

BAND SUMS OF LINKS WHICH YIELD COMPOSITE LINKS. THE CABLING CONJECTURE FOR STRONGLY INVERTIBLE KNOTS

MARIO EUDAVE-MUÑOZ

ABSTRACT. We consider composite links obtained by bandings of another link. It is shown that if a banding of a split link yields a composite knot then there is a decomposing sphere crossing the band in one arc, unless there is such a sphere disjoint from the band. We also prove that if a banding of the trivial knot yields a composite knot or link then there is a decomposing sphere crossing the band in one arc. The last theorem implies, via double branched covers, that the only way we can get a reducible manifold by surgery on a strongly invertible knot is when the knot is cabled and the surgery is via the slope of the cabling annulus.

INTRODUCTION

We consider here the problem of when a banding of a link k yields a composite link. There is a trivial situation in which a composite link is produced. This is when one of the summands of k is not perturbed by the band, i.e., there is a decomposing sphere disjoint from the band. There are other cases which are not trivial, for example the square knot is a band sum of trivial knots (Figure 1).

Definitions. A link k is *split* (or *splittable*) if there is a sphere Q disjoint from k which separates its components. Such a sphere is called a *splitting sphere*. Note that k is split if and only if $S^3 - k$ is reducible.

A link k is *composite* if there is a sphere P intersecting k transversally in two points, such that neither of the 3-balls bounded by P intersects k in a single unknotted spanning arc. Such a sphere is called a *decomposing sphere* for k . A link is *prime* if it is neither composite, nor split, nor trivial.

Let P be a decomposing sphere for a composite link k . Take a tubular neighborhood $\eta(k')$ of the component k' of k which meets P . $\partial\eta(k') - P$ consists of two annuli. Attach one of these annuli to $P - \text{int } \eta(k')$, to get a torus T . Call T a *swallow-follow torus* for k . There is a disk D such that $D \cap T = \partial D$ and k intersects D transversally in one point. By doing 2-surgery on T with D we recover the decomposing sphere. Note that in this construction we get two swallow-follow torus.

Received by the editors May 15, 1989.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 57M25; Secondary 57N10.

Key words and phrases. Band sum, banding, composite knot and link, decomposing sphere, split link, sutured manifold, surgery on knots.

Research supported in part by a National Science Foundation Grant.

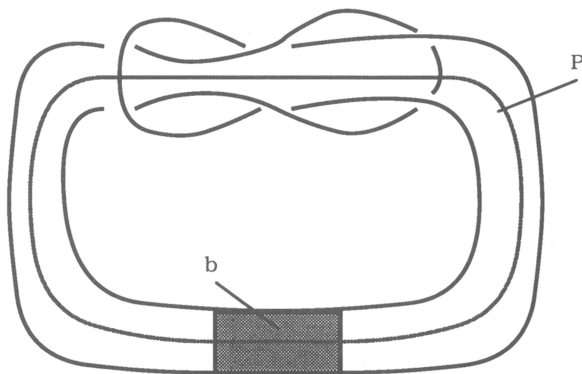


FIGURE 1

Let k be a link in S^3 . Let $b : I \times I \rightarrow S^3$ be an imbedding such that $b(I \times I) \cap k = b(I \times \partial I)$. Let $k_b = \{k - b(I \times \partial I)\} \cup b(\partial I \times I)$. k_b is another link called a *banding* of k . If $k = k_0 \cup k_1$ is a split link, where k_0, k_1 are knots, and $b(I \times \{i\}) \cap k = b(I \times I) \cap k_j$, for $i = 0, 1$, then k_b is called a *band sum* of k_0 and k_1 .

A band is *trivial* if there is a sphere in $S^3 - k$ which intersects the band in one arc parallel to $b(I \times \{1/2\})$. So if k_b is a band sum of k_0 and k_1 , b is trivial if and only if it is a connected sum of k_0 and k_1 .

A *Seifert surface* for a link k is a compact, oriented surface none of whose components are closed and whose boundary is the link. Define $\chi(k)$ to be the maximal Euler characteristic of all Seifert surfaces for k .

Let k be a link, and b a band such that k_b is composite. If P is a decomposing sphere for k_b , then it is possible to isotope P so that it meets k_b in two points not in $b(\partial I \times I)$, and intersects the band in a collection of arcs parallel to $b(\{1/2\} \times I)$. For example, in Figure 1 there is a decomposing sphere crossing the band in one arc. Our main results are the following.

Theorem 1. *Let k be a link and b a band such that $S^3 - (k \cup b(I \times I))$ is irreducible. If k_b is composite then either*

- (a) *there is a decomposing sphere for k_b disjoint from the band; or*
- (b) *there is a decomposing sphere for k_b which crosses the band exactly in one arc parallel to $b(\{1/2\} \times I)$; or*
- (c) *there is a swallow-follow torus for k_b disjoint from the band; or*
- (d) *$S^3 - k$ is irreducible and there is a Seifert surface S for k with $\chi(S) = \chi(k)$, and such that the band intersects S transversally always in the same direction.*

We prove Theorem 1 in §§1–5, using sutured manifold theory and a lengthy combinatorial argument.

Theorem 2. *Let $k = k_0 \cup k_1$ be a split link, for knots k_0, k_1 , and let k_b be a band sum of k_0 and k_1 . If k_b is composite then either there is a decomposing sphere disjoint from the band, or there is a decomposing sphere crossing the band in one arc parallel to $b(\{1/2\} \times I)$.*

Apply Theorem 1 to the more special case of Theorem 2. Note that under the hypothesis of Theorem 2, case (d) of Theorem 1 cannot happen. We use

the combinatorics of [BE] to prove that if (c) happens then this is because we are in the trivial case (a). This is shown in §6.

Theorem 3. *Let k be the trivial knot, and k_b a banding of k . If k_b is composite then there is a decomposing sphere which crosses the band in one arc parallel to $b(\{1/2\} \times I)$.*

Apply Theorem 1 to the case when k is the trivial knot. Note that cases (a) and (c) of Theorem 1 cannot happen. Making a combinatorial argument we prove that if (d) happens then (b) also happens. This is shown in §7.

Corollary 1. *Suppose $k = k_0 \cup k_1$ is a split link, and k_b is a band sum of k_0 and k_1 . Suppose k_b is a composite knot and that there is no decomposing sphere disjoint from the band. Then k_b has two prime summands, and one of them is a two-bridge knot.*

It is proved in [BE] that if a band sum of trivial knots yields a composite knot, then one of the summands is a two-bridge knot.

Corollary 2. *Suppose k is the trivial knot and k_b is a banding of k . If k_b is a composite knot or link, then it has two prime summands, one of them is a two-bridge knot or link.*

S. A. Bleiler [B₂] has proved that if k is the trivial knot and k_b is composite, then k_b has a two-bridge summand.

Theorems 2 and 3 tell us that most of the knots and links obtained by bandings of the trivial knot or a split link are in fact prime, unless there is a summand which is not perturbed by the band. For example, suppose a banding of the trivial knot or a band sum of a split link yields a composite knot or link, and there is no decomposing sphere disjoint from the band. Now do some twists to the band to get a new knot or link; then this new knot or link is in fact prime. This observation is made more precise in Corollary 3, stated in §7.

Corollaries 1, 2, 3 are proved via Theorems 2, 3, and tangle theory [L]. Corollary 1 is proved in §6, and Corollaries 2, 3 in §7. Theorem 3 has an interesting consequence on surgery of knots, specifically we have

Theorem 4. *Suppose k is a nontrivial strongly invertible knot in S^3 . If some surgery on k yields a reducible manifold then k is cabled and the surgery is via the slope of the cabling annulus.*

This is proved using Theorem 3, the fact that the manifolds obtained by surgery in a strongly invertible knot are double branched covers of S^3 [M], and the fact that a link is prime if and only if its double branched cover is irreducible [KT]. This is shown in §8.

Throughout the paper we assume familiarity with the theory and the notation of sutured manifolds as presented in [Sc₃]. The paper should be interpreted as being in the P.L. category. We refer to [R] or [H] for standard facts in knot theory and 3-manifolds.

I would like to express my sincerest gratitude to Professor Martin Scharlemann for his supervision of this work which formed part of my Ph.D. thesis at the University of California at Santa Barbara. I am thankful to the referee for his suggestions.

1. PRELIMINARY ARGUMENTS

1.1. Let β be the planar “eyeglass” 1-complex consisting of two circles β_0 and β_1 and an arc β_α joining them. Regard β as a complex in $\mathbb{R}^2 \subset \mathbb{R}^3$, and let U and W be regular neighborhoods of β in $\mathbb{R}^2$ and $\mathbb{R}^3$ respectively, so that U is a properly embedded planar surface in the genus two handlebody W . $\partial U \subset \partial W$ has three components: two of them, denoted by c_0 and c_1 , are parallel in U to β_0 and β_1 respectively. Denote the third by c_α . (See Figure 2.)

1.2. Let k be a link in S^3 and k_b be a banding of k . Let $M = S^3 - \text{int } \eta_k$. Embed β in M in such a way that β_α runs parallel to $b(\{1/2\} \times I)$, and that β_0 and β_1 are parallel to meridian of k ; we have that $\beta \cap \partial M = \emptyset$. Consider W as a regular neighborhood of β in M . Consider M as a sutured manifold, all of whose boundary is in R_+ or R_- , denote it by (M, γ, β) . Suppose $S^3 - (k \cup b[I \times I])$ is irreducible; this implies that (M, γ, β) is β -irreducible and ∂M is β -incompressible, so (M, γ, β) is β -taut. (M, γ) may not be $\emptyset$ -taut, i.e. may be reducible. ($\emptyset$ denotes the empty set, to be $\emptyset$ -taut means to be taut in the Thurston norm.) With this hypothesis we restate Theorem 1 as follows.

Theorem 1. *Let k be a link and b a band such that k_b is a composite link. Then either*

- (a) *there is a decomposing sphere for k_b disjoint from the band; or*
- (b) *there is a decomposing sphere which crosses the band exactly in one arc parallel to $b(\{1/2\} \times I)$; or*
- (c) *there is a swallow-follow torus for k_b disjoint from the band; or*
- (d) *(M, γ) is $\emptyset$ -taut, and there is a β -taut Seifert surface for k which is also $\emptyset$ -taut; or*
- (e) *$(M, \gamma) = (S^1 \times D^2, \emptyset)$ (i.e. k is the trivial knot), and k bounds a β -taut disk.*

Note that if (a) happens then (c) also happens, but not vice versa. We include both cases because they appear in different ways. (a) is just a trivial case.

It is not difficult to see that (d) and (e) of the above theorem imply (d) of Theorem 1 in the introduction, i.e. if a Seifert surface for k is taut in the Thurston norm then $\chi(S) = \chi(k)$ (for a proof see [ST₂, 1.2]).

The proof of this theorem will be as follows: First we take a β -taut Seifert surface S for k , and a decomposing sphere Q for k_b , which will be considered as a parameterizing surface; then construct a sutured manifold hierarchy,

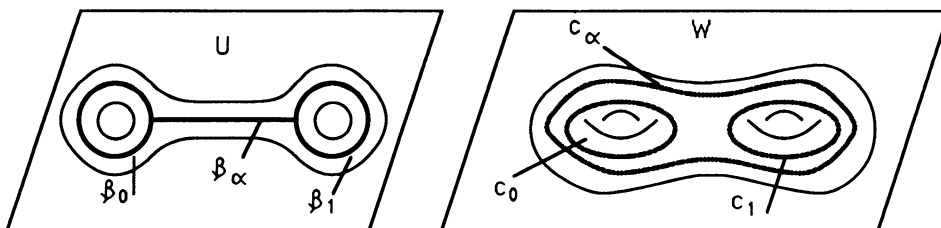


FIGURE 2

starting with S and respecting Q . If the final step in the hierarchy is $\emptyset$ -taut, we use [Sc₄, 2.7 and Sc₃, 3.3] to get (d) or (e). This is shown in §1. If the final step in the hierarchy is not $\emptyset$ -taut, then we do a lengthy combinatorial argument to see how this final step of the hierarchy may look; we rely heavily on 1.15 and 2.3. This is made in §§2–4. In §5 we use the information gathered to conclude (b) or (c).

1.3. Let P be a decomposing sphere for k_b . Suppose there is no decomposing sphere disjoint from the band, otherwise we are done. Suppose P intersects the band in a nonempty collection of arcs parallel to $b(\{1/2\} \times I)$; suppose this number of arcs taken over all decomposing spheres is minimal. If the band is trivial, then obviously there is a decomposing sphere disjoint from the band, so suppose the band is not trivial. Note that if there is a sphere meeting β in one point then the band is trivial.

Assume that the arcs of intersection between P and the band are contained in W (except by a tiny neighborhood of their endpoints). Let $Q = P \cap (M - \text{int } W)$. Q is a planar surface which has several components; some of them are annuli for which one boundary component is in ∂W and the other is a meridian of $\partial \eta(k)$. By eliminating parallel annuli we can suppose that $(Q, \partial Q) \subset (M - \text{int } W, \partial M \cup \partial W)$ consist of three components, Q_0 , Q_1 and Q_α with the following properties:

(a) Q_1 (Q_0) is an annulus for which one boundary component is $c_1 \subset \partial W$ ($c_0 \subset \partial W$) and the other is a meridian of $\partial \eta(k)$.

(b) Q_α is a connected planar surface, two of its boundary components are meridians in ∂M , and the others are parallel to c_α in ∂W .

According to [Sc₃, 7.1], Q is a parameterizing surface for (M, γ, β) . Note that Q is incompressible and ∂ -incompressible in $M - \text{int } W$. This is clear for Q_0 and Q_1 . If Q_α were compressible then we could make an isotopy in $S^3 - k_b$ to reduce the number of intersections between P and the band. If Q_α were ∂ -compressible, then by looking to the possible types of ∂ -compression disks and doing an argument similar to [E₂, 2.6], we could again reduce the number of intersections between P and the band.

1.4. Let S be a β -taut Seifert surface for k ; put S in normal position with respect to Q [Sc₃, 7.2]. (See [Sc₃, 7.7] for the definition of a sutured manifold decomposition respecting a parameterizing surface.) It is not difficult to see and is implicit in [Sc₃] that the sutured manifold decomposition $(M, \gamma) \xrightarrow{S} (M_1, \gamma_1)$ is β -taut and respects Q . Here we use the notion of sutured manifold hierarchy as presented in [Sc₄, 2.1]. Construct a β -taut sutured manifold hierarchy

$$(M, \gamma, \beta) \xrightarrow{S_1=S} (M_1, \gamma_1, \beta_1) \xrightarrow{S_2} \cdots \rightarrow (M_n, \gamma_n, \beta_n)$$

respecting Q [Sc₄, 2.5] (see also [Sc₃, 4.19 and Sc₃, 7.8]). ∂M_n is a collection of spheres. S meets β_0 (β_1) in one point and Q_0 (Q_1) in one arc. Following [Sc₃, 2.4(c), Sc₄, 2.1] the edges β_0 , β_1 and β_α can be oriented so that at any point of intersection with an S_i (hence with $R(\gamma_i) = R_+(\gamma_i) \cup R_-(\partial_i)$) the orientation points in the direction of the normal vector to S_i . We can suppose β has one of the orientations showed in the Figure 3 (the choice of one of them depends on the orientation of S_1). The argument for the two cases is more or less similar, and we will do both together until §4, where some differences



FIGURE 3

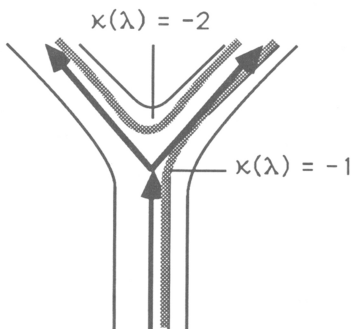


FIGURE 4

arise. Consider $R_+(\gamma_i)$ ($R_-(\gamma_i)$) as the part of ∂M_i in which the orientation points out of (into) M_i . Denote by Q_i and β_i the remnants of Q and β in M_i . Sometimes for simplicity β_i , $R_+(\gamma_i)$, $R_-(\gamma_i)$, $R(\gamma_i)$ will be denoted by β , R_+ , R_- , $R(\gamma)$ respectively.

