabai**, *Foliations and the topology of 3–manifolds: III*, J. Differential Geom. 26 (1987) 479–536

[3] **H Goda**, **M Ozawa**, **M Teragaito**, *On tangle decompositions of tunnel number one links*, J. Knot Theory and its Rami. 8 (1999) 299–320

[4] **C Gordon**, **A Reid**, *Tangle decompositions of tunnel number one knots and links*, J. Knot Theory and its Rami. 4 (1995) 389–409

[5] **T Kobayashi**, *Classification of unknotting tunnels for two bridge knots*, from: "Proceedings of the 1998 Kirbyfest", Geometry and Topology Monographs, 2 (1999) 259–290

[6] **K Morimoto**, *A note on unknotting tunnels for 2–bridge knots*, Bulletin of Faculty of Engineering Takushoku University, 3 (1992) 219–225

[7] **K Morimoto**, *Planar surfaces in a handlebody and a theorem of Gordon–Reid*, from: "Proc. Knots '96", (S Suzuki, editor) World Sci. Publ. Co. Singapore (1997) 127–146

[8] **A Thompson**, *Thin position and bridge number for knots in the 3–sphere*, Topology, 36 (1997) 505–507