Michigan Math. J. 56 (2008)

Conway Products and Links with Multiple Bridge Surfaces

MARTIN SCHARLEMANN & MAGGY TOMOVA

1. Introduction

A link K in a 3-manifold M is said to be in *bridge position* with respect to a Heegaard surface P for M if each arc of $K - P$ is parallel to P, in which case P is called a *bridge surface* for K in M. Given a link in bridge position with respect to P, it is easy to construct more complex bridge surfaces for K from P—for example, by stabilizing the Heegaard surface P or by perturbing K to introduce a minimum and an adjacent maximum. As with Heegaard splitting surfaces for a manifold, it is likely that most links have multiple bridge surfaces even apart from these simple operations. In an effort to understand how two bridge surfaces for the same link might compare, it seems reasonable to follow the program used in [RS] to compare distinct Heegaard splittings of the same non-Haken 3-manifold. The restriction to non-Haken manifolds ensured that the relevant Heegaard splittings were strongly irreducible. In our context the analogous condition is that the bridge surfaces are c-weakly incompressible (definition to follow). The natural analogy to the first step in [RS] would be to demonstrate that any two distinct c-weakly incompressible bridge surfaces for a link K in a closed orientable 3-manifold M can be isotoped so that their intersection consists of a nonempty collection of curves, each of which is essential (including nonmeridional) on both surfaces. In some sense the similar result in [RS] could then be thought of as the special case in which $K = \emptyset$.

Here we demonstrate that this is true when there are no incompressible Conway spheres for the knot K in M (cf. Section 4 and [GL]). In the presence of Conway spheres a slightly different outcome cannot be ruled out: the bridge surfaces each intersect a collar of a Conway sphere in a precise way; outside the collar the bridge surfaces intersect only in curves that are essential on both surfaces; and inside the collar there is inevitably a single circle intersection that is essential in one surface and meridional—and hence inessential—in the other.

2. Definitions and Notation

Suppose that K is a properly embedded collection of 1-manifolds in a compact manifold M. For X any subset of M, let X_K denote $X - K$. A disk that meets K

Received December 19, 2006. Revision received October 25, 2007.
Research of the first author is partially supported by an NSF grant.

transversally and at most once is called a (*punctured*) *disk*. Thus the parenthetical (punctured) means *either* unpunctured or with a single puncture.

Suppose F is a compact surface in M transverse to K. An isotopy of F_K will mean an isotopy of F in M, fixing K setwise, so that F is always transverse to K; in particular, it is a proper isotopy of F_K in M_K. A simple closed curve on F_K is *essential* if it doesn't bound a (punctured) disk on F_K. (In particular, a closed curve on F_K that bounds a once-punctured disk is considered to be inessential.) An embedded disk $D \subset M_K$ is a *compressing disk* for F_K if $D \cap F_K = \partial D$ and ∂D is an essential curve in F_K. A *cut-disk* for F_K (case (3) of [BSc, Def. 2.1]) is an embedded once-punctured disk D^c in M_K such that $D^c \cap F_K = \partial D^c$ and ∂D^c is an essential curve in F_K. A (punctured) disk that is either a cut disk or a compressing disk will be called a *c-disk* for F_K.

Any term describing the compressibility of a surface can be extended to account not only for compressing disks but for all c-disks. For example, we will call a surface *c-incompressible* if it has no c-disks. Let F be a splitting surface for M (i.e., M is the union of two 3-manifolds along F); then we call F_K *c-weakly incompressible* if any pair of c-disks for F_K on opposite sides of the surface intersect. If F_K is not c-weakly incompressible, it is *c-strongly compressible*. (Note the possible confusion with standard Heegaard surface terminology: a Heegaard splitting is weakly reducible (cf. [CG]) if and only if the Heegaard surface is strongly compressible.)

A properly embedded arc $(\delta, \partial\delta) \subset (F_K, \partial F_K)$ is *inessential* if there is a disk in F_K whose boundary is the endpoint union of δ and a subarc of ∂F; otherwise, δ is essential. The surface F_K is ∂-compressible if there is a disk $D \subset M$ such that the boundary of D is the endpoint union of two arcs, $\delta = D \cap F_K$ (an essential arc in F_K) and $\beta = D \cap \partial M$. Note that arcs in F with one or more endpoints at $K \cap F$ do not arise in either context, though such arcs will play a role in our argument. If F is a properly embedded twice-punctured disk in M then a ∂-compression of F_K may create two c-compressing disks when there were none before, in the same way that ∂-compressing a properly embedded annulus may create a compressing disk.

For X any compact manifold, let $|X|$ denote the number of components of X. For example, if S and F are transverse compact surfaces then $|S \cap F|$ is the total number of arcs and circles in which F and S intersect.

3. Bridges and Bridge Surfaces

A properly embedded collection of arcs $T = \bigcup_{i=1}^{n} \alpha_i$ in a compact 3-manifold is called *boundary parallel* if there is an embedded collection $E = \bigcup_{i=1}^{n} E_i$ of disks such that, for each $1 \leq i \leq n$, ∂E_i is the endpoint union of α_i and an arc in the boundary of the 3-manifold. If the manifold is a handlebody A, the arcs are called *bridges* and disks of parallelism are called *bridge disks.*

LEMMA 3.1. *Let A be a handlebody and let $(T, \partial T) \subset (A, \partial A)$ be a collection of bridges in A. Suppose F is a properly embedded surface in A transverse to T that is not a union of (punctured) disks and twice-punctured spheres. If F_T is incompressible in A_T, then $\partial F \neq \emptyset$ and F_T is ∂-compressible.*

Proof. Suppose D is a (punctured) disk component of F or a twice-punctured sphere and let $F' = F - D$. Suppose also there is a compressing disk (resp. ∂-compressing disk) for F'_T. Then a standard cut-and-paste argument shows that there is a compressing disk (resp. ∂-compressing disk) for F'_T that is disjoint from D. So with no loss of generality we may henceforth assume that *no* component of F is a (punctured) disk or twice-punctured sphere.

A standard cut-and-paste argument provides a complete collection of meridian disks for A that is disjoint from a complete collection of bridge disks for T. Let the family Δ of disks be the union of the two collections; in particular, $A - \eta(\Delta)$ is a collection of 3-balls. In fact, choose such a collection, transverse to F_T, so that $|F_T \cap \Delta|$ is minimal. Now Δ cannot be disjoint from F_T, for then F_T would be an incompressible surface in one of the ball components of $A - \eta(\Delta)$ and so a collection of disks, contrary to our assumption. If any component of $F_T \cap \Delta$ were a closed curve then an innermost one on Δ would be inessential in F_T, since F_T is incompressible, and an innermost inessential curve of intersection in F_T (and perhaps more curves of intersection) could be eliminated by rechoosing Δ. We conclude that each component of $F_T \cap \Delta$ is an arc.

Each arc in $F_T \cap \Delta$ can have ends on ∂F, ends on T, or one end on each. A similar cut-and-paste argument shows that, if any arc has both ends on ∂F, then an outermost such arc in Δ would be essential in F_T and so the disk it cuts off in Δ would be a ∂-compressing disk for F_T, as required. Suppose then that all arcs of intersection have at least one end on T. If any arc had both ends on T, then a regular neighborhood of a disk cut off by an outermost such arc in Δ would contain a compressing disk for F_T in its boundary, contradicting the hypothesis. If all arcs of intersection have one end on T and the other on ∂F_T, then a regular neighborhood of a disk cut off by an outermost such arc in Δ would contain a ∂-compressing disk for F_T in its boundary, as required. $\square$

Suppose M is an irreducible 3-manifold and T is a properly embedded collection of arcs in M transverse to a 2-sphere S. Since S bounds a 3-ball it is separating, so $|S \cap T|$ is even. In particular, if $|S \cap T| = 2$ then the core of the annulus S_T cannot bound a disk in M_T, for if it did the result would be a sphere in M intersecting T in a single point. When A is a handlebody and $T \subset A$ is a collection of bridges, more can be said as follows.

LEMMA 3.2. *Suppose A is a handlebody, T is a collection of bridges in A, and $S \subset T$ is a sphere that intersects the link exactly twice transversally. Then T intersects the ball bounded by S in an unknotted arc* (*i.e., a bridge in the ball*).

Proof. We have seen that a 2-sphere in A intersects each bridge in an even number of points, so in particular S intersects exactly one bridge α. Let E be the bridge disk for α. Since the core of the annulus S_T can bound no disk in M_T, a standard innermost disk argument allows E to be chosen so that $S \cap E$ contains no closed curves at all. Any arc of intersection between S and E must have both of its endpoints on α as $S \cap \partial A = \emptyset$. Hence there must be exactly one such arc of intersection β, cutting off a subarc α' of α on the opposite side of S from $\partial\alpha$.

Therefore, α' lies in the ball that S bounds in A. The subdisk of E cut off by β is a bridge disk for α' in that ball. □

COROLLARY 3.3. *Suppose A is a handlebody, T is a collection of bridges in A, and D is a properly embedded (punctured) disk in A whose boundary is inessential in A_K. Then D_K is isotopic, with respect to boundary, to a (punctured) disk D'_K in ∂A. The disk D_K is punctured if and only if D'_K is punctured.*

Proof. Apply Lemma 3.2 to the sphere $D \cup_{\partial D} D'$. □

Suppose M is a closed 3-manifold containing a link K. A Heegaard surface P for M is a *bridge surface* for K in M if K intersects each of the two complementary handlebodies of P in M in a collection of bridges.

Toward explaining how two possibly different bridge surfaces are related, Lemma 3.2 has the following pleasant corollary.

COROLLARY 3.4. *Suppose bridge surfaces P and Q intersect in a collection of curves such that any curve that is inessential in P is also inessential in Q. Then Q can be properly isotoped, without adding curves of intersection and without removing any curve of intersection that is essential in both surfaces, until all curves of intersection are essential in P.*

Proof. The proof is by induction on the number of curves of intersection that are inessential in P. If no such curves exist then there is nothing to prove.

Among curves of intersection that are inessential in P, let γ be innermost in P. By hypothesis, γ is also inessential in Q; let D^P and D^Q be the (punctured) disks that γ bounds in P and Q, respectively. Then $S = D^P \cup_\gamma D^Q$ can be pushed slightly to be a sphere in one of the complementary handlebodies, say X, of Q. Since X is irreducible, the sphere S bounds a ball $\mathcal{B}$ in X. Since K intersects each of D^P and D^Q in at most one point, it follows that K is either disjoint from $\mathcal{B}$ or, following Lemma 3.2, intersects $\mathcal{B}$ in a single bridge. In the latter case, a bridge disk for $\mathcal{B}$ can be isotoped so that it intersects γ in a single point. Thus, in either case, $\mathcal{B}$ defines a proper isotopy of D^Q to D^P keeping γ fixed. Pushing D^Q a bit beyond D^P removes γ as a curve of intersection. This completes the inductive step, once we show that any other curve of intersection that is removed by the isotopy is inessential in Q. But since the interior of $\mathcal{B}$ is entirely disjoint from Q, the only other curves of intersection removed are those in $D^Q \cap P$. Since these lie in D^Q, they are inessential in Q. □

Suppose both P and Q are bridge surfaces for a link K, so that $M = A \cup_P B = X \cup_Q Y$ and the link K is in bridge position with respect to both P and Q. That is, $K \cap A$, $K \cap B$, $K \cap X$, and $K \cap Y$ are all collections of bridges in the respective handlebodies.

LEMMA 3.5. *Suppose that $Q \subset A$ and that P_K compresses in $A_K - Q$. Then either K is the unknot in S^3 or P_K is strongly compressible.*

REMARK. In fact, we will show (Lemma 5.1) that the second alternative holds unless K is in 1-bridge position with respect to $P \cong S^2$.