1.5. Recall from [Sc₃, 7.4] what the index of a parameterizing surface is, $I(Q) = \nu + \mu + \mathcal{K} - 2\chi(Q)$, where ν is the number of sutures and μ the number of edges that ∂Q crosses. For each arc δ of $\partial Q \cap \eta(v)$, where v is a vertex of β , define $\kappa(\delta)$ to be -1 if δ passes between an edge of β pointing into the vertex and one pointing out, and define $\kappa(\delta) = -2$ if δ passes between edges of β both pointing into the vertex or both pointing out (see Figure 4). The curve $c_\alpha \subset \partial W$ has $\nu = 0$, $\mu = 4$, and $\mathcal{K} = -6$, and the curves c_0, c_1 each have $\nu = 0$, $\mu = 1$, $\kappa = -1$. Hence if Q_α has p boundary components in ∂W , $\chi(Q_\alpha) = -p$ and so $I(Q_\alpha) = -2p + 2p = 0$. Also $I(Q_0) = I(Q_1) = 0$.

1.6. Suppose that for an arc λ of β_n , which contains no vertices, there is a disk $D \subset Q_n$, such that $\partial D = \delta_1 \cup \delta_2$, where $\delta_1 \subset \eta(\lambda)$, $\delta_2 \subset \partial M_n$, and δ_2 crosses a suture; λ is called a cancelable arc, and D is called a canceling disk. Then there is a product disk D' which surrounds λ , namely $D' = \text{cl}(\partial \eta(D \cup \lambda) - \partial M_n)$.

The sutured manifold decomposition $(M_n, \gamma_n, \beta_n) \xrightarrow{D'} (M'_n, \gamma'_n, \beta'_n)$ is β -taut [Sc₃, 4.2] and it is not difficult to arrange it so that respects Q_n . Do enough of these operations so that the components of M_n which contain the vertices do not contain cancelable arcs. Still denote the last step as M_n . (Compare to [Sc₃, 4.3] in which a slight different procedure is made to eliminate cancelable arcs.)

1.7. Denote by $Q_{i,n}$, $i = 0, 1$, the parameterizing surface obtained from Q_i at the end of the hierarchy. Similarly denote by $\beta_{i,n}$, $i = 1, 0$, the 1-complex obtained from β_i at the end of the hierarchy. The parameterizing surface $Q_{i,n} \subset Q_i$, $i = 0, 1$, has $I(Q_{i,n}) \leq 0$ [Sc₃, 7.5, 7.6].

Claim. Each component q of $Q_{i,n}$ for which $\partial Q_{i,n} \cap \eta(\beta_{i,n}) \neq \emptyset$ is a disk for which $I(q) = 0$. Indeed $\partial q \cap \beta_{i,n}$ is a single arc and $\partial q \cap \partial M_n$ is a single arc crossing a single suture.

Proof. This is similar to [ST₁, Claim 2]. Note that the surface S_1 meets β_0 and β_1 . $\square$

This claim implies that there is a component q of $Q_{i,n}$, q in a disk, such that ∂q runs through one of the vertices; all the other components of $Q_{i,n}$ are canceling disks for an arc of $\beta_{i,n}$. So the bits of β_0 and β_1 which do not contain a vertex of β can be eliminated as in 1.6, via the remnants of $Q_{i,n}$.

1.8. A component of M_n which does not meet β is a 3-ball with a single suture in its boundary. We will disregard these trivial components and suppose with no loss of generality that all the components of M_n meet β . As a consequence of this we conclude that every disk component of $R_{\pm}$ intersects β . Note that a component of ∂M_n has sutures, for if there is a sphere S without sutures, it has to meet β . But S only can meet β_{α} , for all the arcs of β_0 and β_1 meet components of ∂M_n which contain sutures. For β -tautness β_{α} has to meet S always in the same direction, so if β_{α} meets S twice or more it will imply that there is a nonseparating sphere in S^3 , which is impossible; if β_{α} meet S once then the band has to be trivial, contradicting the hypothesis.

1.9 *Claim.* $\chi(R_+(\gamma_n)) = \chi(R_-(\gamma_n))$.

Proof. $R_{\pm}$ contains no closed component, so by [Sc₃, 1.1], $\chi_{\beta}(R_{\pm}) = |R_{\pm} \cap \beta| - \chi(R_{\pm}) + |\text{disks in } R_{\pm} \text{ disjoint from } \beta|$. By 1.8 R_+ and R_- have the same number of disks disjoint from β . Also in our case $|R_+ \cap \beta| = |R_- \cap \beta|$. By β -tautness of (M_n, γ_n, β_n) , $\chi_{\beta}(R_+) = \chi_{\beta}(R_-)$. This implies that $\chi(R_+) = \chi(R_-)$. $\square$

1.10 *Claim.* Either (M, γ) satisfies (d) or (e) of Theorem 1 or (M_n, γ_n) is not $\emptyset$ -taut.

Proof. If (M_n, γ_n) is $\emptyset$ -taut then by [Sc₄, 2.7] (M_i, γ_i) is $\emptyset$ -taut for all $i \geq 1$, and either (M, γ) is $\emptyset$ -taut, and so is irreducible and ∂ -incompressible or M is a solid torus (i.e. k is the trivial knot). Now in any case, by [Sc₃, 3.3] S is $\emptyset$ -taut. This implies (d) or (e) of Theorem 1. $\square$

So in what follows we made the following

1.11 *Assumption.* (M_n, γ_n) is not $\emptyset$ -taut.

1.12. ∂M_n is a collection of spheres, so (M_n, γ_n) is $\emptyset$ -taut if and only if each sphere has only one suture and it is the boundary of a 3-ball. Note that if a component of M_n has only one sphere in its boundary, then this sphere bounds a 3-ball, for M_n is contained in S^3 .

1.13 *Claim.* Some of the surfaces S_i intersect β_{α} .

Proof. Suppose that none of the S_i 's meets β_{α} .

Case 1. The boundary of the component of M_n which contains β_{α} is a sphere, say T .

By 1.7, 1.8 there are only two disks components in $R(\gamma)$. If there is only one suture, then (M_n, γ_n) is $\emptyset$ -taut, which contradicts 1.11. So T contains more than one suture. Then $R(\gamma)$ consist of two disks and a collection of annuli. By 1.9 $\chi(R_+) = \chi(R_-)$, so it follows that R_+ (R_-) consist of one disk and m

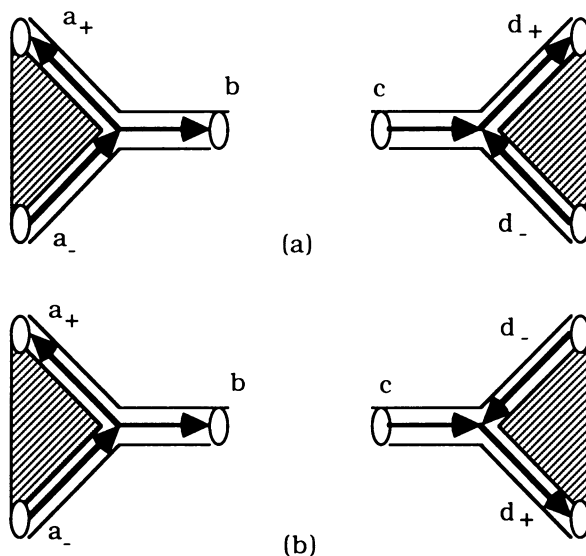


FIGURE 5

annuli. Then the innermost annulus A of R_- must have β -norm equal to 1, because it meets β once. Using the component q of $Q_{i,n}$ described in 1.7, it is not difficult to see that there are two disks D_1, D_2 (take disks parallel to the innermost disk in R_+) such that $\partial(D_1 \cup D_2) = \partial A$, $[D_1 \cup D_2] = [A]$ in $H_2(M_n, \chi(\partial A))$, and each one meets β once. Then $\chi_\beta(D_1 \cup D_2) = 0$, but that implies that R_- is not β -taut.

Case 2. ∂M_n has two components, say two spheres T_1, T_2 . A sphere parallel to T_1 meets β_α in one point, implying that the band is trivial. $\square$

1.14. Following 1.13 there are two components of β_n which have a vertex. Denote the component which has two ends in R_+ (R_-) and one in R_- (R_+) by A_+ (A_-). Denote the ends of A_+ by a_+, a_-, b , and the ends of A_- by d_+, d_-, c , where b, c are the ends which are part of β_α . a_+, d_+, b lie on R_+ , and a_-, d_-, c lie on R_- . a_+, a_- (d_+, d_-) are part of β_0 (β_1). 1.7 shows that a_+ and a_- (d_+ and d_-) lie in adjacent components of $R_\pm$ and there is a disk q component of $Q_{0,n}$ ($Q_{1,n}$) such that $\partial q \cap \partial M_n$ is an arc joining a_+ and a_- (d_+ and d_-) which crosses a suture. See Figure 5.

There is a collection of arcs on A_+ going from a_+ to b ; each such arc λ is contained in the boundary of a component of Q_n , and for this arc $\kappa(\lambda) = -2$. Call the part of A_+ which contains these arcs the negative side of A_+ . Analogously, there is a collection of arcs on A_+ going from a_- to b . For each such arc λ , $\kappa(\lambda) = -1$. Call the part of A_+ which contains these arcs the positive side of A_+ . In a similar way we define the negative and positive side of A_- .

1.15 *Claim.* No component of Q_n has negative index.

Proof. If q is a component of Q_n with $I(q) = \nu + \mu + \mathcal{K} - 2\chi(q) < 0$, then because $\nu + \mu + \mathcal{K} \geq 0$, it follows that $\chi(q) > 0$, so q is a disk, $\nu + \mu + \mathcal{K} = 0$, $I(q) = -2$, and then ∂q crosses no suture, and the only arcs of β that it can cross are the negative side of A_+ or A_- (it cannot cross both of them

because then ∂q would cross sutures or other arcs). So ∂q consist of arcs which are contained in R_+ , say, and arcs over A_+ which go from a_+ to b . An application of [Sc₃, 6.2] shows that either there are compressible β -loops [Sc₃, 6.1] for Q_α in a_+ or b , or there is a sphere intersecting β in one point, or $\partial q \subset R_+$. The first contradicts the ∂ -incompressibility of Q_α , the second implies that the band is trivial, and the third contradicts the fact that q is part of a parameterizing surface. $\square$

1.16 *Claim.* Each component of Q_n is a disk or annulus of index zero.

Proof. $I(Q_n) \leq I(Q) = 0$ and Q_n has no negative index components, then $I(Q_n) = 0$, and each component of Q_n has index zero. For q a component of Q_n , $I(q) = \nu + \mu + \mathcal{K} - 2\chi(q) = 0$, but $\nu + \mu + \mathcal{K} \geq 0$, so $\chi(q) \geq 0$, and then q is a disk or annulus. $\square$

Note that if q is an annulus component of Q_n , then ∂q crosses no suture, and the only arcs of β that it can cross are the negative side of A_+ or A_- .

The following lemma will avoid possible technical problems in §4.

1.17 **Lemma.** *It can be supposed that in each step of the sutured manifold hierarchy $Q_i = Q_{i-1} - \text{int } \eta(S_i)$.*

Proof. According to [Sc₃, 7.5, 7.6] that is the case except possibly when a product disk, product annulus or a disk with boundary in $R(\gamma)$ was used in the hierarchy. In [Sc₃, 7.6] is explained what changes are made to the parameterizing surface in that case, but in general these changes decrease the index. Let S_i be the first of such surfaces used in the hierarchy. If $I(Q)$ does not drop then either there are no arcs of intersection between S_i and Q_{i-1} whose two ends lie in $R_+(R_-)$, or each such arc bounds a disk in Q_{i-1} which crosses no suture and no arc of β . In the last case S_i can be isotoped to not intersect Q_{i-1} . After this move the manifold M_i looks the same as if we had made the changes mentioned in [Sc₃, 7.6]. Do the same to each of these surfaces. $\square$

This observation could be made for any β -taut sutured manifold hierarchy in which the index of the parameterizing surface does not decrease.

2. THE COMBINATORIAL SETTING

2.1. Let T be one of the sphere components of ∂M_n . The points $\beta \cap T$ and the arcs $\partial Q_n \cap T$ can be regarded as a graph Γ in T . A vertex of Γ is a point of $\beta \cap T$ and an edge is an arc component of $\partial Q_n \cap T$, each one of its ends is at a vertex. Denote the components of $\partial Q_\alpha - \partial M$ by $a_1, a_2, \dots, a_p$, labeled so that a_i and a_{i+1} bound an annulus in ∂W disjoint from Q_α , for $1 \leq i \leq p-1$. Denote by $\varepsilon_a = \partial Q_{0,n} \cap T$ ($\varepsilon_d = \partial Q_{1,n} \cap T$) the edge joining a_+ and a_- (d_+ and d_-), this edge exists for 1.7, 1.14; it crosses a suture. Label the end of an edge in Γ , other than ε_a and ε_d , with i (i^*) if this point is over $a_i \cap \beta$ in the top (bottom) part of β (top and bottom according to Figures 3 and 5). The collection of labels around a vertex in Γ , other than $a_\pm, d_\pm$, looks like $1, 2, \dots, p, p^*, \dots, 1^*$; the collection of labels around $a_\pm, d_\pm$ looks like $1, 2, \dots, p$ or $1^*, 2^*, \dots, p^*$. An edge is level if their ends are equally labeled (ignoring the asterisk). a_+, a_-, b, c, d_+, d_- are named special vertices, any other is called a simple vertex. Call the part of a special vertex which is contained in the negative (positive) side of A_+ or

A_- the negative (positive) side of such vertex (see 1.14). b and c have both negative and positive sides, a_+ and d_- have only negative side, a_- and d_+ have only positive side.

Give an orientation to Q_α ; this orientation induces an orientation in the boundary components of Q_α . Two components of $\partial Q_\alpha - \partial M$ are parallel if with the induced orientation they are homologous in ∂W , otherwise they are antiparallel. Put an orientation on T , for example the orientation induced by R_+ , this induces an orientation on the vertices of Γ . Let v_1, v_2 be two vertices of Γ , other than $a_\pm, d_\pm$. v_1 and v_2 are parallel if with the induced orientation they are homologous on $\partial\eta(\beta_\alpha) \cap \partial W$, otherwise they are antiparallel. v_1 is parallel to a_+ or a_- if v_1 is homologous to $a_+ \cup a_-$ in $\partial\eta(\beta_\alpha \cup A_+) \cap \partial W$. Analogously for d_+, d_- . $a_\pm$ is antiparallel to $d_\pm$ for $a_+ \cup a_-$ is not homologous to $d_+ \cup d_-$ in $\partial\eta(A_+ \cup \beta_\alpha \cup A_-) \cap \partial W$. Then all the vertices in R_+ (R_-) with the exception of a_+ (d_-) are parallel. A vertex in R_+ is antiparallel to a vertex in R_- (other than a_+, d_-). We have the following parity rule [Sc1, 2.2]. Let v_1, v_2 be vertices of Γ , $i, j \in \{1, \dots, p\}$.

(1) If an edge joins parallel vertices v_1 and v_2 with labels i and j (i and j^* , or i^* and j^*) respectively, then a_i and a_j are antiparallel (resp. parallel, resp. antiparallel).

(2) If an edge joins antiparallel vertices v_1 and v_2 with labels i and j (i and j^* , or i^* and j^*) respectively, then a_i and a_j are parallel (resp. antiparallel, resp. parallel).

