TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 321, Number 1, September 1990

COMPOSITE RIBBON NUMBER ONE KNOTS HAVE TWO-BRIDGE SUMMANDS

STEVEN A. BLEILER AND MARIO EUDAVE-MUÑOZ

ABSTRACT. A composite ribbon knot which can be sliced with a single band move has a two-bridge summand.

A knot K in the 3-sphere is said to be of *ribbon number* r if K is a ribbon knot which cannot be "sliced" in fewer than r band moves; i.e. r is the minimal number of saddle points on a ribbon disc in the 4-ball having as boundary the knot K. A natural question to ask is if a ribbon number one knot is prime. The answer, of course, is no, as the square knot shows. However, the answer is "nearly" yes. Recall that if a knot K is prime then in every decomposition of K into a connected sum the structure of one of the summands is trivial, i.e. there must be an unknotted summand, or, put another way, a summand with bridge number one. We show for a composite ribbon number one knot, the structure of one of the summands, while possibly nontrivial, must still be "elementary". As advertised in the title, we show

Theorem. *A composite ribbon number one knot has a two-bridge summand.*

The theorem above is a direct corollary of Theorem 1.3 below regarding band sums. The proof of 1.3 uses the combinatorial techniques introduced by M. Scharlemann in [Sc_1] and [Sc_2] to minimize the intersection of planar surfaces in a tangle exterior. A knowledge of these techniques is essential to the understanding of this paper.

The authors wish to thank Marty Scharlemann for helpful conversations and correspondence.

1. THE MAIN THEOREM AND PRELIMINARY ARGUMENTS

Divide a 3-ball into four quadrants by two discs, D_v and D_h, one vertical and the other horizontal. Label the quadrants by the points of the compass NE, NW, SW, SE. Let N be the manifold obtained by attaching two 1-handles to the 3-ball, one connecting NE to SE, the other connecting NW to SW. The

Received by the editors September 30, 1988.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 57M15; Secondary 57M25, 05C10.

Key words and phrases. Ribbon knot, ribbon number, Scharlemann cycle, pure level circuit.

Research supported in part by grants from the National Science Foundation.

1-handles are to be thought of as being attached at points on the boundary of a disc $D_\perp$ in the 3-ball, perpendicular to both D_v and D_h. The disc $D_\perp$ divides the boundary of the 3-ball into two-hemispheres, call that part of the two hemispheres lying in ∂N the *front* face and *back* face of ∂N respectively. Similarly, the disc D_h divides the boundary of the 3-ball into two hemispheres, call that part of these two hemispheres lying in ∂N the *top* face and *bottom* face of ∂N respectively.

1.1. Let A_m be an imbedded family of simple closed curves in ∂N consisting of circles $a_1, \cdots, a_m$ (labelled west to east) parallel to ∂D_v. Let B_n be an imbedded family of simple closed curves in ∂N consisting of circles $b_1, \ldots, b_n$ (labelled south to north) parallel to ∂D_h, together with two circles b_+ and b_- which are meridians of the 1-handles (possibly of the same 1-handle); see Figure 1.

1.2. Imbed into disjoint 3-balls in S^3 two copies γ_0, γ_1 of the unknot, and let $b: I \times I \to S^3$ be an imbedding such that $b^{-1}(\gamma_i) = I \times \{i\}$, $i = 1, 0$. The band sum $\gamma_0 \#_b \gamma_1$ of γ_0 and γ_1 is obtained by joining $\gamma_0 - b(I \times \{0\})$ to $\gamma_1 - b(1 \times \{1\})$ by the arcs $b(\partial I \times I)$.