Proof of Lemma 3.5. With no loss of generality, suppose $P \subset X$. Then $A \cap X$ is a cobordism between P and Q, and the compressing disk for P_K in $A_K - Q$ lies in $(A \cap X)_K$ (see Figure 1). Let P' be the surface obtained from P by maximally compressing P_K in $(A \cap X)_K$. Since P' is a closed surface in the handlebody X, it follows from Lemma 3.1 that either P'_K compresses in X_K or P' is the union of twice-punctured spheres. Moreover, by construction, P' separates Q from P.

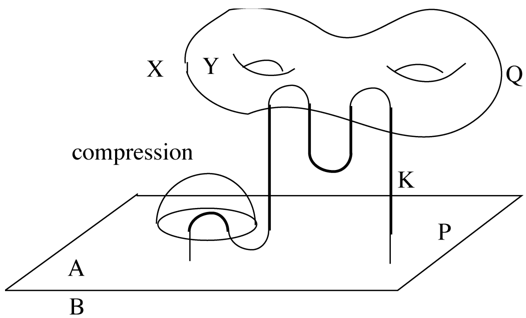

Figure 1

Suppose first that P' is a union of twice-punctured spheres. Since P' separates Q from P, some component P'_0 of P' separates Q from P. (Any path from P to Q intersects P' an odd number of times, so it must intersect some component an odd number of times.) Since $P'_0 \subset A$ and A is irreducible, it follows that P'_0 bounds a ball on the side of P'_0 not containing P—namely, the side containing X. By Lemma 3.2, K intersects that ball in a single unknotted arc. Symmetrically, since $P'_0 \subset X$ and X is irreducible, P'_0 bounds a ball on the side of P'_0 not containing Q—namely, the side containing A. By Lemma 3.2, K intersects that ball in a single unknotted arc. Thus P'_0 bounds a ball on both sides, so M is a 3-sphere and K intersects each side of P'_0 in a single unknotted arc. Hence P'_0 is a 1-bridge sphere for K, which is the unknot in $M = S^3$.

Now suppose instead that P'_K compresses in X_K. By construction, any such compressing disk must lie on the side of P'_K opposite to Q. Denote that side B' because it contains B. In fact, B' is obtained from B by attaching the 2-handles determined by the maximal compression of P_K in $(A \cap X)_K$. Dually, B is obtained from B' by deleting a neighborhood of $P' \cup \Gamma$, where the graph Γ consists of the arcs that are co-cores of these 2-handles and where, by construction, $P' = \partial B'$. Choose Γ (up to slides and isotopies in $P' \cup \Gamma$) and choose the compressing disk $D \subset B'_K$ so as to minimize the number of points $|\Gamma \cap D|$ (see Figure 2). If D is disjoint from Γ, then D lies in B_K and is disjoint from the compressing disks in A_K that are dual to the arcs of Γ, exhibiting a strong compression of P_K.

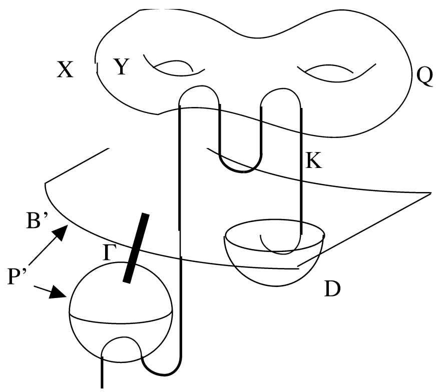

Figure 2

So we henceforth assume that Γ intersects D. Let $\Delta \subset B$ be a complete collection of meridian disks and bridge disks chosen to minimize $|\Delta \cap D|$. If Δ is a single bridge disk then K is the unknot in S^3, as required. So assume hereafter that Δ is more complicated. In particular: within Δ, each disk (or the boundary of its regular neighborhood) is a compressing disk for B_K.

Following [ST, Cor. 2.3], whose proof we now briefly recapitulate, consider the graph Υ in D whose vertices are the points $\Gamma \cap D$ and whose edges are the arc components of $\Delta \cap D$. As in that proof, the minimization of $|\Delta \cap D|$ guarantees that no closed component of $\Delta \cap D$ can bound a disk in the complement of Γ and no loop in Υ cuts off a disk in D that is disjoint from Γ. Hence there is a vertex v of Υ that is incident only to simple edges, perhaps because it is isolated. If v is not isolated then, of all the arcs incident to v, one that is outermost in Δ describes a way to slide the edge of Γ containing v to remove v from $\Gamma \cap D$. This would violate our original minimization of $|\Gamma \cap D|$. Hence v is isolated, so Δ is disjoint from the compressing disk in A_K dual to the edge containing v, exhibiting again a strong compression of P_K. □

LEMMA 3.6. *Suppose that $Q \subset A$ and that P_K c-compresses in $A_K - Q$. Then either K is the unknot in S^3 or P_K is c-strongly compressible.*

Proof. Let D be a c-disk for $A_K - Q$ in P_K. If D is disjoint from K then the result follows from Lemma 3.5, so we may as well assume D is a once-punctured disk. As before, assume without loss of generality that $P \subset X$ and so $D \subset X$. Let $\alpha \subset X$ be the bridge for Q that intersects D and let E be a bridge disk for α, so ∂E is the endpoint union of α and an arc δ in Q. Choose E to minimize $|D \cap E|$. If $D \cap E$ contains a closed curve then an innermost such curve c in D cannot bound a once-punctured disk in D, else the union of this disk together with the disk that c bounds in E would be a once-punctured sphere in X, which is impossible. Nor can c bound a disk in D_K, for otherwise c could be removed by a different choice of E. Any arc component of $D \cap E$ with both endpoints on ∂D can be removed by rechoosing E. We conclude that $D \cap E$ consists of a single arc, with one end at the point $K \cap D$ and the other end at a point $p \in P \cap \partial D$.

Now, retaining the requirement that $E \cap D$ be a single arc, choose E to minimize $|P \cap E|$. Observe that, among the entire collection of curves $P \cap E$, only one curve intersects ∂D and this intersection is the single point $\{p\}$, since $\partial D \cap E = \{p\}$ (see Figure 3). One possibility is that there is a closed curve c of intersection that bounds a disk in P_K. (Such a curve would necessarily intersect ∂D in an even number of points and so, in particular, c does not contain p.) An innermost such curve in P_K could be eliminated by a rechoice of E. It follows that no closed curve of intersection bounds a disk in P_K. If a closed curve of intersection $c \subset E \cap P$ that is innermost in E bounds a disk in A_K, then the result follows from Lemma 3.5. If c bounds a disk D' in B_K and does not contain p, then D and D' are the disjoint disks that are sought. If c does contain p then the boundaries of D and D' intersect precisely in the single point p. The boundary of a regular neighborhood of $\partial D \cup \partial D'$ in P is essential in P_K (since $|K \cap P| \geq 2$) and also bounds a disk D'' in B_K—namely, two copies of D' banded together along ∂D. Then D (possibly punctured) and D'' are the disjoint disks that are sought. So henceforth we assume that there are no closed curves in $P \cap E$.

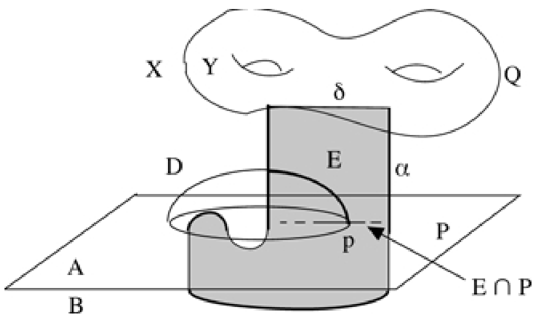

Figure 3

Next consider the arcs of intersection in $P \cap E$. Then ∂E is the endpoint union of two arcs, α a component of $K - Q$ and δ an arc in Q. Since $Q \subset A$, it follows that ∂E intersects P only in points on α. An outermost arc of intersection β in E then cuts off a bridge disk E' for a bridge $\kappa \subset \alpha$ for K with respect to P. In particular, a regular neighborhood of E' contains in its boundary a compressing disk D' for P_K. If $D' \subset A$ then we are done by Lemma 3.5. If $D' \subset B$ and β does not contain p, then D' and D are the required disjoint c-disks for P_K. The only remaining possibility is that the only outermost arc of intersection β contains p and the disk E' that β cuts off lies in B_K. In this case, consider the arc $\gamma = D \cap E \subset A_K$ with one end on α and the other end on p. Then the union of γ and a subarc β' of β cuts off a subdisk $E'' \subset A$ of $E - E'$ (see Figure 4). The union of D and E'' along γ has a regular neighborhood in A consisting of two disks, one parallel to D in A_K and the other, D'', parallel to D in A but disjoint from K.

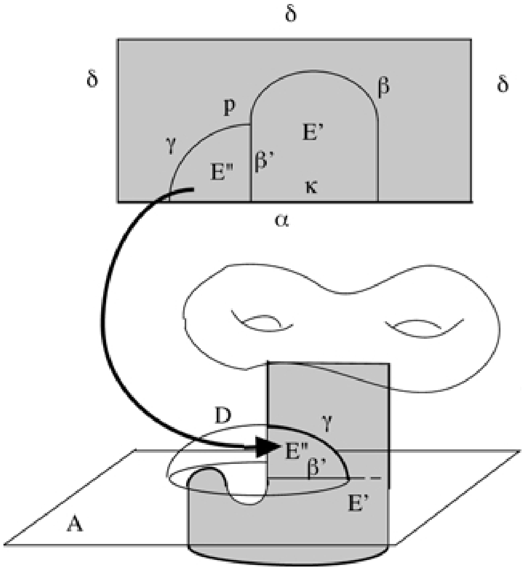

Figure 4

We know that $\partial D''$ must be essential in P_K, for otherwise it would bound a disk in P_K and then ∂D would bound a once-punctured disk in P_K, contradicting the assumption that D is a c-disk. Hence D'' is a compressing disk for P_K and the result once again follows from Lemma 3.5. □

To allow inessential curves in $P_K \cap Q_K$, the hypotheses are weakened further in our next lemma.

LEMMA 3.7. *Suppose*:

- *every curve in $P_K \cap Q_K$ is inessential in both P_K and Q_K*;
- *$Q_K \cap A$ contains some curve that is essential in Q_K*; *and*
- *P_K c-compresses in $A_K - Q$.*

Then either K is the unknot in S^3 or P_K is c-strongly compressible.

Proof. The proof is by induction on $|P_K \cap Q_K|$. If P_K and Q_K are disjoint then the result follows from Lemma 3.6. (Note that if Q_K contains no essential curves then Q_K is a twice-punctured sphere, so K is the unknot in S^3.) Suppose, for the inductive step, that $|P_K \cap Q_K| \geq 1$ and c is a circle of intersection that is innermost in P. Let $D \subset P$ be the (punctured) disk that c bounds in P and let E be the (punctured) disk that c bounds in Q. Clearly the c-disk for P_K disjoint from Q_K is disjoint from D. Although the interior of E may intersect P, the interior of D is disjoint from Q and so, after a slight push on E, $S = D \cap_c E$ is a sphere that is disjoint from Q. Hence S bounds a ball in X or Y and S itself is either disjoint from K or punctured twice. Therefore, S bounds either a ball in X_K or (according to Lemma 3.2) a ball in X that K intersects in an unknotted arc. In either case, the ball describes a proper isotopy of Q_K, stationary away from E, that replaces E with D. After a further small push, c is removed as a curve of intersection. That is, $|P_K \cap Q_K|$ has been reduced by at least one. Any essential curve in Q_K that is disjoint from P_K is clearly unaffected by this isotopy, because the curve must be disjoint from E and so the curve remains in A. Furthermore, the P_K compressing disk in A_K disjoint from Q remains disjoint from Q, since it was disjoint from D. This completes the inductive step. $\square$

4. Conway Spheres and Bridge Position

DEFINITION 4.1. A *Conway sphere* [Co] S for a link K in a 3-manifold M is a sphere $S \subset M$ transverse to K such that $|K \cap S| = 4$. A Conway sphere is an *incompressible* Conway sphere if S_K is incompressible in M_K.