2.2. Let Γ be a subgraph of Γ , and let $x \in \{1, 2, \dots, p, 1^*, \dots, p^*\}$. We say that Λ satisfies condition $P(x)$ if (see [CGLS, 2.6]):

For each vertex y of Λ there exist an edge of Λ incident to y with label x , connecting y to a parallel vertex.

For $x \in \{1, \dots, p^*\}$ an x -path is a path in Γ so that the beginning point of each edge is labeled with x , and all its vertices are parallel. An x -cycle is an x -path which is a cycle. A Scharlemann cycle is an x -cycle which is the boundary of a disk without edges nor vertices in its interior. If a subgraph Λ of Γ satisfies $P(x)$ for some x , then it is possible to construct an x -path beginning at a vertex, and then this will form an x -cycle in Λ .

2.3 Lemma. *Let Λ be a subgraph of Γ which consist of the intersection of Γ with a disk, and such that all its vertices are parallel, and that satisfies condition $P(x)$ for some x , then Λ contains a Scharlemann cycle.*

Proof. It is a straightforward modification of [CGLS, 2.6.1, 2.6.2], or [BE, 4.4]. $\square$

2.4 Claim. There is no Scharlemann cycle in Γ .

Proof. If Γ contains a Scharlemann cycle, then either is a loop, which contradicts the ∂ -incompressibility of Q_α , or it has at least two edges, and then a modification of [Sc2, 5.6] shows that there is a lens space summand in S^3 . One of these Scharlemann cycles may be a cycle in which every edge is labeled with 1 and 1^* (or n and n^*), i.e. a pure level circuit in the terminology of [BE], but in contrast with the argument there, here a pure level circuit can be used to construct a lens space because all its vertices are parallel. $\square$

2.3 and 2.4 imply that there is no subgraph of Γ consisting of the intersection of Γ with a disk, satisfying $P(x)$ and with all its vertices being parallel.

2.5 Lemma. Suppose that D is a disk component of $R(\gamma)$ which does not contain $a_{\pm}$ nor $d_{\pm}$. Then for each $x \in \{1, \dots, p^*\}$ there is at least one edge with label x in D which crosses the suture surrounding D . So there are at least $2p$ arcs crossing that suture.

Proof. If this does not happen for some x then $\Lambda = D \cap \Gamma$ has $P(x)$, contradicting 2.3, 2.4. $\square$

2.6 Lemma. If a disk D of R_+ (R_-) contains only one of a_+ , d_+ (a_- , d_-), then there are at least p edges, other than ε_a , ε_d , crossing the suture surrounding D .

Proof. If a_- or d_+ are in D , then the proof is as in 2.5, for they are parallel to all the other vertices in D . If a_+ (d_-) is in D , and there are no more vertices there then the conclusion is obvious. If there are more vertices in D then there are at least p labels such that no edge adjacent to a_+ (d_-) has one of these labels in its other end, now do for these labels the same argument as in 2.5. $\square$

The following two lemmas will be applied in 3.7, 3.11, 4.1, 4.2, 4.8.

Let D be a disk contained in T so that ∂D cut transversally two special vertices, and so that either the positive or negative part of such vertices is in D . Suppose also that all the vertices in the interior of D are simple, and all the vertices in D are parallel, including the special vertices in the boundary. Suppose that no edge crosses ∂D . Let $\Lambda = \Gamma \cap D$. Call the labels of the two special vertices which meet ∂D , the labels of ∂D .

2.7 Lemma. Suppose ∂D is labeled with $\{1, \dots, p, p^*, \dots, 1^*\}$, in this order, and Λ looks like Figure 6. Then either there is a Scharlemann cycle which intersects at most one side of ∂D , or for each $i \in \{1, \dots, p\}$ there is an i -path beginning at one side of ∂D and finishing in the other side of ∂D with label i^* . Furthermore all the edges adjacent to ∂D have ends labeled $i-i^*$, i.e. they are level.

Proof. Suppose there is no such Scharlemann cycle. For the label 1, take a 1-path γ_1 beginning at ∂D , this path has to finish at the other side of ∂D , otherwise there is a 1-cycle (and by 2.3 a Scharlemann cycle); so it finishes at i_1^* ($i_1^* \geq 1$). The complementary part of γ_1 which contains the other labels j 's of ∂D has $P(2)$, because all the vertices are parallel. Take now a

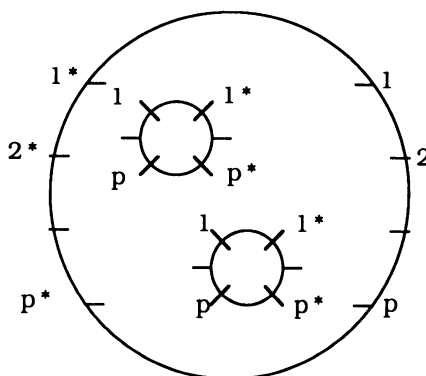


FIGURE 6

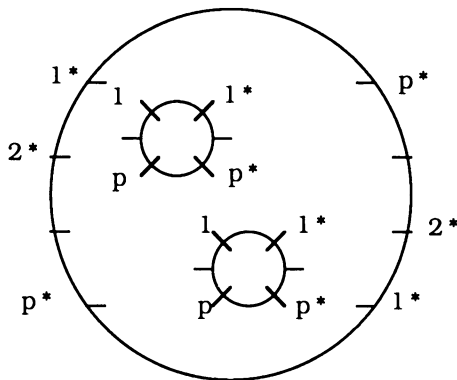


FIGURE 7

2-path γ_2 beginning at ∂D , it finishes at ∂D with label i_2^* , $i_2^* > i_1^*$, and the complementary part of γ_2 has $P(3)$. So by induction we can take a k -path γ_k beginning at ∂D , which finishes at ∂D with label i_k^* , $i_k^* > i_t^*$ if $k > t$, and the complementary region of γ_k has $P(k+1)$. This implies $i_k^* = k^*$ for all k , so the last edge of each of these paths has labels $k - k^*$. Repeating the argument, but now beginning the paths at i^* in ∂D , one conclude that the edges of ∂D with label i , have the other end with label i^* . $\square$

2.8 Lemma. Suppose ∂D is labeled with $\{1^*, \dots, p^*, 1^*, \dots, p^*\}$, in this order (or $\{1, \dots, p, 1, \dots, p\}$), and Λ looks like in Figure 7. Then there is a Scharlemann cycle in D ; if this cycle contains edges which meet ∂D then the cycle does not contain the part of ∂D between the labels 1^* and p^* (1 and p).

Proof. Do the same as in 2.7, construct the paths γ_k . After reaching the other end of ∂D this path can be continued implying the existence of a k^* -cycle, and hence a Scharlemann cycle, unless this path ends with an edge labeled $k^* - k^*$ which is not possible for it contradicts the parity rule. $\square$

3. POSSIBLE CONFIGURATIONS OF (M_n, γ_n, β_n)

The goal of this and next section is to show that there are two vertices of Γ joined by p parallel level edges, one for each label, and such that these edges when viewed in Q are bad loops (see 4.3). In this section we prove that each sphere of ∂M_n has at most two disks components of $R(\gamma)$ and list a finite number of possible configurations of the arcs of β in M_n . In this and subsequent sections $R(\gamma)$, R_+ , R_- will denote $R(\gamma_n)$, $R_+(\gamma_n)$, $R_-(\gamma_n)$ respectively.

3.1 Claim. A component N of M_n which does not contain a vertex of β is a 3-ball with one suture on its boundary, so is $\emptyset$ -taut.

Proof. The proof is by induction on the number r of arcs of β that are contained in N . If $r = 0$ then by β -tautness N is as required. So suppose $r > 0$. Note that by 1.8 each component of ∂N has sutures. There are two possibilities:

(a) There is in N an arc β' which is part of β_0 or β_1 . Then β' is a cancelling arc via Q_0 or Q_1 , and doing the operation described in 1.6, we get

N' , which contains $r - 1$ arcs of β . By induction N' is as required, and then it is not difficult to see that N is also a 3-ball and is $\emptyset$ -taut.

(b) All the arcs in N are part of β_α . Let D be a disk in $R(\gamma)$; by 2.5 there is at least one edge crossing the boundary of D , this edge has to be part of an index zero disk F , and as F does not meet the vertices of β , ∂F crosses a suture and goes through an arc β' of β , so F is a cancelling disk for β' . Now proceed as in case (a). $\square$

We are assuming that M_n is not $\emptyset$ -taut (1.11), and so by 3.1 we can assume now that all the components of M_n contain a vertex of β .

3.2 Claim. A disk D in $R_\pm$ must contain special vertices.

Proof. Suppose this is false. Suppose w.l.o.g. that D is in R_+ . By 2.5 there are at least $2p$ edges crossing the suture ∂D . Each of these edges belong to an index zero disk, and then these edges have to meet c or d_- in its negative part, otherwise these disks would be cancelling disks. But there are only p negative labels in c , and if a disk meet one of these labels it also meets the corresponding label in d_- , so in $c \cup d_-$ there are room only for p edges coming from D , so the other p edges have to be part of cancelling disks, which is a contradiction. $\square$

Claim 3.2 implies that a component of ∂M_n has at most four innermost sutures, i.e. sutures which bound a disk contained in $R_\pm$, since $a_\pm$ ($d_\pm$) lie in adjacent components of $R_\pm$.

3.3 Claim. If a disk D in $R_\pm$ contains a special vertex, then the component adjacent to it also has a special vertex.

Proof. This is clear if D contains one of a_+ , a_- , d_+ , d_- . So suppose that D contains b and no other special vertex. By 2.5 there are at least $2p$ edges which cross the suture ∂D . At most p edges come from the negative side of b , so at least p edges come from a simple vertex or the positive side of b . If an edge come from the positive side of b , its other end meets either a_- or the negative side of a special vertex, and if the edge come from a simple vertex then its other end has to meet the negative side of a vertex, so there is a special vertex in the component adjacent to D . $\square$

3.4 Claim. A disk D in $R_\pm$ contains at most two special vertices.

Proof. Suppose D contains three special vertices, say a_+ , d_+ and b . The component E adjacent to D will contain a_+ and d_- . If c is in E , then by 3.2 there are no more sutures, and then E is a disk, so M_n is a 3-ball with a suture in its boundary, and so is $\emptyset$ -taut by 1.12, which contradicts Assumption 1.11. If c is not in E , then there is another disk D' such that c is in D' . D' cannot be adjacent to E , since both are in R_- , so in the area adjacent to D' there is no special vertex which contradicts 3.3. $\square$

3.5 Claim. A component T of ∂M_n has at most three disks components of $R(\gamma)$.

Proof. By 3.2 T has at most four innermost sutures. So suppose T has four disks of $R(\gamma)$, call them D_1, D_2, D_3, D_4 . Then, say, b is in D_1 , c in D_2 , $a_\pm$ in D_3 , $d_\pm$ in D_4 . Let E_- (E_+) be the area adjacent to D_1 (D_2), it contains special vertices by 3.3. Then E_- (E_+) contains either a_- or d_- (a_+ or d_+)

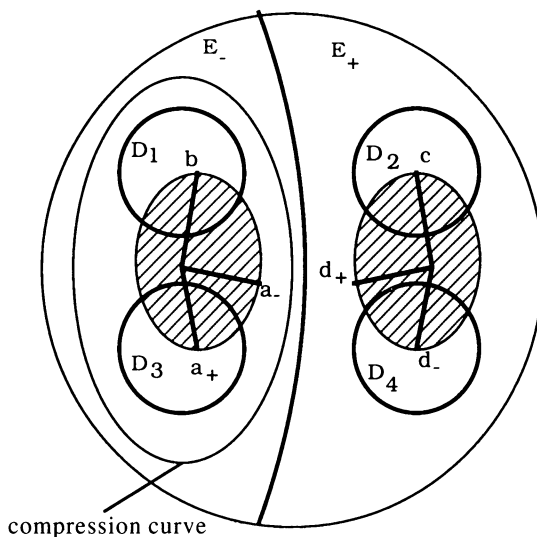


FIGURE 8

and must be adjacent to two disks components of R_+ (R_-). It follows that E_- , E_+ are three-punctured spheres.

Case 1. d_- is in E_- .

Then d_+ is in D_4 , and E_- is adjacent to D_4 . There are, by 2.5, at least $2p$ edges which cross ∂D_1 , and p of them have to meet d_- , for only p may come from the negative side of b . By 2.6 there are p edges which cross ∂D_4 , and they have to meet a special vertex; but there is no special vertex in E_- other than d_- , and there is no place in d_- for these p edges. So this case it is not possible.

Case 2. a_- is in E_- .

Then a_+ is in D_3 , and E_- is adjacent to D_3 . There are $2p$ edges which cross ∂D_1 . If they come from simple vertices, then they must meet the negative side of a vertex, but there is no such vertex in E_- . Then the $2p$ edges come from b ; p of them have to join the positive side of b and a_- . By 2.6 there are p edges which cross ∂D_3 , they cannot come from simple vertices, so the only possibility left is that these edges come from a_+ , cross ∂D_3 , ∂D_1 and then meet the negative side of b . This argument implies that there are no simple vertices in D_1 and D_3 . A similar situation happens in D_2 , D_4 and E_+ . We have a situation as in Figure 8; now it is not difficult to see that E_- is then compressible, which contradicts β -tautness. $\square$

3.6 Lemma. *Let D be a disk in $R_{\pm}$. Suppose that the area adjacent to D is an annulus E . Suppose a_+ (d_-) is the only special vertex in D . Then D contains simple vertices and E contains either c or d_- (b or a_+).*

Proof. a_- (d_+) is in E .

Claim 1. There are simple vertices in D .

If a_+ is the only vertex in D , then E is not β -taut. To see that consider two disks D_1 , D_2 , so that $\partial D_1 \cup \partial D_2 = \partial E$, and $[E] = [D_1 \cup D_2]$, these disks can be chosen so that D_1 meets β in one point and D_2 meets β in as many points

as E does, say n . Then $\chi_\beta(E) = n$, but $\chi_\beta(D_1 \cup D_2) = n - 1$. This contradicts the β -tautness of E . Note that this construction can be done because of the existence of a disk in $Q_{0,n}$ whose boundary go through a_+ , a_- and crosses a suture (cf. 1.7).

Claim 2. No edge adjacent to a_+ crosses ∂D (other than ε_a).

Suppose an edge η adjacent to a_+ crosses ∂D . At most $p - 1$ edges adjacent to a_+ meet simple vertices in D , so there are $p + 1$ labels which are not the other end of an edge adjacent to a_+ ; for each one of these labels we can find an edge which crosses ∂D (the argument is similar to that of 2.6) and then there will be at least $p + 1$ edges coming from simple vertices which cross ∂D . These edges have to meet the negative side of a vertex, but in the negative side of a vertex there is room only for p edges, so this is not possible.

Then the p edges which cross ∂D come from simple vertices and have to meet the negative side of a vertex, the lemma follows. $\square$

3.7 Lemma. Let D be a disk in $R_\pm$. Suppose that the area adjacent to D is an annulus E . Then $a_-(d_+)$ cannot be the only special vertex in D .

Proof. Suppose a_- is the only special vertex in D . Then a_+ is in E . D contains simple vertices and no edge adjacent to a_- crosses ∂D (other than ε_a). The proof is similar to that in 3.6. By 2.6, and because at a special vertex there is room only for p edges, there are exactly p edges which cross ∂D (other than ε_a).

Claim. The p edges which cross ∂D meet the negative side of b .

The other possibility for these edges is that they meet a_+ . For each i , $i = 1, \dots, p$, take a i^* -path λ_i starting at a_- . Each path has to cross ∂D , and then meet b or a_+ . This produce the p edges which cross ∂D . So suppose that for some i , λ_i finishes at a_+ . λ_i is an i^* -path in $D \cup E$ which starts in a_- and finishes in a_+ , and so that $\lambda_i \cup \varepsilon_a$ bound a disk $D' \subset D \cup E$. d_+ is not in D' , for then d_- would be in D . There is no other vertex in the interior of $E \cap D'$, for if there are vertices, less than p edges join a_+ and these vertices. So there are more than p labels j in these vertices, at which no edge coming from a_+ is incident, so for each such label there is a j -path in $E \cap D'$ which has to cross the part of ∂D lying in D' , that is, there are more than p edges crossing ∂D , but this is not possible. Either the label $(i - 1)^*$ or $(i + 1)^*$ at a_- lie in D' , say the label $(i - 1)^*$, i as above (possibly $1^* = i^*$). It is not difficult to see that λ_{i-1} lies in D' , and so it finishes at a_+ . Continuing in this way as in the proof of 2.7, we conclude that either there is a Scharlemann cycle in D , or λ_1 consists of level edges, starts at a_- and finishes at a_+ , and $\lambda_1 \cup \varepsilon_a$ bound a disk D'' with interior disjoint from Γ . D'' can be used, as in 2.4, to construct a lens space in S^3 , which is impossible. So we conclude that the p edges which cross the suture ∂D meet the negative side of b .

If $p = 1$ there is an edge with label 1^* at a simple vertex in D and label 1 at b , which contradicts the parity rule. So suppose $p > 1$. Let η_1 (η_p) = (edge adjacent to b at label 1 (p)) $\cap E$, and let α be the arc in b , which go through the negative side of b , and which joins the labels 1 and p . Then $\eta_1 \cup \eta_p \cup \alpha \cup$ (arc on ∂D joining η_1 and η_p) bounds a disk $F \subset E$. We have two cases:

Case 1. The positive side of b is not in the interior of F .

Then all the edges adjacent to the negative side of b lie in F . No vertex is in the interior of F , for if there is one, then by 2.3 we could find a

Scharlemann cycle inside F . The negative side of b has labels $1, 2, \dots, p$, and b is antiparallel to all vertices in D . Relabel the negative side of b by $1^*, 2^*, \dots, p^*$; with the new labeling the negative side of b appears as a vertex parallel to all the vertices in D . Note that the parity rule still holds under the new labelling. Now an application of 2.8—where a_- and the negative side of b are considered the boundary vertex—yields a contradiction.

Case 2. The positive side of b is in the interior of F .

Then no edge adjacent to the negative side of b is in the interior of F . a_+ must lie in F , for otherwise as the positive side of b is in F , we could find a Scharlemann cycle in F . a_+ has labels $1, 2, \dots, p$, and it is antiparallel to all vertices in F . Relabel a_+ by $1^*, 2^*, \dots, p^*$; with the new labeling a_+ appears as a vertex parallel to all the vertices in F . Now an application of 2.8—where a_+ and the positive side of b are considered the boundary vertex—yields a contradiction. $\square$

3.8 Lemma. *A component T of ∂M_n has at most two innermost sutures.*

Proof. By 3.5 a component T of ∂M_n has at most three innermost sutures, so suppose T has three disks D_1, D_2, D_3 components of $R(\gamma)$. The other components of $R(\gamma)$ are annuli and a three punctured sphere F . By 1.9 $\chi(R_+) = \chi(R_-)$, so necessarily the disks D_1, D_2 and F belong to the same part of $R(\gamma)$, R_+ or R_- , say R_+ , and the remaining disk D_3 belong to R_- . Then F cannot be adjacent to the disks D_i , $i = 1, 2$. The area adjacent to D_1 and D_2 has to consist of annuli. We can assume that b is in D_1 or D_2 , for otherwise d_+ would be the only special vertex in one of D_1 or D_2 contradicting 3.7. So without loss assume that b is in D_1 . If D_2 contains both special vertices a_+ and d_+ , then c is in D_3 , and there will not be a special vertex in the area adjacent to D_3 , contradicting 3.3. So by 3.7 D_2 contains only a_+ , and a_- is in the adjacent annulus.

Case 1. c is in D_3 .

Then by 3.6 the annulus adjacent to D_2 contains d_- as well as a_- . Then the annulus adjacent to D_1 could contain no special vertex, again contradicting 3.3.

Case 2. c is not in D_3 .

Then d_- is in D_3 and by 3.6 c is in the annulus adjacent to D_2 . Again there is no special vertex in the annulus adjacent to D_1 . $\square$

3.9 Lemma. *Let D be a disk in $R_{\pm}$. Suppose that the area adjacent to D is an annulus E . Suppose a_- (d_+) is in E , and $b(c)$ is the only special vertex in D . Then either c or d_- (b or a_+) are in E .*

Proof. Suppose a_- is the only special vertex in E . There are $2p$ edges which cross the suture ∂D , none can come from a simple vertex, for there is no special vertex with negative side in E , so the $2p$ edges crossing ∂D are adjacent to b . It follows that there is no simple vertex in D . The p edges adjacent to the positive side of b have to meet a_- . There is an index zero disk q , with ∂q running from a_- to b along β and then running across a suture. Then using q (as in 3.6, Claim 1), we can find two disks D_1, D_2 , so that $\partial D_1 \cup \partial D_2 = \partial E$, and $[E] = [D_1 \cup D_2]$. These disks can be chosen so that D_1 meets β in one point and D_2 meets β in as many points as E does, say n . Then $\chi_{\beta}(E) = n$, but $\chi_{\beta}(D_1 \cup D_2) = n - 1$. This contradicts the β -tautness of E . Then c or d_- are in E . $\square$

3.10. Suppose first ∂M_n has only one boundary component. This boundary component has two innermost sutures, which determine two disks, and by 1.9 one is in R_+ and one in R_- , denoted by D_+ and D_- respectively. Denote by E_- , E_+ the annulus adjacent to D_+ and D_- respectively. Note that there may be more annuli between E_+ and E_- . D_+ and D_- contain special vertices by 3.2, and E_+ and E_- also contain special vertices by 3.3.

3.11 *Claim.* D_+ (D_-) does not contain both a_+ and b (d_- and c).

Proof. Suppose D_+ contains a_+ and b . It contains no other special vertex by 3.4. a_- is in E_- .

Subclaim 1. No special vertex is in E_- , other than a_- .

If c or d_- are in E_- then d_- or c will be the only special vertex in D_- and E_+ will contain only d_+ , contradicting 3.6 or 3.9.

Subclaim 2. There is no edge joining the positive side of b and a_- .

Suppose η is such an edge, joining a_- and b and crossing ∂D_+ . So there is a zero index disk q whose boundary consists of η plus an arc over β going from a_+ to b (this is the only possibility, for otherwise $I(q) > 0$). η plus the arc joining a_+ and a_- divide the complement of D_+ in two parts. One of these parts is a disk E' contained in E_- . If E' contains simple vertices then there would be more than p edges coming from simple vertices and crossing ∂D_+ , but in D_+ there is room only for p of these edges, so no vertex is in E' . It is not difficult to see, using q and $Q_{0,n}$, that there is a disk $F \subset M_n$, so that $\partial F = \alpha \cup \lambda$, α is an arc over D_+ joining a_+ and b , λ is an arc over β , going from a_+ to b , $(\text{int } F) \cap \beta = \emptyset$; use F to isotope D_+ and find a surface homologous to D_+ but intersecting β in fewer points, contradicting the β -tautness of D_+ (this is the same argument as in [Sc₃, 6.2], which was applied in 1.15).

Subclaim 3. No edge adjacent to the positive side of b or a simple vertex lying in D_+ crosses ∂D_+ .

If there were one it had to meet the negative side of a special vertex lying in E_- , but there is no such special vertex.

Subclaim 4. No edge adjacent to a_+ crosses the suture ∂D_+ (other than ε_a).

If there is one, then by an argument similar to 3.6, Claim 2, there are $p + 1$ edges which cross ∂D_+ , at least one of them has to come from the positive side of b or from a simple vertex, which contradicts Subclaim 3.

Then the p edges adjacent to the negative side of b are the only edges which cross ∂D_+ . a_+ has labels $1, 2, \dots, p$ and is antiparallel to all other vertices in D_+ . Relabel a_+ by $1^*, 2^*, \dots, p^*$, then now it appears like a vertex parallel to all other vertices in D_+ . An application of 2.8—where a_+ and the positive side of b are considered the boundary vertex—yields a contradiction. $\square$

3.12 *Claim.* Suppose that a_+ is the only special vertex in D_+ , then we have the following possible cases:

Case 1. c is in D_- , a_- and d_- are in E_- , b and d_+ are in E_+ .

Case 2. d_- is in D_- , a_- and c are in E_- , d_+ and b are in E_+ .

Proof. If a_+ is the only special vertex in D_+ , then by 3.6 there are simple vertices in D_+ and either c or d_- are in D_+ , but not both of them. If c is in E_- , then d_- is in D_- , d_+ is in E_+ , and b has to be in E_+ too (by 3.6), so

we have Case 2. Note that in this case there may be more annuli between E_+ and E_- .

Suppose now d_- is in E_- and c is in D_- . There are $2p$ edges which cross the suture ∂D_- .

Subclaim. No edge adjacent to the negative side of c crosses ∂D_- .

Proof. There are p edges which come from D_+ and meet d_- , all of them come from simple vertices (3.6, Claim 2). Let η be one of these edges. There is an index zero disk q which contains η in its boundary, ∂q also contains an arc on β joining simple vertices, an arc on β going from d_- to c , and an arc joining the negative side of c and a simple vertex, which crosses no suture (this is the only possibility for otherwise $I(q)$ would be positive or there would be a loop in D_- , contradicting 2.4). This implies that the edges adjacent to the negative part of c meet a simple vertex lying in D_- , so they do not cross ∂D_- .

There are $2p$ edges which cross the suture ∂D_- , which come from simple vertices or the positive side of c . These coming from simple vertices have to meet the negative side of a vertex, so b has to lie in E_+ . Those edges coming from the positive side of c , have to meet d_+ or the negative side of a vertex. Meeting d_+ is the only possibility left, so d_+ has to lie in E_+ . Therefore we have Case 1. Note that in this case there are no more annuli in $R(\gamma)$. $\square$

Dually we have the following

3.13 *Claim.* Suppose that d_- is the only special vertex in D_- , then we have either Case 2 as in 3.12 or

Case 3. b is in D_+ , a_- and c are in E_- , a_+ and d_+ are in E_+ .

3.14 *Claim.* Suppose that b is the only special vertex in D_+ , then one of the following cases occur:

Case 3 as in 3.13.

Case 4. a_- and d_- are in D_- , c is in E_- , a_+ and d_+ are in E_+ .

Case 5. a_- and c are in D_- , d_- is in E_- , a_+ and d_+ are in E_+ .

Case 6. c is in D_- , a_- and d_- are in E_- , a_+ and d_+ are in E_+ .

Proof. Suppose first that a_- is in E_- . Then by 3.9 either c or d_- are in E_- . If c is in E_- , then d_- is the only special vertex in D_- , and by 3.6 a_+ is in E_+ , so we have Case 3. If d_- is in E_- , c is in D_- , and a_+ and d_+ are in the same component, then because there should be special vertices in E_+ it follows that a_+ and d_+ are in E_+ and so we have Case 6.

Now suppose that a_- is not in E_- . c or d_- are in E_- but not both, for a_- would be the only special vertex in D_- . If c is in E_- , then d_- is in D_- , and by 3.13 no case is possible if d_- is the only vertex in D_- , so a_- is in D_- and we have Case 4.

Subclaim. If d_- is in E_- then a_+ is in the area adjacent to E_- .

There are $2p$ edges crossing ∂D_+ , p of them come from simple vertices or the positive side of b and have to meet d_- . The remaining p edges cannot meet a special vertex in E_- , and so come from the negative side of b . Doing an argument similar to 2.6 (recalling ε_d) shows that for each edge adjacent to the negative side of b we can find a path in E_- which eventually has to cross

∂E_- and then meet the negative side of a vertex, that is, it has to meet a_+ . So a_+ is in the area adjacent to E_- .

If d_- is in E_- then c is in D_- , and by the subclaim a_+ and d_+ are in the same component, so E_+ has to be adjacent to E_- , for otherwise there will not be special vertices in E_+ . Then we have Case 5. $\square$

Dually we have the following

3.15 Claim. Suppose that c is the only special vertex in D_- , then one of the following cases occur:

Case 1 as in 3.12.

Case 6 as in 3.14.

Case 7. a_+ and d_+ are in D_+ , b is in E_+ , a_- and d_- are in E_- .

Case 8. b and d_+ are in D_+ , a_+ is in E_+ , a_- and d_- are in E_- .

In 3.12, 3.13, 3.14, 3.15 we have described all the cases in which only one special vertex is in an innermost disk. If an innermost disk contains two of $a_{\pm}, d_{\pm}$, then the other innermost disk contains only a single special vertex, so these cases are covered. Therefore the only remaining case is when b and c are in the innermost disks but are not the only special vertex there. By 3.11 it is not possible that b and a_+ are in the same innermost disk, so the only remaining case is the following:

Case 9. b and d_+ are in D_+ , a_- and c are in D_- , d_- is in E_- , a_+ is in E_+ .

3.16. When ∂M_n has only one boundary component the following are all the possible configurations of the vertices. Recall that this boundary component has two innermost sutures, one is in R_+ and one in R_- , denoted D_+ and D_- respectively, and E_- , E_+ are the annuli adjacent to D_+ and D_- respectively.

Case 1. a_+ is in D_+ , c is in D_- , a_- and d_- are in E_- , b and d_+ are in E_+ .

Case 2. a_+ is in D_+ , d_- is in D_- , a_- and c are in E_- , d_+ and b are in E_+ .

Case 3. b is in D_+ , d_- is in D_- , a_- and c are in E_- , a_+ and d_+ are in E_+ .

Case 4. b is in D_+ , a_- and d_- are in D_- , c is in E_- , a_+ and d_+ are in E_+ .

Case 5. b is in D_+ , a_- and c are in D_- , d_- is in E_- , a_+ and d_+ are in E_+ .

Case 6. b is in D_+ , c is in D_- , a_- and d_- are in E_- , a_+ and d_+ are in E_+ .

Case 7. c is in D_- , a_+ and d_+ are in D_+ , b is in E_+ , a_- and d_- are in E_- .

Case 8. c is in D_- , b and d_+ are in D_+ , a_+ is in E_+ , a_- and d_- are in E_- .

Case 9. b and d_+ are in D_+ , a_- and c are in D_- , d_- is in E_- , a_+ is in E_+ .

Note that in Cases 2, 4, 7, 9 there may be more annuli between E_- and E_+ , but in Cases 1, 3, 5, 6, 8 there are not more annuli. See Figure 9.

Suppose now that ∂M_n has more than one boundary component. Indeed, by 3.3 there are at most three, but we will shortly see that it has only two.