Suppose M_0, M_1 are orientable 3-manifolds containing links K_0, K_1 respectively. For each $i = 0, 1$, let τ_i be an arc in M_i whose ends lie on K_i but that is otherwise disjoint from K_i. Let $\mathcal{B}_i$ be a regular neighborhood of τ_i, a ball intersecting K_i in two arcs, one near either end of τ_i. Then $S_i = \partial\mathcal{B}_i$ is a Conway sphere for K_i in M_i. The arcs $K_i \cap B_i$ are parallel in $\mathcal{B}_i$ to arcs τ_i^e and τ_i^w in S_i. Let τ_i^n and τ_i^s denote a pair of arcs in each S_i that, together with τ_i^e and τ_i^w, form an embedded circle in S_i (see Figure 5).

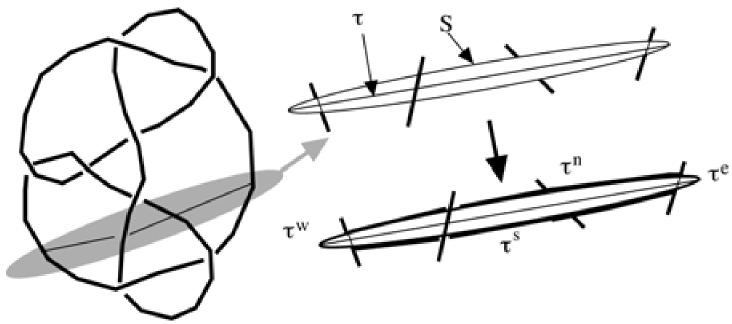

Figure 5

DEFINITION 4.2. Given τ_i as just described, let $K_0 +_c K_1$, called a *Conway sum* of the K_i, denote a link in $M_0 \# M_1$ obtained by removing the interior of $\mathcal{B}_i$ from each M_i and gluing S_0 to S_1 via a homeomorphism that identifies the pair of arcs τ_0^e, τ_0^w with the arcs τ_1^e, τ_1^w.

Similarly, let $K_0 \times_c K_1$, called a *Conway product,* denote a link in $M_0 \# M_1$ obtained by instead gluing S_0 to S_1 via a homeomorphism that identifies the pair of arcs τ_0^e, τ_0^w with the arcs τ_1^n, τ_1^s and the pair of arcs τ_0^n, τ_0^s with the arcs τ_1^e, τ_1^w.

The image S of S_0 and S_1 after their identification is called the *Conway sphere* of the sum (or product).

NOTE. Essentially the same constructions can be done for disjoint arcs τ_0, τ_1 that are contained in the same manifold M and whose ends are on the same link $K \subset M$. In that case, choose the identification $S_0 \cong S_1$ to be orientation reversing so that the resulting manifold is $M \# S^1 \times S^2$. The sum and product links in $M \# S^1 \times S^2$ respectively are denoted $K +_c$ and $K \times_c$, and in this case the Conway sphere of the sum or product is nonseparating.

The definition of Conway sum and product is motivated by analogy to tangle sum and tangle product (cf. [A, pp. 47–48]). Unlike the standard construction of a connected sum of links, a Conway sum or Conway product of links depends on many choices beyond the question of which components of K_i contain the ends of τ_i. Most prominent is the choice of the arcs τ_i defining the operation, but there is also some choice in how S_0 is identified to S_1 beyond the constraints given by the definitions.

In general, one would expect little connection between bridge presentations of K_i in M_i and bridge presentations of $K_0 +_c K_1$ and $K_0 \times_c K_1$ in $M_0 \# M_1$ (or $K \times_c$ in $M \# S^1 \times S^2$). The most obvious problem is that each τ_i may intersect a bridge surface P_i for K_i an unknown number of times (e.g., perhaps $|\tau_0 \cap P_0| \neq |\tau_1 \cap P_1|$), and there is no way of cobbling together the complementary punctured bridge surfaces into a plausible bridge surface for the resulting link. In the case of Conway sum, the problem can be alleviated by limiting the operation to arcs τ_i that lie in the bridge surface P_i and by requiring that the equator curve $S_0 \cap P_0$ be identified with the equator curve $S_1 \cap P_1$. Then the resulting surface $P_0 \# P_1$ is the standard connected sum of Heegaard surfaces for M_0 and M_1, so $P_0 \# P_1$ is a Heegaard surface for $M_0 \# M_1$ and thus a potential bridge surface for $K_0 +_c K_1$. But even in this case, if the arcs τ_i intersect bridge disks for K_i then there is no natural reason why $K_0 +_c K_1$ should be in bridge position with respect to $P_0 \# P_1$. Hence one expects that Conway sums do not in general behave well with respect to bridge number.

The situation is more hopeful for Conway products, to which we now turn.

DEFINITION 4.3. For $i = 0, 1$, suppose that P_i is a bridge surface for the link K_i in M_i and that τ_i is an arc in P_i intersecting K_i exactly in the endpoints of τ_i. Form a Conway product by taking τ_i^n, τ_i^s to be arcs in S_i disjoint from the equator ∂P_i. The result is called a Conway product that *respects the bridge surfaces.* The sphere S is called a *Conway decomposing sphere* for the pair $(P_0 \# P_1, K_0 \times_c K_1)$.

The same terminology is used when P is a bridge surface for the link K in M; τ_0 and τ_1 are disjoint arcs in P intersecting K exactly in their endpoints; and a Conway product is formed by taking τ_i^n, τ_i^s to be arcs in S_i disjoint from the equator ∂P_i. Then S is a Conway decomposing sphere for the pair $(P \# (S^1 \times S^1), K \times_c)$.

The reason for regarding S as a decomposing sphere for the pair is as follows.

PROPOSITION 4.4. *Let $K_0 \times_c K_1$ be a Conway product respecting the bridge surfaces P_i for K_i in M_i $(i = 0, 1)$. Then $P_0 \# P_1$ is a bridge surface for $K_0 \times_c K_1$ in $M_0 \# M_1$. Moreover, the bridge numbers satisfy*

$$\beta(K_0 \times_c K_1) \leq \beta(K_0) + \beta(K_1) - 1.$$

Similarly, if $K \times_c$ is a Conway product respecting the bridge surface P for K in M, then $P \# (S^1 \times S^1)$ is a bridge surface for $K \times_c$ in $M \# (S^1 \times S^2)$ and the bridge number satisfies $\beta(K \times_c) \leq \beta(K) - 1$.

Proof. We consider only the first case, when the Conway product $K = K_0 \times_c K_1$ is of links in two different manifolds; the proof for a product $K \times_c$ is similar.

Examine how the bridge surfaces P_0 and P_1 intersect S_0 once S_0 and S_1 are identified. Each P_i intersects S_0 in a single closed curve $c_i = S_0 \cap P_i$. Since τ_0 lies in P_0, it follows that c_0 intersects each of τ_0^e, τ_0^w in a single point and, by hypothesis, is disjoint from τ_0^n, τ_0^s. Similarly c_1 intersects each of τ_0^n, τ_0^s in a single point and is disjoint from each of τ_0^e, τ_0^w. It follows that the c_i can be isotoped in $S_0 = S_1 = S$ with respect to the circle $\bar{\tau} = \tau_0^e \cup \tau_0^n \cup \tau_0^w \cup \tau_0^s$ in such a way that (a) $c_0 \cap c_1$ consists of a single point in each of the disks $S - \bar{\tau}$ and (b) together the curves $c_0 \cup c_1$ divide S into quadrants, each containing a single puncture. The three curves $c_0 \cup c_1 \cup \bar{\tau}$ divide S into octants (see Figure 6).

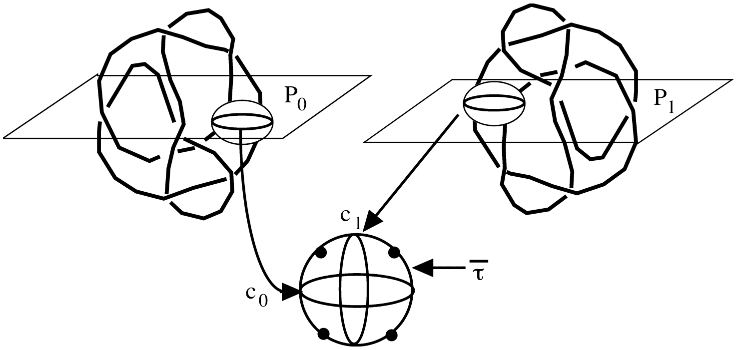

Figure 6

With no loss of generality, assume that $\bar{\tau}$ and the c_i are mutually orthogonal great circles of S. Let $\rho_t \colon S^1 \to S$ be an isotopy of c_0 that rotates c_0 through an angle of $\pi/2$ to c_1 around the two fixed points $c_0 \cap c_1$. Thus $\rho_i(S^1) = c_i$ for $i = 0, 1$. During the isotopy, the image of ρ_t will cross exactly two of the four punctures $S \cap K$. (There are two choices for such an isotopy, one for each direction of rotation; opposite choices will carry c_0 to anti-parallel copies of c_1 and will cross the complementary pair of punctures.)

Let $S \times I$ be a collar of S in $M_0 \# M_1$ and use it to consider an alternate construction of $M_0 \# M_1$—namely, identify $S_i = \partial M_i$ with $S \times \{i\} \subset S \times I$ for $i = 0, 1$. Connect $\partial(P_0 - \mathcal{B}_0) = c_0$ to $\partial(P_1 - \mathcal{B}_1) = c_1$ in $S \times I$ via the embedded annulus $S^1 \times I \subset S \times I$ given by $(x, t) \to (\rho_t(x), t)$ and connect the four points $\partial(K_0 - \mathcal{B}_0)$ to $\partial(K_1 - \mathcal{B}_1)$ in $S \times I$ by the four arcs $(K \cap S) \times I$ (see Figure 7). The result is a specific embedding of the surface $P_0 \# P_1$ and K in $M_0 \# M_1$ such that $|(P_0 \# P_1) \cap K| = |P_0 \cap K_0| + |P_1 \cap K_1| - 2$. Thus it remains only to show that $P_0 \# P_1$ is a bridge surface for K in $M_0 \# M_1$. Part of this is easy: as observed previously, $P_0 \# P_1$ is the standard connected sum of the Heegaard surfaces P_i for M_i, so it is a Heegaard surface $M_0 \# M_1$. The remaining problem is then to exhibit

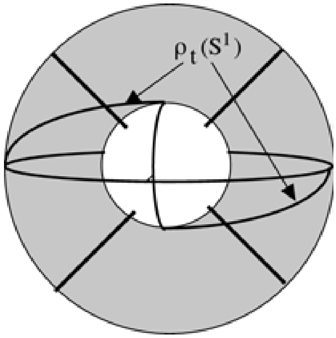

Figure 7

bridge disks for each arc component of $K - (P_0 \# P_1)$. This requires a more concrete description, as follows.

We follow convention in the theory of tangles (see [GL]) and label the four points of $K \cap S$ according to the four quadrants in which we imagine they lie: NW, NE, SE, SW. Assume without loss of generality that c_0 is horizontal, separating the northern hemisphere from the southern and so separating the pair NW, NE from the pair SE, SW. Similarly assume that c_1 is vertical, separating the western hemisphere from the eastern and so separating the pair NW, SW from the pair NE, SE. Further assume that ρ_t isotopes c_0 across the points NW and SE to c_1.

In order to exhibit bridge disks, consider the types of bridges (i.e., arc components of $K - (P_0 \# P_1)$) that can arise. First consider a bridge that is disjoint from $S \times I$—for example, a component α of $K_0 - P_0$ that is disjoint from S_0. Let E be a bridge disk for α with respect to the original splitting surface P_0. By general position, ∂E intersects the original arc $t_0 \subset P_0$ transversally in a number of points. Corresponding to each point is a half-meridian disk (denoted μ_0 in Figure 8) in which E intersects $\mathcal{B}_0 - P_0$. Thus E intersects S_0 in a collection of parallel arcs in one of the twice-punctured hemisphere components (say the northern hemisphere) of $S_0 - P_0 \cong S - c_0$. Push all these arcs $E \cap (S_0 - K)$ to the west of c_1 in the northern hemisphere. The isotopy ρ_t sweeps c_0 entirely across all the

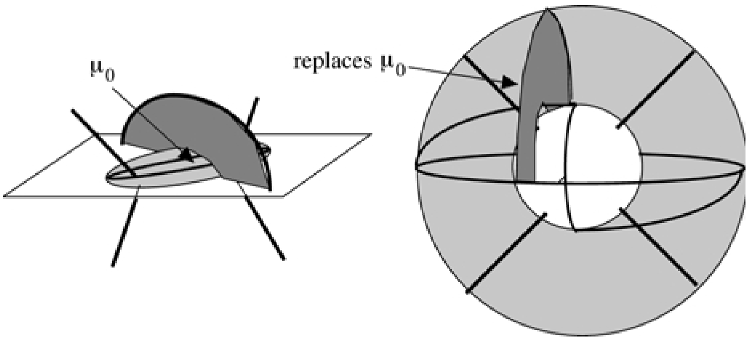

Figure 8

arcs $E \cap (S_0 - K)$ and, in doing so, defines a collection of disks that replace the half-meridians of τ_0 to give a bridge disk for α.

The argument is little different for a bridge α of $K - (P_0 \# P_1)$ that contains exactly one of the four points NW $\times\, \partial I$ or SE $\times\, \partial I$, say NW $\times \{0\} \in S \times \{0\}$. Let E be the bridge disk in $M_0 - P_0$ for the bridge α_0 that contains the arc τ_0. Then the northern hemisphere of $S_0 - P_0$ cuts off from E a collection of half-meridians of τ_0 just as before, together with a disk whose boundary consists of three arcs: an arc lying on $P_0 \cap \mathcal{B}_0$, the segment $\alpha_0 \cap \mathcal{B}_0$, and an arc γ from NW to c_0 in $S \times \{0\}$ that we may as well take to be part of τ_0^w. But the sweep of c_0 across γ defines a disk that can be used (together with the half-meridians of the previous case) to complete $E - \mathcal{B}_0$ to give a bridge disk for α (see Figure 9).

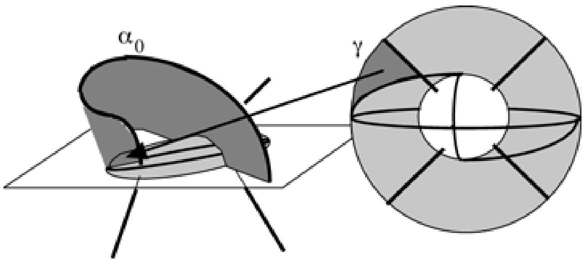

Figure 9

Next consider the case when α intersects $S \times \{0\}$ in exactly one point (say, NE $\times \{0\}$) that the isotopy ρ_t does *not* sweep across. Then all of the arc NE $\times\, I \subset S \times I$ lies in α. Assuming (for the moment) that α otherwise does not intersect $S \times \partial I$, we have that α is the union of NE $\times\, I$ and two other segments, α_0 and α_1, where $\alpha_i = \alpha \cap (M_i - \mathcal{B}_i)$. Let E_i be a bridge disk for the bridge in $K_i - P_i$ that contains α_i. Much as before, we can arrange that E_0 intersects S in a collection of arcs (with both ends on c_0) in the western half of the northern hemisphere, together with a single arc γ connecting NE to c_0. It will be useful in this phase to take for γ an arc whose other end is at one of the two rotation points $c_0 \cap c_1$ (see Figure 10). Likewise, E_1 intersects S in a collection of arcs (with both ends

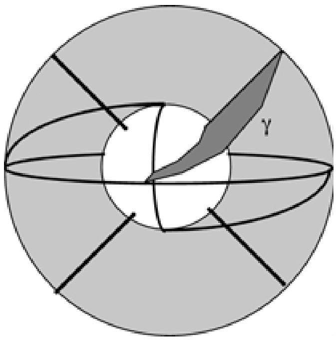

Figure 10

on c_1) in the eastern half of the southern hemisphere, together with an arc (e.g. γ) connecting NE to c_1. This arc γ is unaffected by the isotopy of c_0 to c_1, so $\gamma \times I$ attaches $E_0 - \mathcal{B}_0$ to $E_1 - \mathcal{B}_1$ and so creates a single bridge disk for α.

The two other points of $S \times \partial I$ that the bridge α containing NE $\times\, I$ could also contain are NW $\times \{0\}$ and SE $\times \{1\}$. In this case, the bridge disk for α is assembled by combining the previous arguments. $\square$

It is natural to ask whether the inequality in Proposition 4.4 could be an equality, just as Schubert showed the analogous equality for a connected sum. (See [Schu] or [Sch] for a modern proof.) R. Blair pointed out an example in S^3 where equality does not hold. The factor knots are 3-bridge knots incorporating rational tangles that cancel when the product is constructed. The resulting Conway product is a 4-bridge link (see Figure 11).

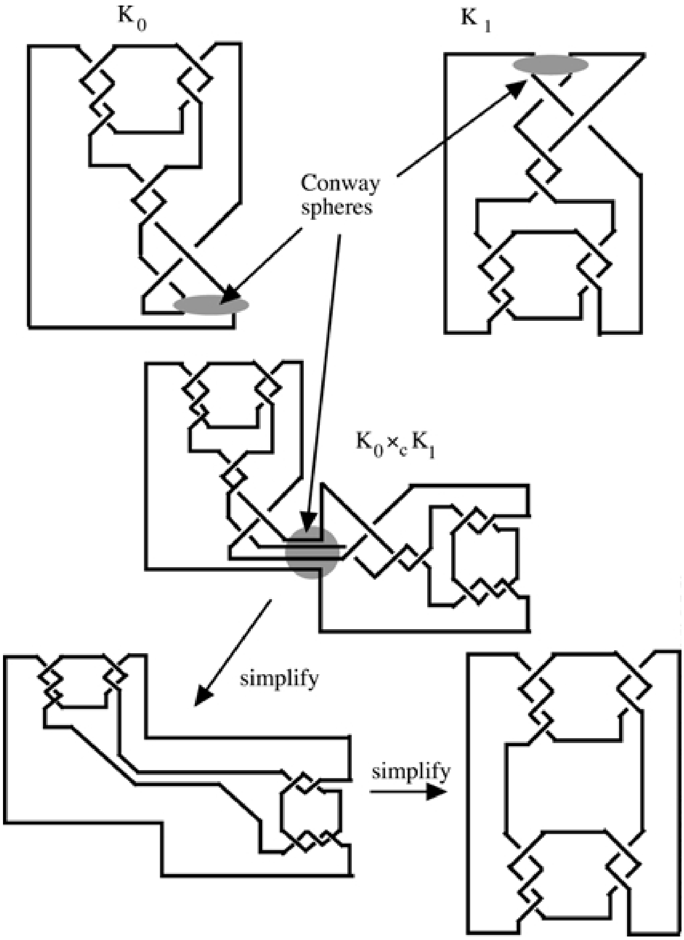

Figure 11

5. Spines and Sweep-outs

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

Suppose that a link $K \subset M$ is in bridge position with respect to a Heegaard surface P for M. Then the closed complementary components of P are handlebodies A and B that K intersects in a collection of bridges. Let $\Sigma_{(A,K)}$ (resp. $\Sigma_{(B,K)}$) denote a spine in A (resp. B) for $K \cap A$ (resp. $K \cap B$). Then, following our previous remarks, there exists 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 H gives a mapping cylinder structure to a neighborhood of $\Sigma_{(A,K)} \cup \Sigma_{(B,K)}$. Little is lost and some brevity gained if we restrict H to $P_K \times (I, \partial I) \to (M_K, \Sigma_{(A,K)} \cup \Sigma_{(B,K)})$. Then H is called a *sweep-out* associated to P.

As a warm-up, here is a classical application of sweep-outs. In effect it strengthens somewhat the conclusions of Lemmas 3.5–3.7 for it shows, in those lemmas, that if P_K is c-weakly incompressible then K is not only the unknot in S^3 but also in 1-bridge position with respect to P.

LEMMA 5.1. *Suppose $K \subset S^3$ is the unknot and is in bridge position with respect to P. Then: either P is a sphere and K is the unknot in 1-bridge position with respect to P, or P_K is strongly compressible.*

Proof. We will assume that either the number of bridges $m \geq 2$ or $P \neq S^2$, for otherwise K is 1-bridge with respect to a sphere P and we are done. In particular, we can assume that the neighborhood of any bridge disk for K contains a compressing disk for P_K in S^3_K.

Choose spines $\Sigma_{(A,K)}$ and $\Sigma_{(B,K)}$ as described before. In particular, the end points of $\Sigma_{(A,K)}$ (resp. $\Sigma_{(B,K)}$) are just the collection of m maxima (resp. minima) of K with respect to a sweep-out $H\colon P_K \times (I, \partial I) \to (S^3_K, \Sigma_{(A,K)} \cup \Sigma_{(B,K)})$. Let D be the disk that K bounds. For ε very small, $P^\varepsilon_K = H(P_K \times \{\varepsilon\})$ is the boundary of a regular neighborhood of $\Sigma_{(B,K)}$. By transversality (of D with $\Sigma_{(B,K)}$), the curves of intersection $D \cap P^\varepsilon_K$ consist of a family of m arcs (the family is ∂-parallel in D, and each corresponds to a minimum of K) and an unnested collection of simple closed curves, each corresponding to a point in $D \cap \Sigma_{(B,K)}$. Each arc and simple closed curve cuts off a disk from D; the former are bridge disks for $\Sigma_{(B,K)}$ and the latter are compressing disks for P_K in B_K.

Similar comments hold for $P_K^{1-\varepsilon} = H(P_K \times \{1-\varepsilon\})$, except all the subdisks of D that are cut off lie in A_K. Now consider the intersection with D of a generic $P_K^t = H(P_K \times \{t\})$. Any circle or arc in D is clearly inessential in D, and P_K^t necessarily intersects D in m arcs. After all circles in $D \cap P_K^t$ that are inessential in P_K^t are removed by an isotopy of D, either an innermost circle of intersection or an outermost arc of intersection cuts off a subdisk of D that is either a bridge disk or a compressing disk for P_K^t. Furthermore, the neighborhood of a bridge disk contains a compressing disk for P_K^t. We thus conclude that, for any generic t, there is some curve of intersection that defines a compressing disk for P_K^t lying either in A_K or in B_K.