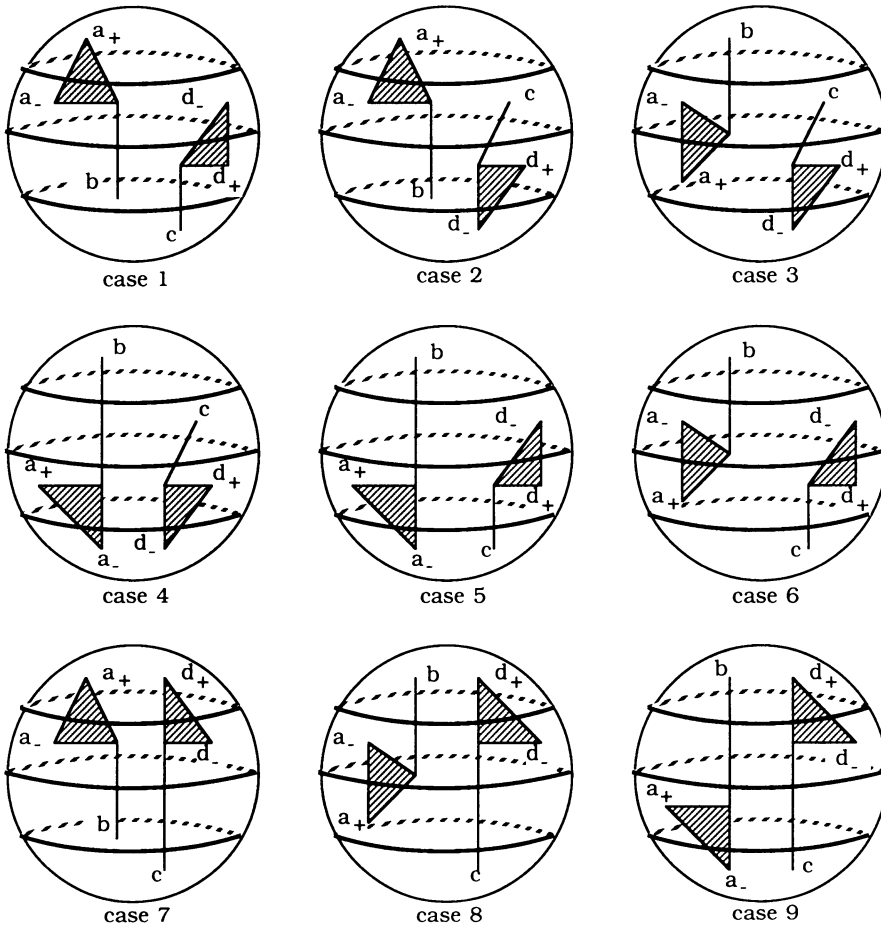


FIGURE 9

3.17 *Claim.* A component of ∂M_n meets both A_+ and A_- .

Proof. Suppose a boundary component T meets only A_+ . We have the following cases:

Case 1. T meets only a_+ and a_- .

Then T has only one suture γ , and two disks D_+, D_- . If T does not meet simple vertices, then there is a sphere parallel to T meeting β_α in one point, showing that the band is trivial. If D_+ meets simple vertices, then there are p edges coming from simple vertices which cross ∂D_+ and they have to meet the negative side of a vertex in D_- , but there is no such vertex. If D_- contains simple vertices and D_+ does not, then an argument similar to that in 3.7 (Claim) shows that either there is a Scharlemann cycle in D_- or a 1^* -path joining a_- and a_+ , which consists of level edges and bounds, together with ε_a , a disk with interior disjoint from Γ ; as in 3.7 this yields a contradiction.

Case 2. T meets only b .

This contradicts 3.3 (note that by 1.8 T has sutures).

Case 3. T meets a_+, a_- , and b .

If a_+ and b are in the same innermost disk, then because T does not meet

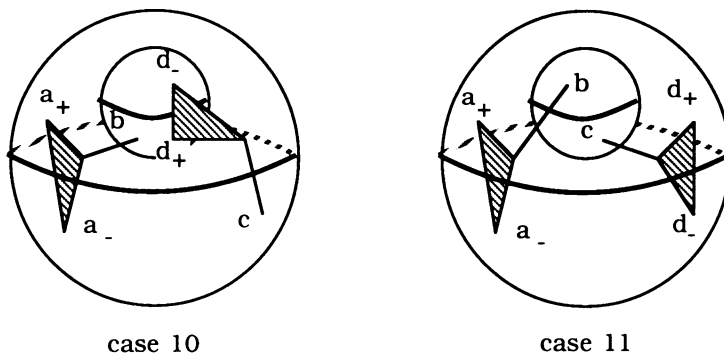


FIGURE 10

A_- , an argument similar to 3.11 yields a contradiction. If b is in an innermost disk and a_+ is in the other, then there is an annulus between these disks, which only contains a_- , which contradicts 3.9. If b is an innermost disk and a_- is in the other, then there would not be special vertices in the area adjacent to b , which contradicts 3.3. $\square$

3.18. Note that 3.17 implies that ∂M_n has only two components, denote them by T and T' . Applying 3.17 it is not difficult to see that we have only the following two cases. See Figure 10.

Case 10. T meets a_+ , a_- and c , and T' meets d_+ , d_- and b . By 3.6 and 3.7 each sphere contains only one suture; a_- and c are contained in a disk of T , and a_+ in the other; d_+ and b are in a disk of T' , and d_- in the other.

Case 11. T meets a_+ , a_- , d_+ , d_- and T' meets b and c . By 3.3, 3.6 and 3.7 each sphere contains only one suture; a disk of T contains a_+ and d_+ , and the other contains a_- and d_- ; a disk of T' contains b and c is in the other.

4. FINDING A SET OF PARALLEL EDGES

We will concentrate on Case 4 of 3.16, because it is representative, and the proofs for the remaining cases are similar.

4.1 *Claim.* If Case 4 of 3.16 happens then the orientation of β is as in the Figure 3(a).

Proof. 2.5 implies that there are at least $2p$ edges crossing the suture ∂D_+ . p of them come from the positive side of b or from simple vertices and have to meet c in its negative side, for in the negative side of c there is room only for p edges. Then $2p$ edges across ∂D_+ , p comes from simple vertices or the positive side of b , and p from the negative side of b . This implies that no edge in D_- crosses ∂D_- (other than $\varepsilon_a, \varepsilon_d$), for otherwise there would be a component of Q_n with positive index. If β has the orientation of Figure 3(b), d_- has labels $1, 2, \dots, p$. d_- is antiparallel to all vertices in D_- . Relabel d_- by $1^*, 2^*, \dots, p^*$, so that it comports like a vertex parallel to all vertices in D_- . Note that the parity rule still holds under the new labelling. Apply 2.8 to this situation, considering a_- and d_- as the vertex in the boundary. So 2.8 implies that there is a Scharlemann cycle in D_- , which is a contradiction.

Then the orientation of β is as in Figure 3(a), and the negative side of c is in the bottom of β . $\square$

4.2 *Claim.* All the edges adjacent to the positive side of b , the negative side of c , a_- and d_- are level.

Proof. As in the proof of 4.1, $2p$ edges cross ∂D_+ , p come from simple vertices or the positive side of b and meet the negative side of c , and p come from the negative side of b . So no edge crosses ∂D_- (other than $\varepsilon_a, \varepsilon_d$). As β has the orientation of Figure 3(a), d_- has labels $1^*, 2^*, \dots, p^*$. Relabel d_- by $1, 2, \dots, p$, so that it comports like a vertex parallel to all the vertices in D_- . Apply 2.7 to this situation, considering a_- and d_- as the vertex in the boundary. Then 2.7 shows that the edges adjacent to a_- and d_- are level. This implies that the edges adjacent to the negative side of c and to the positive side of b are level. $\square$

4.3. $I(Q) = I(Q_n) = 0$, so by 1.17 we can assume that $Q_1 = Q - \text{int } \eta(S)$, $Q_i = Q_{i-1} - \text{int } \eta(S_i)$, for all Q_i in the hierarchy, and then $\partial Q_n = (Q \cap \partial M_n) \cup (Q \cap \beta_n)$. Also $Q \cap T = \partial Q_n \cap T$.

Look at the graph Λ in the planar surface Q_α formed by its intersection with the sphere $T \subset \partial M_n$. A vertex of Λ is a circle of intersection between β and Q_α , and the edges are the arcs of intersection between Q_α and T . As in 2.1 denote the vertices of Q_α by $a_1, a_2, \dots, a_p$. We say that a loop η in Λ is bad if the two circles of intersection of Q_α with ∂M are in distinct components of $Q_\alpha - \eta$, otherwise it is good. Note that a level edge in Γ correspond to a loop in Λ . Choose one of the boundary components of Q_α which lie in ∂M , denote it by Y ; define the interior of a loop in Q_α as the component of its complement which does not contain Y . Label the vertices of T with $1, 2, \dots, n$; 1 corresponds to $a_\pm$, 2 to b , $n-1$ to c , and n to $d_\pm$, label the other vertices in the obvious way. Label the end of an edge in a_i with j (j^*) if this point corresponds to an intersection between a_i and the vertex in T labeled with j , and it lies in the top (bottom) of β .

4.4 *Claim.* The level edges adjacent to the positive side of b , the negative side of c , a_- , and d_- are bad loops in Λ .

Proof. We distinguish two cases.

Case 1. There are arcs of β running from D_+ to D_- . So there are more vertices in Γ .

In this case there are $2p$ edges crossing ∂D_+ , p coming from the negative side of b and p coming from simple vertices.

Step 1. The level edges adjacent to the positive side of b correspond to bad loops in Λ .

If one of them is a good loop, then there is an innermost one; let η be this loop. η in Γ has an end in the positive side of b and the other end in the top part of a simple vertex, its ends in Γ are labeled with i^*-i , say. η in Λ is a loop based at a_i . η in Λ looks as in Figure 11. Now because η is not adjacent to $a_\pm, d_\pm$, the Seifert surface S has to meet Q_α in an arc σ as shown in Figure 11, σ has one end in β_0 , say, and the other in β_α . If the other end of σ is in the bottom part of β , this would contradict that the orientation of β points in the direction of the normal vector to S ; if this end of σ is in the top part of β , then σ will cut a disk q_1 , with $I(q_1) = -2$, which will be

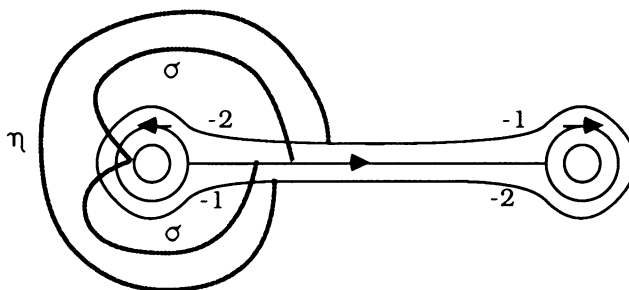


FIGURE 11

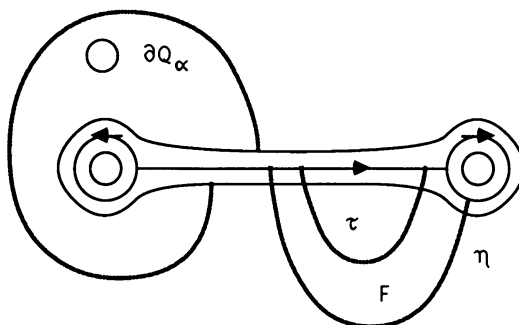


FIGURE 12

a parameterizing surface in its own right in M_1 , and then $I(q_n) \leq -2$, which contradicts 1.15 (an alternate argument: Note that in this situation there is an isotopy which moves the Seifert surface intersecting β in fewer points, which contradicts the β -tautness of S).

Note that the loops corresponding to the level edges in a_- are parallel to these corresponding to b , so they have to be bad loops (this is because there is a disk q of Q_n which made them parallel).

Step 2. The level edges ending in the negative side of c are bad loops in Λ .

Each of these edges join c and a simple vertex, with both ends in the bottom part of these vertices. Suppose one of these edges corresponds to a good loop τ in Λ ; let η be the loop parallel to τ which corresponds to a level edge adjacent to d_- , so η is a good loop. We have something similar in Figure 12. Let F be the disk in Q_α bounded by η . τ is inside F and there are no vertices inside F because each vertex in Λ has a bad loop by Step 1. Note that η in T crosses no suture. Let S_i be the first surface in the hierarchy which meets the interior of F . Then either S_i meets F in a closed curve, which will form a disk of Q_i with boundary contained in $R(\gamma)$, contradicting that Q_i is a parameterizing surface, or S_i meets F in a collection of arcs, each with ends in the bottom part of β_α . But the two intersections of β and S_i corresponding to the ends of one of these arcs have a different sign, which contradicts the β -tautness of S_i . This implies that the loops in Λ corresponding to the level edges adjacent to the negative side of c and d_- are bad loops.

Case 2. There is no simple vertex in D_+ , D_- .

Then the positive side of b and the negative of c are joined up by p level edges, and so are a_- and d_- . If one of these level edges corresponds to a good

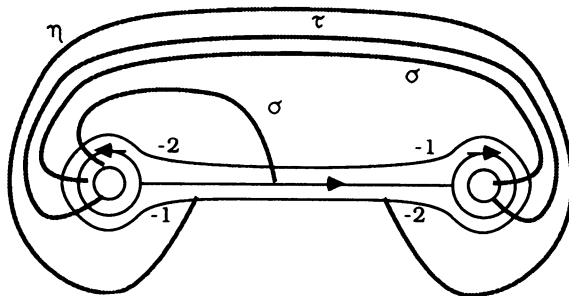


FIGURE 13

loop in Λ , then there is an innermost one. Let η be a level edge joining b and c which corresponds to an innermost good loop in Λ ; let τ be the parallel loop corresponding to an edge joining a_- and d_- , and note that τ crosses no suture. If η is inside τ then proceed as in Step 2 above. So suppose τ is inside η , we have something similar in Figure 13. Look how the surface S meets Q_α in the vertex in question; it meets Q_α in one arc σ , with one end in β_0 , and the other end either in β_α or β_1 . In any case S cut a disk q_1 with $I(q_1) = -2$ which is a parameterizing surface in its own right in M_1 , therefore q_n has negative index, which contradicts 1.15. $\square$

4.5 Remark. If $p > 1$, there are two possibilities of how the negative side of c joins the vertices of D_+ . Let η_1 (η_p) = (edge adjacent to c at label 1^* (p^*)) $\cap E_-$. Let α be the arc in c , which go through the negative side of c , and which joins the labels 1^* and p^* . $\eta_1 \cup \eta_p \cup \alpha \cup (\text{arc in } \partial D_+ \text{ joining } \eta_1 \text{ and } \eta_p)$ bounds a disk $F \subset E_-$. If the positive side of c is in the interior of F , so are the edges coming from the negative side of b . Doing an argument as in 4.2 we can show that all the edges adjacent to the positive side of c and negative side of b are level, and doing an argument as in 4.4 they correspond to bad loops in Q_α . In this case 4.6 and 4.7 are easier. If the positive side of c is not in F , all the edges adjacent to the negative side of c lie in F , and no other vertex or edge is in F . In this case we cannot assure that the edges adjacent to the negative side of b are level. This made the complications of 4.6. Also an edge adjacent to the negative side of b may cross two sutures and meet a_+ . An edge adjacent to the positive side of c may cross a suture and meet d_+ .

4.6 Claim. Suppose that there are simple vertices in D_+, D_- . Then there is a vertex $v = a_x$ in Q_α so that all the edges adjacent to v are loops, except possibly those with labels $1, 2, 3, 4$ (or $(n-2)^*, (n-3)^*$), and all the loops are bad loops, except possibly one with labels $n, n-1$ or $1, 2$, which correspond in Γ to an edge joining the positive side of c and d_+ , or an edge joining the negative side of b and a_+ .

Proof. 4.4 says that in Λ each vertex has bad loops with one end labeled $1^*, 2^*, (n-1)^*, n^*$. Take an innermost vertex $v = a_x$ in Λ , so in one of the sides determined by the loops at v there are no more vertices. Let η_i be the loop at v with one label i^* , for $i = 1, 2, n-1, n$. Then either η_1 or η_{n-1} contains the other three loops in its interior, suppose first that η_1 is the outermost of these loops. The other end of η_1 (η_2) is labeled j ($j+1$), $j < n-2$.

The other end of η_{n-1} (η_n) is labeled $(k+1)^*$ (k^*). If $k+2 \neq n-1$ then there is a loop at v with labels s^* , t^* , $k+2 < s < t < n-1$ but this implies that in T there are simple vertices joined by an edge with labels $x^* - x^*$; this edge cannot cross a suture, for no edge joining simple vertices crosses a suture, so the two vertices are parallel, which contradicts the parity rule. So $k+2 = n-1$, and η_{n-1} is an innermost loop.