Since for small t such a compressing disk lies in B^K and for large t such a disk lies in A_K and since there is always some compressing disk, it follows that there is a (possibly nongeneric) t at which there are compressing disks for P_K both in A_K and B_K. Since these two disks are defined by disjoint arcs of intersection, the disks themselves will be disjoint and so define a strong compression of P_K. (A short and standard argument shows this to be true even when t is nongeneric; see Lemma 5.4.) $\square$

Suppose both P and Q are possibly different bridge surfaces for the link, so that $M = A \cup_P B = X \cup_Q Y$ and the link K is in bridge position with respect to both P and Q. Choose spines $\Sigma_{(A,K)}, \Sigma_{(B,K)}, \Sigma_{(X,K)}, \Sigma_{(Y,K)}$ in each handlebody, all in general position (and hence disjoint) from each other in M. Using these spines, choose (as before) sweep-outs for P_K and Q_K such that (a) each spine is also in general position with respect to the sweep-out coming from the other bridge surface and (b) the two sweep-outs themselves are in general position with respect to each other. This operation will be referred to as a *two-parameter sweep-out*. Two-parameter sweep-outs and their associated graphs are defined in detail in [To], so we provide only a brief overview here.

Associated to a two-parameter sweep-out is a square $I \times I$, where each point (s,t) in the square represents a configuration of P_K and Q_K during the sweep-out. Cerf theory [Ce] describes a graph Γ in the square, a graph whose relevant vertices are all of degree 4. The edges of Γ correspond to configurations where either

- P_K and Q_K intersect in a saddle singularity or
- P_K, Q_K, and K intersect transversally at a point.

The former edges are called *saddle edges* and the latter (which are best thought of as points where a transverse intersection curve of P_K and Q_K intersects the knot K) are called *K-edges.* Vertices in the graph correspond to configurations where two of these events occur simultaneously. That is, either (a) P_K and Q_K intersect in two saddle singularities, or (b) they intersect in a single saddle singularity and elsewhere a curve of $P_K \cap Q_K$ intersects K, or (c) $P_K \cap Q_K$ intersects K in two different points. (According to Cerf theory, some edges may also contain 2-valent "birth–death" vertices, but these are irrelevant to our argument.) Each component of $(I \times I) - \Gamma$, called a *region,* corresponds to a configuration in which P_K and Q_K are transverse.

Given a region of a two-parameter sweep-out of P_K and Q_K as described previously, let $\mathcal{C}_P$ (resp. $\mathcal{C}_Q$) be the set of all curves of $P_K \cap Q_K$ that are essential in

P_K (resp. Q_K) and let $\mathcal{C} = \mathcal{C}_P \cup \mathcal{C}_Q$. Associate to each region of $(I \times I) - \Gamma$ one or more labels in the following manner.

- If there exist curves in $\mathcal{C}$ that are essential on P_K and inessential on Q_K, then pick an innermost such curve c on Q_K and let $D^c \subset Q$ be the (punctured) disk that c bounds. The interior of D^c might intersect P_K, but only in curves that are inessential in both surfaces. If a neighborhood of ∂D^c lies in A (resp. B), label the region A (resp. B). Analogously label regions with X and/or Y.
- If $\mathcal{C} = \emptyset$ (i.e., if P_K and Q_K are either disjoint or intersect only in curves that are inessential on both surfaces), label the region a if some essential curve on Q_K lies entirely in B. (Note the switch: essential curves of Q_K in B result in the label a.) Use the analogous rule to label regions b, x, and y.

LEMMA 5.2. *If a region in the graph has label A then there is a curve $c \subset P_K \cap Q_K$ that c-compresses in A.*

Proof. Let D^A denote a (punctured) disk in M transverse to P_K such that:

- $\partial D^A = c^A \subset P_K \cap Q_K$ is essential in P_K;
- interior(D^A) intersects P_K only in inessential curves;
- a neighborhood of ∂D^A in D^A lies in A; and
- among all such disks with boundary c^A, $|D^A \cap P|$ is minimal.

Such a (punctured) disk exists by definition of label A; in fact, it lies in Q. Among all the components of interior$(D^A) \cap P$, choose a component c' that is innermost in P. Let D'^A and D^P denote the disjoint (punctured) disks that c' bounds in D^A and P, respectively. Then $S' = D'^A \cup_{c'} D^P$ is a sphere in M with at most two punctures. Although S' may not lie completely in any handlebody, it does intersect P only in inessential curves. It follows that S' is separating: Any closed curve in M is homologous to a closed curve in P, and such a closed curve will intersect S' an even number of times. Since S' is separating, $|K \cap S'|$ is either 0 or 2, so D'^A is punctured if and only if D^P is punctured.

Now, in D^A replace D'^A with D^P. The new (punctured) disk D' still has boundary c^A and all its curves of intersection with P are still inessential in P_K (though curves of $D' \cap Q$ could be essential in Q_K), but $|D' \cap P| < |D^A \cap P|$ because the curve c' and perhaps more have been eliminated. From this contradiction we deduce that P is disjoint from interior(D^A); that is, $D^A \subset A$ as required. □

LEMMA 5.3. *If a region in the graph has both labels A and B, then P_K is c-strongly compressible.*

Proof. By Lemma 5.2 there exist curves c^A and c^B in $P_K \cap Q_K$ that c-compress in A and B, respectively. Because they arise as curves of the intersection of two embedded surfaces, c^A and c^B are disjoint and so the (punctured) disks they bound comprise a c-strong compression of P. □

LEMMA 5.4. *If the union of the labels in adjacent regions in the graph contains both labels A and B, then P_K is c-strongly compressible.*

Proof. If both labels are in the same region, the result follows from Lemma 5.3. Otherwise, one region has label A but not B and the adjacent region has label B but not A.

Suppose first that the edge is a saddle edge. Going through the edge of the graph, separating the regions corresponds to banding together two curves c_+ and c_- to yield a curve c. The three curves co-bound a pair of pants on P_K and are thus disjoint. Therefore, the curve c_A that gives rise to the label A is disjoint from the curve c_B that gives rise to the label B. Now the argument in the proof of Lemma 5.2 shows that these curves bound c-disks for P_K on opposite sides, so P_K is c-strongly compressible.

Now suppose that the edge is a K-edge. In this case, going through the edge of the graph corresponds to isotoping a curve of intersection c across a point in K. Beforehand, c gives rise to a label A and afterward to a label B. But the curves of intersection (before and after the isotopy) are disjoint in P, so the same argument applies. □

LEMMA 5.5. *If the union of the labels in adjacent regions in the graph contains both labels A and b, then P_K is c-strongly compressible.*

Proof. No single region can have both labels, since one implies there is a curve of intersection that is essential in P_K and the other implies that all curves of intersection are inessential in both surfaces. So suppose one region R^A of the graph has label A and an adjacent region R^b has label b. In the first region, there is at least one curve in $\mathcal{C}$ that is essential on P_K and inessential on Q_K; in the second region, all curves of $P_K \cap Q_K$ are inessential on both surfaces. Here again a saddle edge is both representative and more difficult, so we examine just that surface.

Passing through the edge of the graph separating the regions corresponds to a saddle tangency where curves c_+ and c_- join together to form c. Let $R \subset P_K$ and $T \subset Q_K$ be the pairs of pants bounded by the three curves $c_\pm, c$ in the two surfaces.

Case 1: There are no other curves of intersection in the regions R^A and R^b. In this case, suppose first that the region in which both c_+ and c_- appear is R^b and that the region in which c appears is R^A. Then the (punctured) disk D in Q_K bounded by c lies entirely in A, since there are no other curves of intersection. After the saddle move that creates $c_\pm$, D remains as a c-disk for P_K that is disjoint from Q_K. The result then follows from Lemma 3.7.

Next suppose that the region in which both c_+ and c_- appear is R^A and that the region in which c appears is R^b. With no loss of generality, suppose c_+ is the curve that gives rise to the label A, so c_+ is essential on P_K but inessential on Q_K. If the (punctured) disk D that c_+ bounds in Q does not contain c_-, the proof is just as before. If it does contain c_-, then c_- is also inessential in Q_K and bounds a (punctured) disk in B. If c_- is essential in P_K then the result follows from Lemma 5.3. So suppose c_- is inessential in P_K. Since c_+ and c_- are nested in D, either c_- or the annulus between c_+ and c_- in $D \subset Q_K$ contains no puncture. In other words, either the (punctured) disk in $Q_K \cap B$ bounded by c_- or the (punctured) disk in $Q_K \cap A$ bounded by c contains no puncture. It follows that either the (punctured) disk in P_K bounded by c_- or the (punctured) disk in P_K bounded by c (both of which curves we are now assuming to be inessential in P_K) contains no puncture. But then c_+ also bounds a (punctured) disk in P_K, contradicting the fact that it is essential.

Case 2: There are curves of intersection other than $c_\pm, c$ in the regions R^A and R^b. The proof is by induction on the number of such curves. Because of the label b, all such curves are inessential in both surfaces; let α be an innermost one in Q. If the (punctured) disk D^Q that α bounds in Q contains T (but, by assumption, contains no other curves of intersection), then the proof is essentially as in Case 1. If D^Q does not contain T, apply the argument of Corollary 3.4 to α, reversing the roles of P and Q. The disk D^P that α bounds in P cannot contain R because R contains essential curves. Hence the isotopy of D^P across D^Q that removes α (and perhaps other curves of $D^P \cap Q$) has no effect on the three relevant curves $c_\pm, c$. Therefore, α can be removed without affecting the hypotheses. $\square$

6. Vertices in the Graph with Four Adjacent Labels

In order to consider labels around vertices in the graph, we now return to Conway spheres.

Definition 6.1. A collar $(S, K \cap S) \times I$ of a Conway sphere S is *well-placed* with respect to a bridge surface P for K if P intersects the collar as in a Conway product. That is, up to homeomorphism, $P \cap S \times \{0\}$ is a horizontal great circle c_0, $P \cap S \times \{1\}$ is a vertical great circle c_1, and $P \cap S \times I$ is a twice-punctured spanning annulus that is the trace of a rotation ρ_t from c_0 to c_1. (See Figure 7.)

Corollary 6.2. *If a bridge surface P is well-placed with respect to a collar $S \times I$ of a Conway sphere S and if neither component of $P \cap (S \times \partial I)$ bounds an unpunctured disk in P_K, then either P_K is c-strongly compressible or the surface $P_K - (S \times I)$ is c-incompressible in $M_K - (S \times I)$.*

Proof. First note that there are c-disks in both $A \cap (S \times I)$ and $B \cap (S \times I)$. Figure 12 shows a c-disk on one side; one on the other side is symmetrically placed.

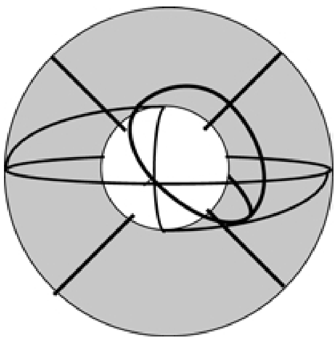

Figure 12

(The condition that neither component of $P \cap (S \times \partial I)$ bound an unpunctured disk in P_K guarantees that the boundary of each disk is essential in P_K; the boundaries of the two disks intersect in two points, so they do not themselves give a c-strong compression.) To prove the corollary, we will show that if the surface $P_K - (S \times I)$ is c-compressible in $M_K - (S \times I)$ then P_K is c-strongly compressible.