Suppose first that there is a good loop in v , with labels n , $n-1$ (note that this is the only possible good loop in the interior of η_1 , any other good loop would contradict the parity rule, as above; this loop corresponds to a level edge with labels x joining c and d_+). So because $k+2 = n-1$ the number of labels between the labels 2^* and k^* , which is $n-6$, has to be equal to the number of labels between $j+1$ and $n-1$, and this forces that $j=3$, i.e. all the edges incident to v are loops, except perhaps those incident to the labels $1, 2$.

Suppose now that there is no good loop at v . Then incident to the labels n , $n-1$ are two bad loops, for these labels are the interior of η_1 . The number of labels between the labels 2^* and k^* , which is $n-6$, has to be equal to the number of labels between $j+1$ and n^* , and this forces that $j=5$, and then all the edges incident to v are loops, except perhaps the edges with labels $1, 2, 3, 4$.

If η_{n-1} is the outermost loop then a similar argument shows that all the edges adjacent to v are loops except perhaps those with labels $(n-2)^*$, $(n-3)^*$. The only possible good loops have ends labeled $1, 2$ or $n, n-1$. This completes the proof. $\square$

4.7 Claim. In Γ there are two vertices joined by p parallel level edges, one for each label, which cross no suture, and these edges correspond to bad loops in Λ .

Proof. Suppose first that there are simple vertices in D_+ , D_- . Note that in T two vertices labelled $i, i+1$ are antiparallel, so only one of $3, 4$ (or $n-2, n-3$) can be at D_+ . 4.6 implies that there is a label $x \in \{1, \dots, p\}$, so that all the edges in D_+ incident to a label x or to a label x^* have to be level, except possibly one, i.e. the incident to the vertex $3, 4, n-2$ or $n-3$. But in D_+ there is the same number of labels x and x^* , and there is exactly one edge with label x (the adjacent to b) and one with label x^* (joining a simple vertex and c) in E_+ which cross ∂D_+ , which forces that all the edges adjacent to a label x^* or x are level. Take a x^* -path γ in D_+ beginning at b . It is formed by level edges and ends at c , but it must meet all the vertices in D_+ , for otherwise there would be a x^* -cycle formed with the vertices missed by the path. Say γ meets the vertices in the order $v_0 = b, v_1, v_2, \dots, v_m = c$. $\gamma \cup$ (edge incident to b at label x) separates D_+ in two parts. Take now an $(x-1)^*$ -path λ (or a $(x+1)^*$) beginning at b ; all its edges lie in one side of γ . λ eventually crosses ∂D_+ and meets the negative side of c . If λ does not meet the vertices in D_+ in the same order as γ , then there is a first vertex at which they disagree, say λ meets the vertices in the order $v_0, v_1, \dots, v_k, v_s, s \neq k+1$. If $s < k$, then λ is a i^* -cycle; if $s > k$, then the part of γ from v_k to v_s and the edge of λ joining v_k and v_s enclose an area which has $P(x-1)$, so there would be an $(x-1)$ -cycle, and then a Scharlemann cycle (cf. Lemma 2.3). Then λ meets the vertices in the same order as γ does, so it consists of level edges. Repeat

the same process for $(x - 2)^*$, etc. This implies that all the edges in D_+ are level and that two consecutive vertices in the path γ are joined by p parallel level edges. In particular this shows that the positive side of b is joined to a simple vertex by p level edges. By 4.4, Case 1, these edges correspond to bad loops in Λ .

If there is no simple vertex in D_+ , D_- , then a_- and d_- are joined by p level edges which cross no suture, and by 4.4, Case 2, these edges correspond to bad loops in Λ . $\square$

4.8. Now we outline a proof of 4.6 for the remaining cases. Cases 5, 7, 8 are similar to Case 4. In these cases b or c is the only special vertex in D_+ , there are $2p$ edges which cross the suture $\partial D_{\pm}$, p come from the negative side of b or c , p come from simple vertices or the positive side of b or c and meet a special vertex, and in $E_{\mp}$ there is only one special vertex. An argument similar to 4.1 shows that Case 7 is possible only when β is oriented as in Figure 3(a), and Cases 5, 8 are possible only when β is oriented as in Figure 3(b). Now we proceed with 4.2, 4.4, 4.6, 4.7, as before, interchanging the roles of b and c when necessary.

In cases 1, 2, 3, a_+ or d_- is the only special vertex in $D_{\pm}$, so 3.6 implies that $D_{\pm}$ has simple vertices and no edge adjacent to a_+ or d_- crosses the suture $\partial D_{\pm}$, so there are p edges which cross $\partial D_{\pm}$, all coming from simple vertices and have to meet a special vertex in $E_{\pm}$. This implies that in Case 1(3) there are $2p$ edges which cross ∂D_- (∂D_+), p come from simple vertices and meet the negative side of b (c), p come from the positive side of c (b) and meet d_+ (a_-). Also in Case 1(3) the p edges adjacent to the positive side of b (c) have to cross a suture and join a_- (d_+). An argument as in 4.1 shows that if β is oriented as in Figure 3(a) then we only have Case 2; and if β is oriented as in Figure 3(b) we have Cases 1, 3. Now proceed with 4.2, 4.4, 4.6, 4.7 as before.

Now consider Case 6. If there are no simple vertices in D_+ (D_-), then the $2p$ edges adjacent to b (c) cross ∂D_+ (∂D_-); the p edges coming from the positive side of b (c) meet a_- (d_+), by index restrictions. Then there is a disk D , so that $\partial D = \alpha \cup \lambda$, α is an arc in ∂T joining b and a_- and λ is an arc in $\eta(\beta)$ going from b to a_- . An argument similar to 3.6, Claim 1, shows that E_- is not β -taut, which is a contradiction. Therefore there are simple vertices in D_+ (D_-). There are $2p$ edges which cross ∂D_+ (∂D_-). If more than p of these edges come from b (c), then an argument similar to 3.6, Claim 2, shows that there are more than p edges coming from simple vertices which cross the suture ∂D_+ (∂D_-), but in d_- (a_+) there is only room for p of them. Then p edges come from simple vertices and meet d_- (a_+), and p come from b (c). If an edge α joins a simple vertex in D_+ (D_-) and d_- (a_+), then α is part of a index zero disk q , whose boundary consist of α , an arc on the negative side of A_- (A_+), an edge in D_- (D_+) joining the negative side of c (b) and a simple vertex and which crosses no suture, and another arc on $\eta(\beta)$. This implies that the p edges adjacent to the negative side of c (b) do not cross the suture ∂D_- (∂D_+). Hence the p edges adjacent to the positive side of b (c) cross ∂D_+ (∂D_-), and they have to meet a_- (d_+). An argument similar to 4.1 shows that Case 6 is possible only when β is oriented as in Figure 3(a). Now we follow with 4.2, 4.4, 4.6, 4.7.

In Case 9, if an edge adjacent to d_+ (a_-) crosses ∂D_+ (∂D_-), then it meets d_- (a_+), which implies that there is a loop at c (b), so no edge adjacent to d_+ (a_-) crosses ∂D_+ (∂D_-). Take an i -path or i^* -path λ_i starting at d_+ (a_-) for each i . This path has to cross ∂D_+ (∂D_-), and if the edge crossing ∂D_+ (∂D_-) is incident in D_+ (D_-) to a simple vertex or the positive side of b (c), then that path will finish at d_- (a_+). Doing an argument as in 3.7 (Claim), λ_1 or λ_p starts at d_+ and finishes at d_- , and consists of level edges. $\lambda_1 \cap \varepsilon_d$ bounds a disk D with interior disjoint from Γ , as in 3.7 this is not possible. Then the edge of the path λ_i crossing ∂D_+ (∂D_-) is incident to the negative side of b (c). Then the p edges adjacent to the negative side of b (c) cross ∂D_+ (∂D_-). In particular this implies that no edge adjacent to the positive side of b (c) crosses ∂D_+ (∂D_-). This also implies that no edge adjacent to a simple vertex in D_+ (D_-) can cross ∂D_+ (∂D_-), by index restrictions. Then the only edges which cross ∂D_+ (∂D_-) are those adjacent to the negative side of b (c). Doing something like in 4.1 it can be seen that this case is only possible when β is oriented as in Figure 3(a). Now follow with 4.2, 4.4, 4.6, 4.7 as before.

Case 10 is similar to Case 9, and Case 11 is similar to Case 4. These cases are only possible when β is oriented as in Figure 3(a).

5. CONSTRUCTING A SWALLOW-FOLLOW TORUS

Recall that P denotes the decomposing sphere from which Q_α was obtained. It was proved in 4.7 that there are two vertices in T ($= \partial M_n$) which are joined by p consecutive parallel level edges, one for each label i , $i = 1, \dots, p$, and these edges correspond to bad loops in Q_α . Call these vertices v_1 and v_2 . In what follows we do not distinguish between the labels i and i^* .

Denote by e_i the edge joining v_1 and v_2 with label i , for $i = 1, \dots, p$; and denote by α_i (η_i) the arc inn v_1 (v_2) joining the labels i and $i+1$, for $i = 1, \dots, p-1$. Let C_i be the disk in T bounded by α_i , η_i , e_i , and e_{i+1} . The arc e_i in Q_α (or in P) is a loop based at a_i . Choose for each a_i an arc γ_i in a_i having the same ends as e_i , and so that α_i , γ_i , η_i , γ_{i+1} bound a disk in $\partial\eta(\beta)$. Let H_i be the disk in $\partial\eta(\beta)$ bounded by α_i , γ_i , η_i , γ_{i+1} . Then the interior of H_i and C_i do not meet P . Let A_i be the annulus obtained by gluing C_i and H_i along $\alpha_i \cup \eta_i$; the interior of A_i meets neither P nor int $\eta(\beta)$. Let F_i be the annulus in P bounded by e_i , γ_i , e_{i+1} , γ_{i+1} ; note that this annulus may intersect $\eta(\beta)$. It follows that $\partial A_i = \partial F_i$. See Figure 14. Note that there are two choices for γ_i , which produce two choices for H_i , A_i and F_i ; denote by γ'_i , H'_i , A'_i , F'_i the alternative choice.

Note that A_i and A_{i+1} have a component of its boundary in common, so A_i and A_{i+1} lie in different side of P in S^3 . It is not difficult to see that there is a F_j which is innermost, i.e. its interior misses $\eta(\beta)$ or meets it in two disks, and in the last case the interior of F'_j misses $\eta(\beta)$. Do 2-surgery in P with A_j ; then we obtain a sphere P' intersecting k_b in two points and a torus R disjoint from k_b . P' and R are isotopic to $(P - F_j) \cup A_j$, $F_j \cup A_j$, respectively. If P' is a decomposing sphere for k_b , and F_j meets $\eta(\beta)$ in two disks, then P' has fewer intersections with $\eta(\beta)$ than P does, contradicting the minimality of the intersections between P and $\eta(\beta)$; if the interior of F_j misses $\eta(\beta)$, then using A'_j instead of A_j , we get a decomposing sphere P''

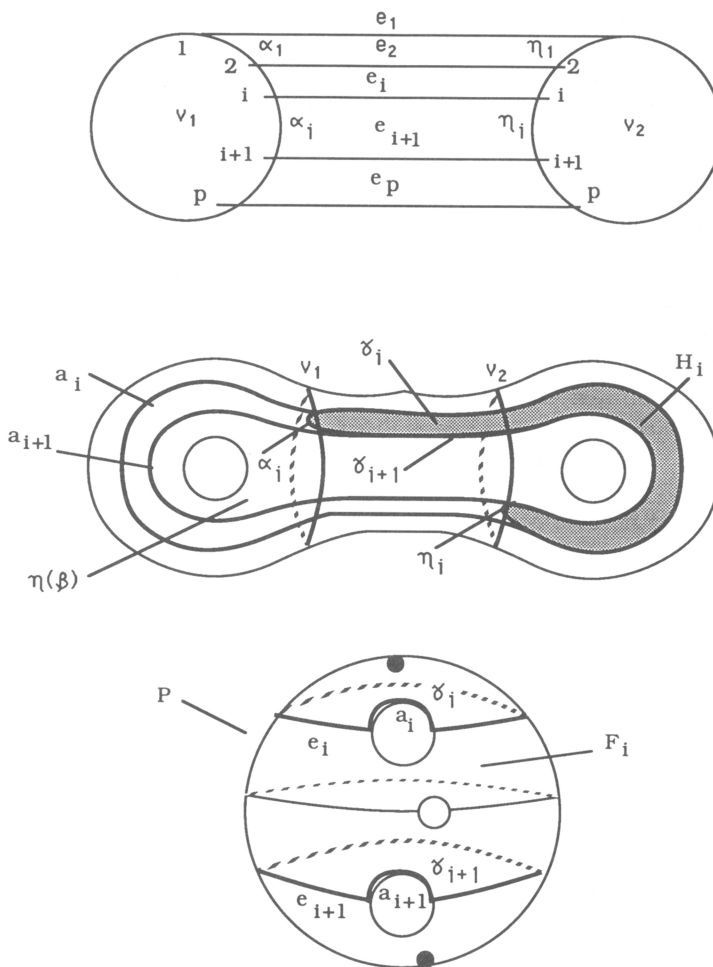


FIGURE 14

(in fact isotopic to P'), having fewer intersections with $\eta(\beta)$ than P does. So P' is not a decomposing sphere and the annulus A_i runs parallel to one arc of one of the summands of k_b . Then R is a swallow-follow torus for k_b . If the interior of F_j misses $\eta(\beta)$ then clearly R is disjoint from $\eta(\beta)$. If F_j meets $\eta(\beta)$ in two disks, then using A'_j instead of A_j , we get a swallow-follow torus R' (isotopic to R) which is disjoint from $\eta(\beta)$.

This construction is possible only if $p > 1$; so we conclude that either $p = 1$, or there is a swallow-follow torus disjoint from the band.

This completes the proof of Theorem 1.

6. BAND SUMS OF SPLIT LINKS WHICH YIELD COMPOSITE KNOTS

Theorem 2. *Let k_b be a band sum of k_0 and k_1 , for $k_0 \cup k_1$ a split link. If k_b is composite, then there is a decomposing sphere disjoint from the band or there is a decomposing sphere crossing the band in one arc parallel to $b(\{1/2\} \times I)$.*

Proof. Clearly $S^3 - (k_0 \cup k_1 \cup b(I \times I))$ is irreducible. Now apply Theorem 1, as stated in the introduction. (d) of Theorem 1 cannot happen because $S^3 - (k_0 \cup k_1)$ is reducible. If (a) or (b) happen then we are finished, so suppose (c) happens. Let T be a swallow-follow torus disjoint from the band. T is a torus in the complement of $k_0 \cup k_1$. T bounds a solid torus T' , k_b lies in T' . The winding number of k in T' is 1, so the algebraic winding number of $k_0 \cup k_1$ is also 1; so at least one of the components of this link is not trivial inside T .

Let S be a splitting sphere for $k_0 \cup k_1$. Recall that $b : I \times I \rightarrow S^3$ is a map such that $b^{-1}\{k_i\} = I \times \{i\}$, $i = 0, 1$. Isotope S so that intersects the band in a collection of arcs parallel to $I \times \{1/2\}$; take a splitting sphere S which minimizes this number of arcs.

Claim. T and S can be chosen so that they do not intersect.

Proof of claim. Suppose they intersect in a collection of circles, take one innermost circle α in S , which is boundary of a disk D . If D lies outside of T' , then because T is an essential torus, α is a trivial curve on T , which bounds a disk in T , and because $S^3 - T'$ is irreducible we can isotope D to remove the circle α . If D lies inside T' , then because the winding number of $k_0 \cup k_1$ in T' is 1, D cannot be a meridian of T' , so α is trivial in T and is boundary of a disk $D' \subset T$. Take a curve of intersection between S and D' which is innermost on D' ; such a curve is a boundary of a disk $D'' \subset D'$ whose interior is disjoint from S . By doing surgery in S with D'' , we get two spheres, at least one of them is a splitting sphere which has fewer intersections with T , and it has equal or fewer intersections with the band than S . $\square$

As T is a swallow-follow torus for k_b , there is a meridian disk of T' which intersects k_b in one point, call it Q . By isotoping Q , suppose that it meets the band in a collection of arcs parallel to $b(\{1/2\} \times I)$; choose Q so that it minimizes this number of arcs. Now follow the combinatorial route of [BE], the situation in here is very similar to the situation in there, because the disk Q , which intersects the knot in one point is like a decomposing sphere for a knot; note also that S does not meet ∂Q . So we consider graphs Γ_S and Γ_Q , in S and Q respectively, whose fat vertices are the disks of intersection between a regular neighborhood of $b(I \times I)$ and S and Q respectively, and their edges are the intersection arcs between $S - \text{int } \eta(b(I \times I))$ and $Q - \text{int } \eta(b(I \times I))$. The whole combinatorial argument is now similar to [BE]. The strategy of the proof is as follows: Use [BE, §4] (which is essentially 2.3, i.e. [CGLS, 2.6.2]) to find a pure level circuit or a (Scharlemann) cycle λ in Γ_S . If λ is a pure level circuit in Γ_S , then it corresponds in Γ_Q to a collection of loops; if one of these loops is good then we do a path threading argument, as in [Sc₁, §6, 7, BE, 2.8]. If all these loops are bad, then in [BE, §3] λ is used to construct two spheres which meet a composite knot k in two points, and one of them had to be a decomposing sphere unless k had a two-bridge knot summand. Here the same construction yields a sphere intersecting k_b in two points, and a disk with the same boundary as Q , intersecting k_b in one point, so it is a meridian disk of T' . This new disk has fewer intersections with the band than Q , contradicting its minimality. If λ is a cycle, we first do 1-surgery on Q to get a punctured torus, and then do 2-surgery with the disk in S determined by λ , to get a new meridian disk having fewer intersections with the band (see [Sc₁, 4.7, Sc₂, 5.6] for details of this construction).

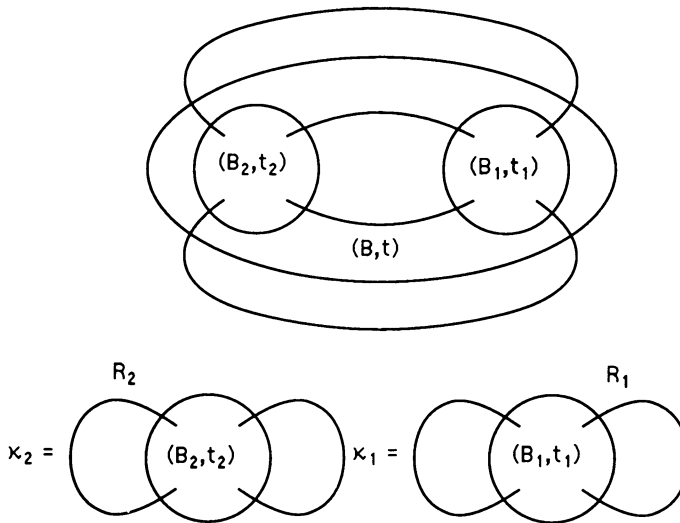


FIGURE 15

This argument implies that the band does not meet Q . Doing 2-surgery on T with Q , we get a sphere disjoint from the band which intersects k_b in two points. This is a decomposing sphere for k_b , since k_b is not a core of T , so we have (a) of Theorem 1. $\square$

Remark. Actually we are proving a little more: Suppose k is any split link, and b is a band such that $S^3 - k \cup b(I \times I)$ is irreducible. If the banding k_b is composite then the same conclusions as in Theorem 2 happen.

We refer to [L] or [E₂] for definitions and facts about tangles not found here.

Corollary 1. *Let k_b, k_0, k_1 be as in Theorem 2. If k_b is composite and there is no decomposing sphere disjoint from the band, then k_b has two prime summands, and one is a two-bridge knot.*

Proof. Let N be a regular neighborhood of the band. It intersects $k_0 \cup k_1$ in two arcs. Its complement can be seen as a locally prime tangle (B, t) , i.e. it does not have local summands for no decomposing sphere is disjoint from the band. Then k_b and $k_0 \cup k_1$ can be seen as links obtained by summing a rational tangle and (B, t) . By Theorem 2 there is a decomposing sphere P intersecting the band in one arc; then $P \cap N$ is a single disk, and $P \cap B$ is disk properly embedded in (B, t) intersecting the strings of (B, t) in two points. This shows that (B, t) is a partial sum of two tangles, say (B_1, t_1) and (B_2, t_2) (see Figure 15). Note that $k_0 \cup k_1$ can be expressed as a tangle sum of (B_1, t_1) and (B_2, t_2) . If both tangles are prime, Theorem 1 in [L] implies that $k_0 \cup k_1$ is a prime link which is impossible, so one of them, say (B_1, t_1) is a rational tangle. Let (N, s) be the rational tangle obtained by the intersection of k_b and N . k_b can be expressed as $k_b = (B, t) + (N, s) = \kappa_1 \# \kappa_2$, where $\kappa_1 = (B_1, t_1) + R_1$, and $\kappa_2 = (B_2, t_2) + R_2$, where R_1, R_2 are rational tangles (see Figure 15). κ_1 is a two-bridge knot because it is the sum of two rational tangles. If (B_2, t_2) is also a rational tangle then κ_2 is a two-bridge knot and then it is prime. If (B_2, t_2) is a prime tangle, then note that $(B_2, t_2) + (B_1, t_1) = k_0 \cup k_1$ is a split link and $d((B_1, t_1); R_2) \geq 2$ (i.e. the boundaries of the central disks of (B_1, t_1) and

R_2 , which lie in ∂B_2 , meet in at least four points, see [E₂], for otherwise κ_1 would be a trivial knot. Then Theorem 2 in [E₂] implies that $\kappa_2 = (B_2, t_2) + R_2$ is a prime knot. $\square$

It is proved in [BE] that if k_b is a composite knot and k_0 and k_1 are trivial, then one of the summands of k_b is a two-bridge knot. (In fact, we do not need the fact that k_0 and k_1 are trivial knots, but only that no decomposing sphere is disjoint from the band.)

7. BANDINGS OF TRIVIAL KNOTS WHICH YIELD COMPOSITE LINKS

Theorem 3. *Suppose k is the trivial knot and k_b is a banding of k . If k_b is a composite knot or link, then there is a decomposing sphere which crosses the band in one arc parallel to $b(\{1/2\} \times I)$.*

7.1 Proof. Apply Theorem 1, as stated in the introduction. (a) cannot happen because this would imply that k has a nontrivial summand. (c) does not happen for there is no essential torus in the exterior of k . So either (b) or (d) happens. If (b) holds we are finished, so suppose we have (d).

7.2. So assume that:

There is a disk Q' with $\partial Q' = k$, so that it has a minimal number of intersections with the band among all such disks, and such that the band intersects it always in the same direction.

Remark. Let (M, γ, β) as in 1.2. Note that the sutured manifold decomposition $(M, \gamma, \beta) \xrightarrow{Q'} (M_1, \gamma_1, \beta_1)$ is β -taut, ∂M_1 is a sphere with only one suture, and then (M_1, γ_1) is $\emptyset$ -taut; i.e. the sutured manifold hierarchy as constructed in 1.4 consist of only one step.

7.3. Let P' be a decomposing sphere for k_b . In what follows we want to prove that either the band meets P' in one arc, or that 7.2 is impossible. Let N be a regular neighborhood of $k \cup b(I \times I)$. The last is homeomorphic to a 3-ball with two 1-handles. By general position P' can be isotoped so that $P' \cap b(I \times I) = b(\{a_i\} \times I)$ for some finite $\{a_i\} \subset \text{int}(I)$. Similarly, by general position, Q' can be isotoped so that $Q' \cap b(I \times I) = b(I \times \partial I) \cup b(I \times \{b_s\})$, where $\{b_s\}$ is some finite subset of I . Suppose P' has a minimal number of intersections arcs with the band. Let $P = P' - \text{int}(N)$, $Q = Q' - \text{int}(N)$; P, Q are connected planar surfaces in the complement of N . ∂P consist of n curves $a_1, \dots, a_n$, plus two curves a_+, a_- , which are meridians of the 1-handles, as in Figure 16. ∂Q consist of m curves $b_1, \dots, b_m$, plus one curve b_0 which goes once over the 1-handles; we have a situation as in Figure 16. See [Sc₁, Sc₂] for more details over N and the positions of the boundary curves of P and Q .

Give an orientation to P and Q . This induces an orientation in their boundaries. Two curves a_i and a_j are parallel if with the induced orientation they are parallel in ∂N , the same for the b_i 's. By 7.2 we may assume that all the b_i 's, $1 \leq i \leq m$, are parallel. b_0 also has an induced orientation. a_i and b_j meet in two points, one in the front of N and one in the back. Label the point with i in Q , and with j in P , if this point is in the front of N , and label it with i^* in Q and j^* in P if it is in the back of N , for $1 \leq i \leq n$, $1 \leq j \leq m$. Label the point of intersection between b_0 and a_i with i in Q and with 0

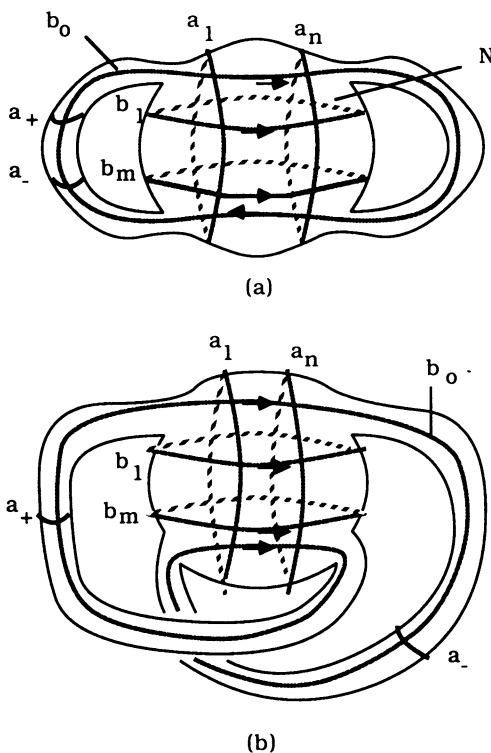


FIGURE 16

in P if it lies in the top of N , and label it with i in Q and $m+1$ in P if it lies in the bottom of N . Denote the points of intersection between b_0 and a_+ , a_- with $+$, $-$ in both P and Q . See Figure 17. Consider the graphs of intersection between P and Q like in [Sc₁], denote them by Γ_P , Γ_Q . Define loop, good loop, bad loop, circuit, cycle, i -path, Scharlemann cycle, level edge, etc., as in [Sc₁] or as in 2.2, 4.2. Define the interior of a circuit or path going from b_0 to b_0 in Γ_Q as the component of its complement which does not contain the point x on the east of b_0 as in the Figure 17.

7.4. We have two cases, that is, k_b is either a knot or a link of two components.

Case 1. k_b is a link.

In this case N appears as in Figure 16(a). a_+ , a_- lie in the same 1-handle, suppose they lie in the west side of N , as in Figure 16(a). The top (bottom) part of b_0 with the induced orientation is either parallel to the front part or to the back part of the b_i 's. Suppose with no loss of generality that the top part of b_0 is parallel to the front part of the b_i 's and the bottom part of b_0 is parallel to the back part of the b_i 's. Relabel a point of intersection between b_0 and a_i with i in Q and with 0 in P if this point lies in the top of N , and label it with i^* in Q and with $(m+1)^*$ in P if it lies in the bottom of N . It is not difficult to see that Γ_Q appears as in Figure 18.

7.5. Here we have the same parity rule as in 2.1. Note that in Γ_Q there is no edge with labels $i-i$ (i^*-i^*), i.e. a level edge has to be labeled $i-i^*$. Note that the top part of b_0 has labels $1, 2, \dots, n$, and the bottom part has labels $1^*, 2^*, \dots, n^*$. This implies that a loop in Γ_P has ends labeled $i-j^*$.

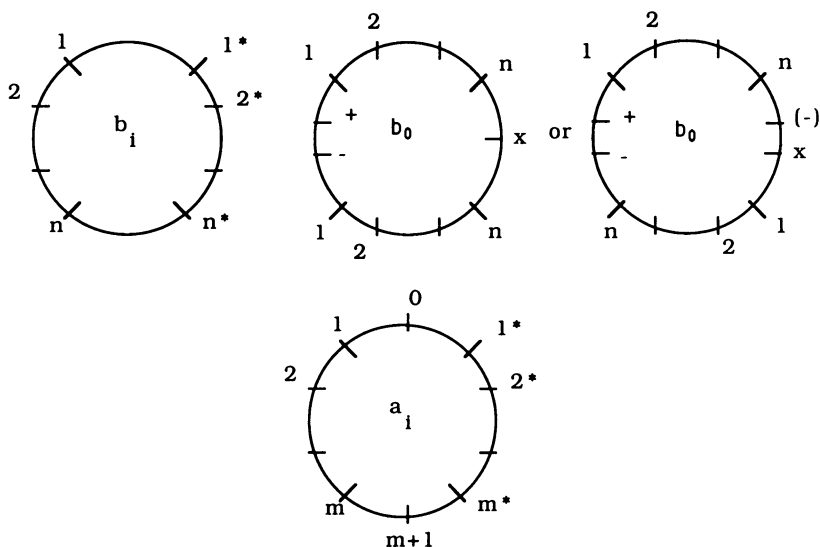


FIGURE 17

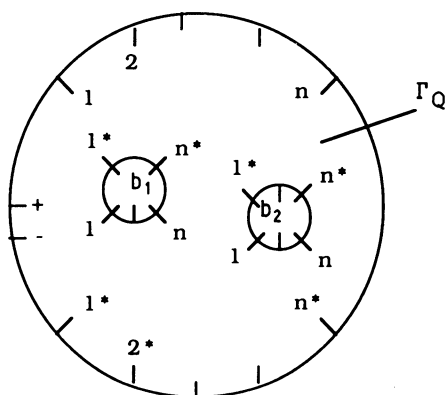


FIGURE 18

7.6 Lemma. A loop (good loop) in Γ_Q (Γ_P) has interior vertices.

Proof. It is as in [Sc₂, 5.1] or [E₁, 3.2]. $\square$

7.7 Lemma. There is no i -cycle in Γ_Q . So there is no Scharlemann cycle in Γ_Q . In particular there is no loop in Γ_Q .

Proof. It is the same as in 2.3, 2.4. $\square$

The following lemmas look like 2.7, 4.6 and 4.7.

7.8 Lemma. There is a 1-path (or a 1*-path) which starts at b_0 and finishes at $+$ or at $-$.

Proof. Take a 1-path γ starting at b_0 . Because there is no 1-cycle, γ either finishes at $+$ or $-$, or in the other side of b_0 . If γ finishes at the bottom part of b_0 , take a 1*-path γ' starting at b_0 , γ' is in the interior of γ . γ' either finishes at $+$, $-$ or at b_0 . If γ' finishes at b_0 then γ and γ' are just the

same path, and they have all their edges labeled $1-1^*$. In this case if there are vertices in the interior of γ , there will be an i (i^*)-cycle, for some i , unless $n = 1$, and so we will be done; if there is no vertex in the interior of γ , then there is a loop joining $+$ and $-$, which contradicts 7.6. So either γ or γ' finishes at $+$ or $-$. $\square$

Suppose without loss of generality that there is a 1-path γ which starts at b_0 and finishes at $+$. Note that this path cannot be just an edge joining b_0 and $+$.