Consider a c-compressing disk D for $P_K - (S \times I)$ in $M_K - (S \times I)$. With no loss of generality, $D \subset A$. Since P_K intersects $S \times I$ in a twice-punctured annulus, ∂D is essential not only in $P_K - (S \times I)$ but also in P_K. Thus D is a c-disk for P_K lying in A. We have already exhibited (Figure 12) a c-disk for P_K lying in B, and we know that it is disjoint from D (because its boundary lies in $S \times I$) while ∂D is disjoint from $S \times I$. Since there are c-disks for P_K lying on opposite sides of P_K and since the boundaries are disjoint, P_K is strongly compressible. □

LEMMA 6.3. *Suppose $M = A \cup_P B$ is a bridge decomposition of $K \subset M$, and suppose there is a 4-punctured sphere $S_K \subset M_K$ intersecting P in a single circle c that is essential in S_K. Then*:

(i) *c bounds a disk in A_K or B_K*;
(ii) *S_K is incompressible (and so is an incompressible Conway sphere)*;
(iii) *a component of $S_K - c$ is parallel to a component of $P_K - c$; or*
(iv) *P_K is c-strongly compressible.*

Proof. We know that c divides S into the twice-punctured disks $D^A \subset A$ and $D^B \subset B$. If either of these disks is c-compressible in its handlebodies, then the c-disk must have boundary parallel to c in the disk and so, by parity, it must be an unpunctured disk. This gives part (i).

So henceforth we assume that both D^A and D^B are c-incompressible in their respective handlebodies. On the other hand, by Lemma 3.1, D^A_K (resp. D^B_K) is ∂-compressible in A (resp. B). Either any boundary compression of D^A_K into A produces a c-disk for P_K in $A_K - S$, or D^A_K is parallel in A_K to a twice-punctured subdisk of P_K, yielding part (iii). So, assume hereafter that any boundary compression of D^A_K into $A_K - S$ (or, symmetrically, D^B_K into $B_K - S$) produces a c-disk for P_K. (In particular, following the claim, $P_K - S$ is c-compressible into both $A_K - S$ and $B_K - S$.) If ∂-compressing disks for D^A_K and D^B_K abut $P_K - S$ in disjoint arcs (so that, in the terminology of [S], P_K is strongly ∂-compressible to S_K), then these boundary compressions create a pair of c-strong compressing disks for P, giving part (iv). Therefore, to prove the lemma it suffices to show that if S_K is compressible in M_K then P_K is strongly ∂-compressible to S_K. This we now do.

Consider the intersection of P with a compressing disk D for S_K. The argument is long and complex, but it is analogous to the proof of [S, Thm. 5.4] (see also the previous proof of Lemma 3.5), to which we mostly defer. Since D^A_K and D^B_K are incompressible in M_K, it follows that ∂D must intersect c and so $D \cap P$ contains arcs of intersection. If it contains closed curves of intersection, if the innermost ones reflect compressing disks for P_K in either $A_K - S$ or $B_K - S$, and if both types arise, then P_K is indeed strongly compressible. If, on the other hand, all innermost disks of $D - P$ cut off compressing disks for P_K into A, say, then we apply the argument of [S, Lemma 5.5]. To be specific: let P_A denote P_K maximally compressed into A, and use P_A instead of Q_X in that argument. We may conclude that either P_K is c-strongly compressible or D, when chosen to minimize $|D \cap P_K|$, yields $D \cap P_K$ consisting entirely of arcs.

Now consider these arcs of intersection. Suppose without loss of generality that some arc of $D \cap P_K$ outermost in D cuts off from D an outermost disk that lies in A. If there is a similar outermost disk cut off that lies in B, then together these describe a strong ∂-compression of P_K to S_K and we are done. So henceforth assume that all arcs of $D \cap P_K$ that are outermost in D cut off a disk in A. Let Δ^B be a c-disk for $P_K - S$ in $B_K - S$ and consider how Δ^B intersects D. A standard innermost disk argument dispenses with any circles of intersection. If the arcs $\Delta^B \cap D$ are all disjoint from any outermost arc of $D \cap P_K$, then the boundary compression of D_K^A given by the latter exhibits a c-disk for $P_K - S$ in $A_K - S$ that is disjoint from Δ^B. Thus P_K would be c-strongly compressible (part (iv) again). So we may as well assume that each outermost arc of $D \cap P_K$ is incident to an arc of $\Delta^B \cap D$.

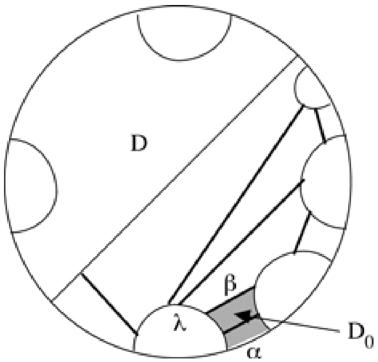

Figure 13

Now the argument is analogous to Case 3 in the proof of [S, Thm. 5.4], mostly picking up just below the figure [S, p. 347] (shown as Figure 13). Here Δ^B plays the role of the disk E in the cited proof. The construction given there exhibits a pair of disjoint ∂-compressing disks for P_K, one in A_K (cut off by λ in Figure 13) and one in B_K (containing the rectangle cut off by β in Figure 13). This shows that P_K is strongly ∂-compressible to S_K, as required. □

LEMMA 6.4. *Suppose there is a* 4*-valent vertex in the graph, the adjacent four regions each have a single label, and* (*in order around the vertex*) *these labels are* A, X, B, Y. *Then*:

(i) P_K *or* Q_K *is c-strongly compressible*;

(ii) P_K *and* Q_K *can be isotoped to intersect in a nonempty collection of curves that are essential in both surfaces*; *or*

(iii) *there is an incompressible Conway sphere $S \subset M$ with a collar $S \times I$ that is well-placed with respect to both P and Q. The pairs of curves $P \cap (S \times \partial I)$ and $Q \cap (S \times \partial I)$ are parallel in one component of $S_K \times \partial I$ and anti-parallel in the other. Both $P_K - (S \times I)$ and $Q_K - (S \times I)$ are c-incompressible in $M_K - (S \times I)$. The surfaces P and Q can be properly isotoped outside the collar so that all curves of intersection outside the collar are essential in both P_K and Q_K.*

Proof. We will suppose both P_K and Q_K to be c-weakly incompressible and show that only (ii) or (iii) is possible. Since all four labels appear, it follows from Lemma 5.1 that K is not the unknot in S^3. It is fairly easy to rule out the possibility that one or more edges is a K-edge, so again we focus on where each edge adjacent to the 4-valent vertex in the graph corresponds to a saddle move—that is, we attach a band to a curve or a pair of curves. Let c^A, c^X, c^B, c^Y denote the collection of all curves of $P_K \cap Q_K$ in each of the corresponding regions. The proof of Lemma 5.3 shows that no curve in c^A responsible for the label A can be isotoped to be disjoint from a curve in c^B responsible for the label B. It follows that the two disjoint bands by which c^A is converted to c^B must be incident to the opposite side in P_K of some curve c_0 in c^A, a curve that is essential in P_K and inessential in Q_K. Let b^X (resp. b^Y) be the band attached to c^A that changes c^A to c^X (resp. c^Y).

Consider how many curves in c^A are involved in the band moves at the vertex. There are two bands and each has two ends. Because c_0 is incident to two of those ends, at most two other curves are involved.

Case 1: c_0 is the only curve in c^A that is involved in the band moves, and only one curve in c^B is involved in the band moves. This means that both ends of b^X and both ends of b^Y are attached to c_0. If the ends of the two bands are not interleaved in c_0 (i.e., if the ends of b^X lie in a subinterval of c_0 that is disjoint from the ends of b^Y) then the result of attaching both bands would be three curves in c^B, contrary to the assumption in this case. Hence the ends of b^X and b^Y are interleaved in c_0. Viewed in Q_K, where c_0 bounds a (punctured) disk, the band moves appear as follows. Two bands are attached to the (punctured) disk of $A \cap Q$ bounded by c_0, with the ends of the bands interleaved. Moreover, the result of attaching both of them is an inessential curve bounding a (once-) punctured disk in B. This is possible only if Q is a twice-punctured torus and the bands correspond to orthogonal curves in the torus (call them the *meridian* and *longitude*). The labels X and Y show that the meridian and longitude each bound (punctured) disks on one side or the other, so we further conclude that Q is an unknotted torus in S^3. Since each of the four quadrants has only one label, every curve of intersection that is not incident to either band is inessential in both P_K and Q_K. In particular, in the region labeled A, all curves of intersection other than c_0 can be removed by a proper isotopy of P (following Corollary 3.4).

Now do the band move that creates the pair of curves c^X: two parallel essential curves in Q with a single puncture in each annulus between them in the torus Q. At that point, consider the surfaces $P^X = P \cap X$ and $P^Y = P \cap Y$. Exploiting the bridge positioning of K in X and Y, let D^X and D^Y be meridian disks for the solid tori X and Y (respectively) chosen to be disjoint from K and to have minimal intersection with P^X and P^Y (respectively).

Since the region is labeled X, one component P^D of P^X is a (punctured) disk; let P^0 be the other component. We will show that P^0_K, and hence P^X_K, is compressible in X_K. First note that P^X cannot be simply the union of an unpunctured disk and a possibly punctured disk, since c_0 bounds the band sum of the two and is essential in P_K. Suppose first that P^D is unpunctured and so P^0 is not a (punctured) disk. Then P^0_K has non-abelian fundamental group. On the other hand, K intersects the 3-ball $X - P^D$ in an unknotted arc, so $\pi_1(X_K - P^D) = \mathbb{Z}$. It follows that P^0_K is compressible in X_K, as claimed. Suppose next that P^D is punctured, so P^0 is not an unpunctured disk. A standard cut-and-paste argument shows that D^X is disjoint from P^D. It follows that ∂D^X is parallel to ∂P^0 in Q_K and so reveals a compressing disk for P^0_K in X_K. Again we conclude that P^0_K is compressible in X_K.

Similarly, P^Y_K is at least c-compressible in Y_K. (Recall that we are still considering the region with label X, so there is not necessarily any symmetry with P^X_K.) For consider the intersection of P^Y with D^Y. If there are any closed curves in $P^Y \cap D^Y$, then an innermost one in D^Y would give a compressing disk for P^Y_K in Y_K, as we are seeking. If there are only arcs of intersection, then ∂-compression of P^Y_K via an outermost disk cut off in D^Y would reveal either a c-disk for P^Y_K in Y_K or that P^Y is a ∂-parallel punctured annulus. The latter is impossible: the dual band to b^X would span the annulus P^Y and its complement in P^Y would be a punctured disk in P whose boundary is c_0, contradicting that c_0 is essential in P_K.

Having established that P^X_K is compressible in X_K and that P^Y_K is at least c-compressible in Y_K, we now consider where these c-disks for P^X_K and P^Y_K must lie. If one lies in $X_K \cap A$ and the other on $Y_K \cap B$, then they would give a c-strong compression of P_K. So they both lie on the same side of P—say, in $X_K \cap A$ and $Y_K \cap A$. A single band move changes c^X to c^B. If that band is in X_K then the c-disk in $Y_K \cap A$ is undisturbed. Similarly, if that band is in Y_K then the c-disk in $X_K \cap A$ is undisturbed. In any case, there is a c-disk for P_K in A that is disjoint from c^B, contradicting the c-weak incompressibility of P.

Case 2: Exactly two curves are involved in the band move at c^A and each band b^X and b^Y has one end on each curve. Then the ends of the bands b^X and b^Y that are not in c_0 both lie in a curve c_1. But this implies that the only curve in c^X involved in the band move is the curve obtained by banding c_0 and c_1 together along b^X. Similarly, only one curve in c^Y is involved in the band moves. So this is exactly Case 1 with the roles of P and Q reversed.

Claim 1: All other cases are equivalent to the case where exactly three curves of c^A are involved in the band move. If instead c_0 is the only curve of c^A involved in the band move and if (unlike in Case 1) the ends of the bands are not interleaved, then the result is that exactly three curves of c^B are involved in the band move. Switching A and B in the argument establishes the claim in this case. If exactly one other curve c_1 is involved in the band move and if (unlike in Case 2) the end of one band (say b^Y) has both ends on c_0, then exactly three curves of c^Y are involved in the band moves. Thus, switching the roles of P and Q and of A and Y establishes the claim also in this case.