7.9 Lemma. *All the edges adjacent to the top part of b_0 are level, i.e., for each i , $1 \leq i \leq n$, there is an edge with label i at b_0 and label i^* at some other vertex.*

Proof. Take a 1^* -path γ' starting at b_0 . It finishes at $-$ or at the other side of b_0 . Suppose first that it finishes at the other side of b_0 . Take now an i^* -path starting at b_0 for each $i > 1$. As in the proof of Lemma 2.7 each one of these paths finishes at b_0 with a level edge i^*-i , for $i \geq 1$. Suppose now that γ' finishes at $-$. So $-$ is not in the interior of γ ; this implies that all the edges of γ are level and that there is no vertex in its interior, for otherwise there will be a 1^* -cycle in its interior. In particular the edge adjacent to b_0 with label 1 has label 1^* in its other end. Take now an i^* -path starting at b_0 for each $i > 1$, these paths finish at the other side of b_0 , and as before each of these paths finishes with a level edge $i-i^*$. $\square$

7.10 Lemma. *Each vertex in Γ_P has m bad loops, its labels m and 1^* are the only ones which are not adjacent to loops.*

Proof. 7.9 implies that each vertex in Γ_P has a loop, and then all the loops are bad, for if there was a good loop there will be one with no interior vertex, contradicting 7.6. All the loops at the same vertex are concentric. There is an edge η joining a_1 and a_+ , η has label i at a_1 , say, $i \neq 0, m+1$. Then because there is a loop in a_1 , there must be a loop with labels $i-1, i+1$; this contradicts the parity rule unless $i = m$. So η has label m at a_1 , and there is a loop with labels $m-1, (m+1)^*$; by 7.9 there is a loop in a_1 with labels $0, j^*$. The only way that this is possible is that $j^* = 2^*$, and so 1^* and m are the only labels at a_1 not adjacent to some loop; then there are m loops at a_1 . Let η_1 be the edge adjacent to a_1 at the label 1^* . If η_1 meets a_- then $n = 1$, for otherwise there would be good loops, so suppose η_1 meets a vertex a_j . Repeating for a_j the argument made for a_1 , we get that η_1 has label 1^* or m at a_j ; the first is not possible for it implies that there is a loop in Γ_Q . Then we conclude that η_1 has label m at a_j , and that all the labels at a_j are adjacent to loops except 1^* . Repeating the same argument for the other vertices completes the proof. $\square$

7.11 Conclusion of Case 1. 7.10 shows that all the edges in Γ_Q with one end labeled 1, have their other end labeled with 1^* , except at b_m , which has an edge adjacent at its label 1 which meets $+$. Similarly, all the edges with one end labeled 1^* have the other end labeled 1 except the edge adjacent to b_1 , which has its other end at b_m labeled j . This implies that the 1-path γ has all its edges with labels $1-1^*$, except the last one which joins b_m and $+$. Take a 1^* -path γ' starting at b_0 . γ' has to reach b_1 , for otherwise it could be continued indefinitely. So γ' reach b_1 and the next vertex it meets is b_m . 7.10 implies

Proof. For each i , take an i -path starting at the top side of b_0 , say. If the path reaches the other side of b_0 , then its label there is not i , and then the path can be continued, and because there is no i -cycle in Γ_Q , that path has to reach $+$, or $-$. $\square$

7.14 Lemma. Γ_P has at most two vertices.

Proof. Lemma 7.13 implies that $n \leq \#\{+, -\}$. $\square$

It can be assumed that there is a 1-path γ_1 , starting at the top of part of b_0 and finishing at $+$, which is in the west of b_0 ; and there is a 2-path γ_2 , starting at b_0 and finishing at $-$.

7.15 Lemma. γ_1 has only level edges, it has no vertices in its interior, and it has two or more edges.

Proof. If there are vertices in the interior of γ_1 or some edge of γ_1 is not level, then all vertices in the interior of γ_1 (including these vertices in γ_1 which have interior labels) have a label 1^* , and then there is a 1^* -cycle, for a label 1^* cannot be joined to $-$. If γ_1 is just an edge joining b_0 and $+$, this contradicts the fact that all the loops have interior vertices. $\square$

7.16 Conclusion of Case 2. Lemma 7.13 implies that there is an edge η in Γ_P joining a_1 and a_+ , and 7.15 implies that there is a loop λ in a_1 with labels $0, j^*$. If the label of η in a_1 is i , then there should be a loop at a_1 with labels $i-1, i+1$, but this contradicts the parity rule unless $i = m+1$ ($i = 3$ in our case). So η has label $m+1$ at a_1 . This implies that there is an edge η in Γ_Q joining the bottom side of b_0 and $+$. This is in contradiction with 7.15, for both γ_1 and η reach $+$, but by construction η cannot be part of γ_1 . This shows that when k_b is a knot 7.2 cannot happen. This completes the proof of Theorem 3. $\square$

Corollary 2. Suppose k is the trivial and k_b is a composite knot or link. Then k_b has two prime summands, and one of them is a two-bridge knot or link.

Proof. It is similar to the proof of Corollary 1. Here we apply $[E_1]$ (or $[E_2]$, Theorem 6]) to prove that one of the summands is prime. $\square$

S. A. Bleiler has proved in $[B_2]$, using a purely combinatorial argument, that if a banding k_b of a trivial knot k is composite, then k_b has a two-bridge summand.

7.17. It is asked in $[L]$ how many ways there are of summing a rational tangle to a given prime tangle to get a nonprime knot or link. The first examples of prime tangles admitting two different summings of a rational tangle which yield a nonprime knot or link were given in $[B_1]$. In $[E_2]$, Corollary 1] it is shown that there are at most three such summings of rational tangles, and they are at a distance 1 from each other, i.e. they differ by a half twist. Theorems 2, 3 give the following refinement:

Corollary 3. Let (B, t) be a prime tangle. Suppose it can be summed with a rational tangle to get the trivial knot or a split link. Then there is at most one more way of summing a rational tangle to get a nonprime knot or link (in fact it has to be a composite one). If this happens then (B, t) is a partial sum of two tangles, one of them a rational tangle.

Proof. Suppose that $(B, t) + (B_1, t_1) = k_1$ is the trivial knot or a split link, for (B_1, t_1) a rational tangle. If there is another way of summing a rational tangle (B_2, t_2) to (B, t) to get a nonprime knot or link k_2 , then by [E₂, Theorem 3], [Sc₁] or [BS] k_2 has to be composite. Now by [E₂, Theorems 2, 6], (B_1, t_1) and (B_2, t_2) are at distance 1, then k_2 is obtained by a banding of k_1 . So by Theorems 2, 3 there is a decomposing sphere P_2 for k_2 so that $P_2 \cap B_2$ is a single disk, and $P_2 \cap B$ is a disk meeting t in two points, so (B, t) is a partial sum of two tangles, and as in the proof of Corollary 1 one of the tangles is a rational one. If for another rational tangle (B_3, t_3) we have that $(B, t) + (B_3, t_3) = k_3$ is another nonprime link, then as before k_3 is composite, (B_3, t_3) is at distance 1 from (B_1, t_1) , and there is a decomposing sphere P_3 for k_3 such that $P_3 \cap B_3$ is a single disk. By [E₂, Theorem 1] (B_2, t_2) and (B_3, t_3) are also at distance 1. Now it is not difficult to observe that it is not possible that both decomposing spheres P_2 and P_3 intersect B_2 and B_3 respectively in one disk. This observation also follows from [E₂, 2.6, 2.7]. $\square$

Remark. In the case when we have a trivial knot and a composite link, the disk which realizes the partial sum of the tangles intersects only one of the strings of (B, t) , so this is not a partial sum in the sense of [L].

This corollary shows that if there is a prime tangle which yields three nonprime knots or links by attachments of rational tangles, then these knots and links have to be composite, but it is still unknown if such a tangle exists. A positive solution to the following conjecture would imply that such a tangle does not exist.

Conjecture. *Let k be a link and b a band. Suppose that k and k_b are both composite. Suppose that there is no decomposing sphere for k or k_b disjoint from the band. Then there is a decomposing sphere for k or one for k_b which intersects the band in one arc parallel to $b(I \times \{1/2\})$ or to $b(\{1/2\} \times I)$.*

8. SURGERY ON STRONGLY INVERTIBLE KNOTS

In this section we apply Theorem 3 to surgery on a knot. We refer to [R] for definitions and facts about Dehn surgery.

It is known that $(pq, 1)$ surgery on a (p, q) -cable knot gives a nonprime manifold with a lens space summand (see [G]). It is conjectured in [GoS] that those are all the surgeries on a knot which yield a reducible manifold. In other words, the conjecture says that if surgery on a knot in S^3 yields a reducible 3-manifold then the knot is cabled, with the cabling annulus part of the reducing sphere. The conjecture has been proved for several classes of knots, in fact it was recently proved for alternating knots [MT]. See [Sc₄] or [E₃] for more references in this problem. Here we prove the conjecture for strongly invertible knots.

Definition. A knot k in S^3 is strongly invertible if there is an orientation preserving involution of S^3 which carries k to itself, reversing its orientation. Waldhausen [W] shows that such an involution is equivalent to a π -rotation whose axis meets k in exactly two points.

Theorem 4. *Suppose k is a nontrivial strongly invertible knot in S^3 . If some surgery on k yields a reducible manifold then k is cabled, and the surgery is via the slope of the cabling annulus.*

Proof. Let $M(k; r)$ denote the closed 3-manifold obtained by Dehn surgery with coefficient r on k . As k is strongly invertible, $M(k; r)$ double branch covers the 3-sphere [M]. Furthermore the base space and branch set $B(k; r)$ of the double branch covering $p : M(k; r) \rightarrow S^3$ decomposes, in an appropriate projection, as the r -rational tangle attached to a locally prime tangle $\mathbb{B} = (B, t)$, and so that $\mathbb{B}$ plus the $1/0$ -rational tangle is the trivial knot $[B_1]$. $\mathbb{B}$ is in fact a prime tangle if k is nontrivial.

Let $M = M(k; r)$ be a reducible manifold. Let κ denote the knot or link of two components which is the branch set of the respective cover. Suppose first that $M = S^1 \times S^2 \# M'$. [KT, Corollary 4] says that a link is prime if and only if its double branch cover is irreducible. So κ is either split or composite. If κ is composite but nonsplit, say $\kappa = \kappa_1 \# \kappa_2$, then its double branch cover is $M = M_1 \# M_2$, where M_i is the double branch cover of κ_i , so this would imply that $S^1 \times S^2$ is the double branch cover of a nonsplit link, which contradicts the fact that $S^1 \times S^2$ uniquely double branch covers the unlink [T]. So κ is a split link, and κ is obtained by summing $\mathbb{B}$ and the 0 -rational tangle. But $\mathbb{B}$ plus the $1/0$ -rational tangle is the trivial knot, so this says that the unknot can be obtained by a band sum of a split link; $[Sc_1]$ implies that the band and the split link have to be trivial, and then $\mathbb{B}$ is in fact a trivial tangle, implying that k is the trivial knot.

Suppose now that M is reducible but does not contain $S^1 \times S^2$ as a summand. Then [KT, Corollary 4] implies that κ is a composite link. κ is obtained by summing $\mathbb{B}$ and the r -rational tangle, then $[E_1]$ implies that r is an integer. This is equivalent to saying that κ is obtained by a banding of the trivial knot. Theorem 3 implies that there is a decomposing sphere P for κ so that $P \cap (r\text{-rational tangle})$ is a disk disjoint from κ , and $P \cap B$ is a disk meeting κ twice. $A = p^{-1}(P \cap B)$ is an annulus, and each one of its boundary components is a circle of slope r on $\partial \eta(k)$, where r is integer. It is not difficult to see that then k is a cable knot and the slope of the surgery which yields M is that of the annulus. $\square$

Theorem 2 and $[E_2]$ can be used to get results about Dehn fillings on manifolds which are double covers of a prime tangle (B, t) branched along t ; see $[E_3]$

REFERENCES

- [B₁] S. A. Bleiler, *Prime tangles and composite knots*, Knot Theory and Manifolds, Lecture Notes in Math., vol. 1144, Springer-Verlag, Berlin and New York, 1985, pp. 1–13.
- [B₂] ———, *Banding, twisted ribbon knots, and producing reducible manifolds via Dehn surgery*, Math. Ann. **286** (1990), 679–696.
- [BE] S. A. Bleiler and M. Eudave-Muñoz, *Composite ribbon number one knots have two-bridge summands*, Trans. Amer. Math. Soc. **321** (1990), 231–244.
- [BS] S. A. Bleiler and M. Scharlemann, *A projective plane in R^4 with three critical points is standard. Strongly invertible knots have property P* , Topology **27** (1988), 519–540.
- [CGLS] M. Culler, C. Gordon, J. Luecke, and P. Shalen, *Dehn Surgery on knots*, Ann. of Math. **125** (1987), 237–300.
- [E₁] M. Eudave-Muñoz, *Cirugía en nudos fuertemente invertibles*, An. Inst. Mat. Univ. Nac. Autónoma de México **26** (1986), 41–57.
- [E₂] ———, *Primeness and sums of tangles*, Trans. Amer. Math. Soc. **306** (1988), 773–790.

- [E₃] ———, *Band sums of knots and composite links. The cabling conjecture for strongly invertible knots*, Ph.D. thesis, UCSB, 1990.
- [GoS] F. Gonzalez-Acuña and H. Short, *Knot surgery and primeness*, Math. Proc. Cambridge Philos. Soc. **102** (1986), 89–102.
- [G] C. Gordon, *Dehn surgery and satellite knots*, Trans. Amer. Math. Soc. **275** (1983), 687–708.
- [H] J. Hempel, *3-manifolds*, Ann. of Math. Studies, no. 86, Princeton Univ. Press, Princeton, N.J., 1976.
- [KT] P. K. Kim and J. L. Tollefson, *Splitting the P. L. involutions of nonprime 3-manifolds*, Michigan Math. J. **27** (1980), 259–274.
- [L] W. B. R. Lickorish, *Prime knots and tangles*, Trans. Amer. Math. Soc. **267** (1981), 321–332.
- [MT] W. Menasco and M. Thistlethwaite, *Surfaces with boundary in alternating knots exteriors*, preprint.
- [M] J. M. Montesinos, *Surgery on links and double branched covers of S^3* , Ann. of Math. Studies, no. 84, Princeton Univ. Press, Princeton, N.J., 1975, pp. 227–260.
- [R] D. Rolfsen, *Knots and links*, Publish or Perish, 1976.
- [Sc₁] M. Scharlemann, *Smooth spheres in R^4 with four critical points are standard*, Invent. Math. **79** (1985), 125–141.
- [Sc₂] ———, *Unknotting number one knots are prime*, Invent. Math. **82** (1985), 37–55.
- [Sc₃] ———, *Sutured manifolds and generalized Thurston norms*, J. Differential Geometry **29** (1989), 557–614.
- [Sc₄] ———, *Producing reducible 3-manifolds by surgery on a knot*, Topology **29** (1990), 481–500.
- [ST₁] M. Scharlemann and A. Thompson, *Unknotting number, genus, and companion tori*, Math. Ann. **280** (1988), 191–205.
- [ST₂] ———, *Link genus and the Conway moves*, Comment. Math. Helv. **64** (1989), 527–535.
- [T] J. L. Tollefson, *Involutions on $S^1 \times S^2$ and other 3-manifolds*, Trans. Amer. Math. Soc. **183** (1973), 139–152.
- [W] F. Waldhausen, *Über Involutionen der 3-sphere*, Topology **8** (1969), 81–91.

SCHOOL OF MATHEMATICS, INSTITUTE FOR ADVANCED STUDY, PRINCETON, NEW JERSEY 08540
 Current address: Instituto de Matematicas, Universidad Nacional Autonoma de Mexico, 04510 Mexico D.F., Mexico