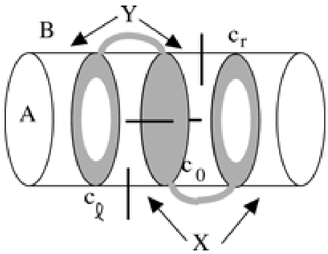

Figure 14

Having established the claim, we now proceed under the assumption that exactly three curves of c^A are involved in the band move. Then b^X attaches c_0 to another curve c_ℓ. In order for this band move to create the new label X, the band b^X must span a (punctured) annulus in P between c_0 and c_ℓ. The same is true for b^Y, so all three curves c_ℓ, c_0, c_r in c^A that are involved in the band moves are co-annular in P; c_0 is in the middle, and the entire annulus contains at most two punctures (see Figure 14).

Claim 2: Both c_ℓ and c_r are essential in both surfaces. If (say) c_r were essential in exactly one surface, it would give a label (by hypothesis, the label A) and that label would persist when b^X is attached, since attaching the band has no effect on c_r. This would contradict the hypothesis that there is only one label in each region. Suppose, on the other hand, that c_r were inessential in both surfaces. If it bounded an unpunctured disk in either surface (and hence in both surfaces), then the band move on b^Y could have no effect; the resulting curve in $P \cap Q$ would be isotopic to c_0. Hence we suppose that c_r bounds a punctured disk in both surfaces. Then, by parity, the (punctured) annulus between c_0 and c_r and the (punctured) disk in Q bounded by c_0 must together have exactly one puncture. If the annulus had no puncture then c_0 would, against hypothesis, be inessential in P. If the disk in Q bounded by c_0 contained no puncture, then when c_0 is banded to c_r the result would be, against hypothesis, inessential in Q. These contradictions establish the claim.

Claim 3: All curves of intersection that are inessential in both P_K and Q_K can be removed by isotopies of P and Q that do not affect c_0, c_ℓ, or c_r or either band b^X, b^Y. The proof is by induction on the number of curves that are inessential in both surfaces. Let α be an innermost one in Q. The (punctured) disk D that α bounds in Q_K cannot contain c_ℓ or c_r, since these are essential in Q. Nor can it contain c_0, since c_0 is banded via the bands b^X and b^Y to the essential curves c_ℓ and c_r in the complement of α. Apply the argument of Corollary 3.4 to α, reversing the roles of P and Q. The disk D^P that α bounds in P cannot contain c_0, c_ℓ, or c_r nor the bands b^X, b^Y between them—because the curves are essential in P_K. Thus the isotopy of D^P across D^Q that removes α (and perhaps other curves of $D^P \cap Q$) has no effect on the relevant curves and bands. This isotopy removes α and completes the inductive step.

Claim 4: Neither c_ℓ nor c_r can bound a (punctured) disk in any of X, Y, A, or B. Suppose that, say, c_ℓ bounded a (punctured) disk in X or Y. Note that c_ℓ can be isotoped in Q_K to be disjoint from the curve obtained by banding c_0 to either c_ℓ or c_r and so, by Claim 3, it is disjoint from curves bounding punctured disks in Y and X. It would follow that Q_K is strongly compressible, a contradiction.

To show that c_ℓ can't bound a (punctured) disk in A or B, note that c_ℓ can be isotoped in P_K to be disjoint from c_0 and from the curve obtained by banding c_0 simultaneously to both c_ℓ and c_r. Then the same argument applies.

Following the claims, we consider the appearance of the three curves in Q. First, c_0 bounds a (punctured) disk that lies in $Q \cap A$. When c_0 is banded to essential curve c_r the result is still essential in Q (though now inessential in P), but when this curve is further banded to c_ℓ the resulting single curve bounds a (punctured) disk that lies in $Q \cap B$. Hence c_r and c_ℓ are also co-annular in Q, with the annulus between them containing c_0 and at most two punctures of which no more than one is outside the (punctured) disk bounded by c_0.

In fact, all possible punctures in the preceding descriptions must appear. If, for example, either the disk bounded by c_0 in Q or the annulus bounded by c_r and c_0 in P contains no puncture, then c_r also bounds a (punctured) disk in A_K—namely, the disk obtained by combining the annulus in P and the disk in Q bounded by c_0. But such a disk would contradict Claim 4.

The next step is either to rearrange the intersection so that P_K and Q_K meet only in curves that are essential in both surfaces (as in part (ii)) or to exhibit the incompressible Conway sphere S of part (iii). We consider the most unusual case first.

Claim 5: If c_ℓ bounds twice-punctured disks D_K^P and D_K^Q in P_K and Q_K (respectively) and if these twice-punctured disks are parallel in $A \cap Y$, then Lemma 6.4(ii) holds. To demonstrate the claim, use the parallelism to perform a K-isotopy of c_ℓ across one of the punctures in D_K^P and D_K^Q, adding an extra puncture in the annulus between c_ℓ and c_r in both P and Q. Then do the band move along b^X (see Figure 15). The result is to make a curve of intersection that is essential in P_K instead of one that gives rise to the label X, since the disks that the curve bounds in P_K and Q_K are now both twice-punctured disks. By hypothesis, all other curves of intersection are also essential in both surfaces, yielding part (ii).

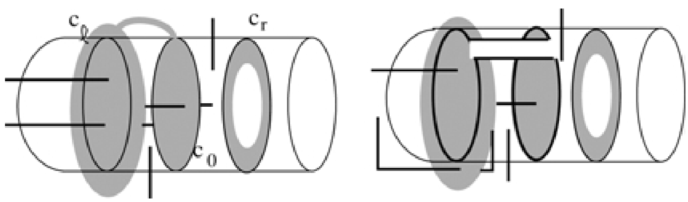

Figure 15

A symmetric argument applies if c_r bounds a twice-punctured disk in both P_K and Q_K and if these two disks are parallel.

We now construct the relevant Conway sphere. Start from our previous description of how P and Q intersect in the region labeled A. The band b^X must lie in

X because a (punctured) disk in X is obtained by removing the band from the annulus in P between c_0 and c_ℓ. Similarly, b^Y lies in Y. Push the bands b^X and b^Y close to Q so that both lie in a collar $C(Q)$ of Q. The end of the collar Q^X in X intersects P as if the band move had been done on b^X; the end of the collar Q^Y in Y intersects P as if the band move had been done on b^Y.

Let $V \subset Q$ be the twice-punctured annulus bounded by $c_\ell \cup c_r$; here V contains c_0 and the once-punctured disk in $Q \cap A$ that it bounds. Choose a curve c in Q just outside of V that is parallel in Q_K to c_ℓ. Use the collar structure to push c to the copy Q^X of Q. Since in Q^X the curve c_0 has been banded to c_ℓ, the copy of c in Q^X can be isotoped, intersecting K once but disjoint from P, to a curve that bounds a once-punctured disk in X disjoint from P—namely, a parallel copy of the disk that gives rise to the label X. Thus c bounds a twice-punctured disk D^X in X that is disjoint from P. On the other hand, a parallel copy c' of c_ℓ just *inside* V, when pushed to the copy Q^Y of Q, can be isotoped in Q^Y (crossing K once but disjoint from P) to the curve obtained when b^Y connects c_0 to c_r. It follows that c' bounds a twice-punctured disk D^Y in Y disjoint from P. If the disks D^X and D^Y are attached to the annulus in Q_K between c and c', the result is a 4-punctured sphere S_ℓ that intersects both P and Q in a single circle. The circle is essential in P_K and Q_K because c_ℓ is. The same construction can be done at c_r instead of c_l. The resulting sphere S_r is parallel to S_ℓ in M_K, and the construction explicitly shows that the region between them is well-placed with respect to both P and Q (see Figure 16).

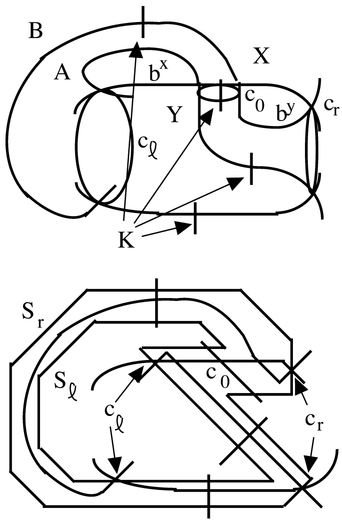

Figure 16

To determine whether S_ℓ (or S_r) is an incompressible Conway sphere for P and Q, we apply Lemma 6.3. There are three cases, as follows.

Case a: No component of $S_{\ell K} - c_\ell$ is parallel to a component of $P_K - c_\ell$ or $Q_K - c_\ell$. Then, by Lemma 6.3, S_ℓ is a incompressible Conway sphere for both surfaces. Moreover, Corollary 6.2 shows that the surfaces $P_K - (S \times I)$ and $Q_K - (S \times I)$ are each c-incompressible in $M_K - (S \times I)$. This yields part (iii) of the lemma.

Case b: c_ℓ cuts off a component Q^ℓ of $Q_K \cap A$ that is parallel to $S_\ell \cap B$ (see Figure 17). In this case, the component P^ℓ of $P_K \cap Y$ cut off by c_ℓ lies between the two parallel twice-punctured disks. This component is also incompressible in that region, essentially by Corollary 6.2. It follows easily that P^ℓ is also a parallel twice-punctured disk and so, following Claim 5, part (ii) holds.

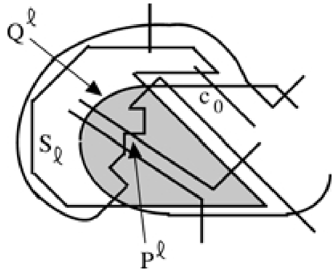

Figure 17

The case in which c_ℓ cuts off a component of $P_K \cap Y$ that is parallel to $S_\ell \cap X$ is handled symmetrically. Here c_ℓ cannot bound a twice-punctured disk on the other side in either surface (i.e., the B-side of c_ℓ in Q_K or the X side of c_ℓ in P_K) because c_r is essential in both surfaces.

Case c: c_ℓ cuts off a component Q^ℓ of $Q_K \cap A$ that is parallel to $S_\ell \cap A$ (see Figure 18). The argument in this case is analogous to that in Claim 5: the parallelism between the two disks allows us to define a K-isotopy that moves the puncture

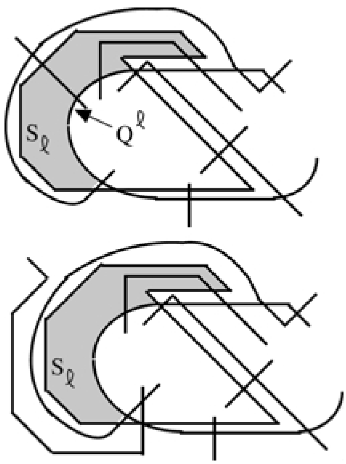

Figure 18

in the annulus between c_0 and c_ℓ in P to the annulus between c_ℓ and c_r in Q. Once both band moves on b^X and b^Y are completed (i.e., in the region labeled B at the vertex), what was previously a punctured disk component of $B \cap Q$ becomes a twice-punctured disk and so its boundary is essential in Q. This change, when viewed in P_K, simply replaces the old curve of intersection with a curve parallel to c_r that is essential in P_K. Hence part (ii) holds again.

The case when c_ℓ cuts off a component of $P_K \cap Y$ that is parallel to $S_\ell \cap Y$ is handled symmetrically. □

COROLLARY 6.5. *Suppose both P_K and Q_K are c-weakly incompressible, there is a* 4*-valent vertex in the graph, the adjacent four regions each have a single label, and* (*in order around the vertex*) *these labels are A, X, B, Y. Then*:

(i) *P_K and Q_K can be isotoped to intersect in a nonempty collection of curves, each essential in both surfaces*; *or*

(ii) *P_K and Q_K have the same Conway decomposing sphere S. The curves in which the surfaces intersect a collar of S are parallel at one end of the collar and are anti-parallel at the other. Furthermore, the surfaces can be isotoped outside the collar so that all curves of intersection outside the collar are essential in both surfaces.*

Now we can prove the main result of this paper.

THEOREM 6.6. *Suppose P and Q are bridge surfaces for a link K in a closed orientable* 3*-manifold M. Assume further that P_K and Q_K are both c-weakly incompressible and that neither is a twice-punctured sphere. Then P_K may be properly isotoped so that P_K and Q_K intersect in a nonempty collection of curves such that*:

(i) *all curves of intersection are essential in both surfaces*; *or*

(ii) *P_K and Q_K have the same Conway decomposing sphere S. The curves in which the surfaces intersect the collar of S are parallel at one end of the collar and are anti-parallel at the other. Outside the collar, all curves of intersection are essential in both surfaces.*

An important special case is the following corollary.

COROLLARY 6.7. *Suppose P and Q are bridge surfaces for a link K in a closed orientable* 3*-manifold M. Assume further that $K \subset M$ is not the unknot in S^3, that there is no incompressible Conway sphere for K, and that both P_K and Q_K are c-weakly incompressible. Then P_K may be properly isotoped so that P_K and Q_K intersect in a nonempty collection of curves that are essential in both surfaces.*

Proof. This is immediate from Theorem 6.6. □

Proof of Theorem 6.6. Because neither P_K nor Q_K is a twice-punctured sphere, the neighborhood of any bridge disk for any bridge surface P contains a compressing disk for P_K. Since P_K is c-weakly incompressible, it follows from Lemma 5.1 that $K \subset M$ is not the unknot in S^3.

Consider a two-parameter sweep-out of P_K and Q_K together with the labeling scheme for its graph described previously. The theorem reduces to proving that either there is an unlabeled region or there is a 4-valent vertex as described in Corollary 6.5. We will prove the existence of such a region or such a vertex via a sequence of claims. The labeling is symmetric with respect to the surface that we consider, so any statement regarding labels A and B (say) has an equivalent statement for labels X and Y.

Claim 1: The union of the labels of two adjacent regions cannot contain both labels A and B. This follows immediately from Lemma 5.4.

Claim 2: The union of the labels of two adjacent regions cannot contain both labels a and b. As in the proof of Lemma 5.4, going through the edge of the graph separating the regions corresponds to banding together two curves c_+ and c_- to yield a curve c. Since both regions have small labels, all curves of $P_K \cap Q_K$ before and after passing through the saddle are inessential on both surfaces. Thus the three curves $c_\pm, c$ bound a pair of pants $F \subset Q$, and each curve bounds a (punctured) disk in Q. The (punctured) disks bounded by the three curves cannot all be disjoint from F, else Q would be a twice-punctured sphere, so one of the (punctured) disks $D \subset Q$ bounded by the three curves must contain all of F.

No essential curve of Q_K can lie inside of D, so at least some parts of both the essential curves $c_a \subset Q_K \cap B$ and $c_b \subset Q_K \cap A$ that give rise to the labels a and b lie outside D. Note that c_a and c_b are also automatically disjoint from all the other punctured disks in Q that are bounded by components of $P \cap Q$, since c_a and c_b are disjoint from P. But removing these disks from Q_K leaves a connected surface, so parts of c_a and c_b lie in the same component of $Q - P$. This is impossible, since P separates A from B in M.

Claim 3: The union of the labels of two adjacent regions cannot contain both labels A and b. This follows immediately from Lemma 5.5.

Claim 4: The theorem is true if there is an unlabeled region. In the corresponding region, all curves of intersection are either essential in both surfaces or inessential in both surfaces, and some must be essential in both surfaces to avoid a label x or y. Apply Corollary 3.4 to remove all inessential circles of intersection but leave all essential curves of intersection. This gives Theorem 6.6(i).

Relabel the regions of the graph as follows. Assign a region label $\textsc{a}$ if A or a is among the labels of the region. Similarly, replace B and b by $\textsc{b}$, x and X by $\textsc{x}$, and y and Y by $\textsc{y}$.

By Claims 1–3, labels $\textsc{a}$ and $\textsc{b}$ *never* appear as labels of the same region or labels of adjacent regions. The same holds for labels $\textsc{x}$ and $\textsc{y}$.

Claim 5: The theorem is true if there is a vertex surrounded by regions that have all four labels $\textsc{a}$, $\textsc{b}$, $\textsc{x}$, *and* $\textsc{y}$. Suppose such a vertex exists, and label the four regions (clockwise) as R_n, R_e, R_s, and R_w. Without loss of generality, suppose that R_n is labeled $\textsc{a}$. Then, according to Claims 1–3, R_s must be the region that carries the label $\textsc{b}$ and, furthermore, neither R_e nor R_w can contain either the label $\textsc{a}$

or the label B. If either R_e or R_w is unlabeled then we are done by Case 4, so without loss of generality assume that R_e carries the label X. Then the symmetric argument shows that R_w must carry the label Y.

All the labels must come from uppercase letters: Suppose, for example, that R_n were labeled a. This would imply that all curves of intersection are inessential in both surfaces and so some essential curve of P lies in either X or Y. Hence R_n would also have the label y or x. Then R_n and either R_w or R_e would contradict Claim 2.

Any curve of intersection other than the ones involved in the saddle moves around the vertex must be either essential in both surfaces or inessential in both; otherwise, there would be a specific additional label in all four regions around the vertex, implying one region would have both labels X and Y or both labels A and B. Either case contradicts Claims 1–3. Since the labels are in fact uppercase A, B, X, Y, the claim follows from Corollary 6.5.

Consider the labeling of the regions adjacent to $\partial(I \times I)$. In each of these regions, one of the surfaces P_K or Q_K is the boundary of a small neighborhood of one of the spines. Suppose that the north and south boundaries of the square correspond to the spines $\Sigma_{(A,K)}$ and $\Sigma_{(B,K)}$, respectively, and that the east and west boundaries correspond to $\Sigma_{(X,K)}$ and $\Sigma_{(Y,K)}$, respectively. By general position we may arrange that all four spines are disjoint in M.

Claim 6: No region adjacent to the north boundary of the square has label B. In such a region, A_K is a thin neighborhood of the spine $\Sigma_{(A,K)}$ and Q_K sweeps from a neighborhood of $\Sigma_{(X,K)}$ to a neighborhood of $\Sigma_{(Y,K)}$. When Q_K is near $\Sigma_{(X,K)}$ or $\Sigma_{(Y,K)}$ it lies entirely in the complement of A_K, so these regions are both labeled a. It follows from Claim 1 that these regions cannot also have label B. To understand the labeling in other regions near the spine, apply general position to the sweep of Q_K across $\Sigma_{(A,K)}$.

During the sweep by Q_K, the surface Q goes through no tangencies with K but can go through vertices of $\Sigma_{(A,K)}$ and have tangencies with edges in $\Sigma_{(A,K)}$. With no loss of generality, assume that all vertices of $\Sigma_{(A,K)}$ are of valence 1 (where the spine attaches to the extrema of K) or valence 3.

When Q passes through either a tangency point with an edge of $\Sigma_{(A,K)}$ or through a valence 3-vertex of $\Sigma_{(A,K)}$, the effect is to change the number of curves of $P_K \cap Q_K$ that are essential in P_K and bound a disk in $A \cap Q_K$. (This may require two stages because a center tangency may be involved, but the center tangency merely creates a curve inessential in both surfaces and so has no effect on the labeling.) Thus, on one region or the other on either side of the corresponding edge(s) in the graph (and perhaps on both), there is a label A. It follows from Claims 1 and 2 that on neither region can there be a label B.

Suppose that Q passes through a valence-1 vertex of $\Sigma_{(A,K)}$. The sweep-out of Q is always transverse to K, so the sweep-out locally appears as in Figure 19. This is a three-stage process (in which two K-edges in the graph are crossed) that ultimately adds (or removes) the label A but otherwise involves only curves that are inessential in both P_K and Q_K. So if there were also a label B in any region

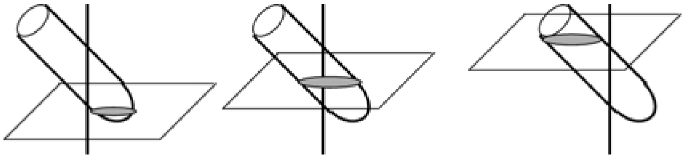

Figure 19

involved then there would be one on the region adjacent to label A, again contradicting Claim 1 or Claim 2.

The proof is now a consequence of the following quadrilateral variant of Sperner's lemma (see [ST, Apx.]).

LEMMA 6.8. *Given a graph Γ that is properly embedded in a square $I \times I$, suppose that all vertices of Γ are of valence 4 or 1 and that all valence-1 vertices are contained in $\partial(I \times I)$. Suppose further that each region of $I \times I - \Gamma$ is labeled with letters from the set* {A, B, X, Y}, *allowing unlabeled regions and regions with multiple labels, so that*:

- *the union of the labels of two adjacent regions never contains both* A *and* B *labels or both* X *and* Y *labels*; *and*
- *no region adjacent to the north boundary* (*resp. south, east, and west boundaries*) *of $I \times I$ is labeled* B (*resp.* A, Y, *and* X).

Then, either some region of $(I \times I) - \Gamma$ is unlabeled or all four labels occur in the four regions surrounding some vertex of Γ.

References

[A] C. Adams, *The knot book. An elementary introduction to the mathematical theory of knots,* Amer. Math. Soc., Providence, RI, 2004.

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

[CG] A. Casson and C. McA. Gordon, *Reducing Heegaard splittings,* Topology Appl. 27 (1987), 275–283.

[Ce] J. Cerf, *Sur les difféomorphismes de la sphère de dimension trois* ($\Gamma_4 = 0$), Lecture Notes in Math., 53, Springer-Verlag, Berlin, 1968.

[Co] J. H. Conway, *An enumeration of knots and links, and some of their algebraic properties,* 1970 Computational Problems in Abstract Algebra (Conf. Proc., Oxford, 1967), pp. 329–358, Pergamon, Oxford, 1970.

[GL] C. Gordon and J. Luecke, *Knots with unknotting number 1 and essential Conway spheres,* Algebr. Geom. Topol. 6 (2006), 2051–2116.

[RS] H. Rubinstein and M. Scharlemann, *Comparing Heegaard splittings of non-Haken 3-manifolds,* Topology 35 (1997), 1005–1026.

[S] M. Scharlemann, *Proximity in the curve complex: Boundary reduction and bicompressible surfaces,* Pacific J. Math. 228 (2006), 325–348.

[ST] M. Scharlemann and A. Thompson, *Thin position and Heegaard splittings of the 3-sphere,* J. Differential Geom. 39 (1994), 343–357.

[Schu] H. Schubert, *Über eine numerische Knoteninvariante,* Math. Z. 61 (1954), 245–288.

[Sch] J. Schultens, *Additivity of bridge numbers of knots,* Math. Proc. Cambridge Philos. Soc. 135 (2003), 539–544.

[To] M. Tomova, *Multiple bridge surfaces restrict knot distance,* Algebr. Geom. Topol. 7 (2007), 957–1006.

M. Scharlemann
Mathematics Department
University of California, Santa Barbara
Santa Barbara, CA 93106

mgscharl@math.ucsb.edu

M. Tomova
Mathematics Department
Rice University
Houston, TX 77005

mt2@rice.edu