1.3 **Theorem.** *Suppose* γ_0 *and* γ_1 *are unknots in the* 3*-sphere and a certain band sum* $K = \gamma_0 \#_b \gamma_1$ *is composite. Then* K *has a two-bridge summand.*

1.4. Let K be as in 1.3. As in [Sc_1] we obtain an imbedding of a genus two handlebody N into S^3 and planar surfaces $P(Q)$ in closure($S^3 - N$) with boundary A_m (B_n). The surface P arises from a 2-sphere S_p separating γ_0

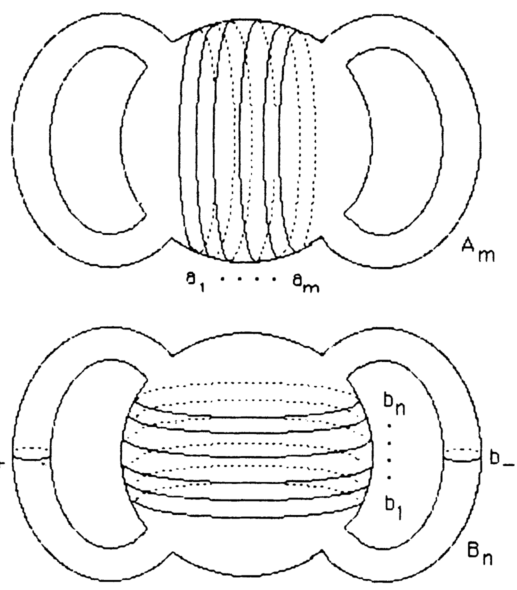


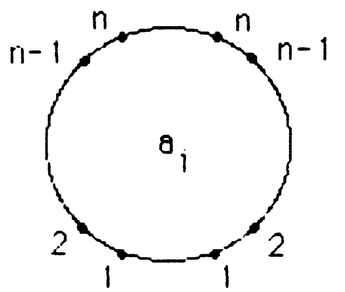
FIGURE 1

from γ_1, the surface Q from a 2-sphere S_Q decomposing K into a nontrivial connected sum. That is, S_Q meets K in two points and separates S^3 into 3-balls B_1 and B_2; each of which meets K in a nontrivial knotted arc. As both γ_0 and γ_1 are unknotted, $n > 0$.

Assume that $m+n$ is minimal with respect to the conditions on the surfaces S_p and S_Q, and that S_Q decomposes K into summands K_1 and K_2, each with bridge number at least three. Further assume that P and Q are in general position so that $P \cap Q$ consists of arcs and circles, and that the number of components of $P \cap Q$ is minimal. In particular, any circle left in $P \cap Q$ is essential in both P and Q.

Construct as follows semioriented graphs Γ_P and Γ_Q in the 2-spheres S_P and S_Q. Regard each a_i (b_s) as a fat vertex in Γ_P (Γ_Q). Regard b_+ and b_- as valence zero vertices in Γ_Q. Finally, regard the arc components of $P \cap Q$ as edges of the graphs Γ_P and Γ_Q. For $1 \leq i \leq m$, $1 \leq s \leq n$ there are two points of intersection of a_i with b_s, each of which label (i, s). The end of an edge in Γ_P (Γ_Q) is assigned the second (first) coordinate of the label above for the corresponding point in $A_m \cap B_n$. In $S_P(S_Q - \{b_+, b_-\})$ therer are $m(n)$ vertices of $\Gamma_P(\Gamma_Q - \{b_+, b_-\})$, each of valence $2n(2m)$. Figure 2 shows how the labels sit around the vertices.

1.5. Orient edges from higher labels to lower; those edges running between identical labels are called *level*. Define a circuit in Γ_P (Γ_Q) to be a subgraph which is an imbedded circle. In Γ_p choose a point x in $S_P - \Gamma_P$ and define the interior of a circuit as the complementary region not containing x. In Γ_Q define the *interior* of a circuit as the complementary region not containing b_-. Further define interior vertex, chord, spoke, loop, base of loop, cycle, unicycle, semicycle, label sequence, interior label, sink, and source as in [Sc_1]. A *level circuit* is a circuit with all edges level. A level circuit is *pure* if every edge has the same label, otherwise a level circuit is *mixed.* The label of the edges of a pure



Labelling about a_i
Labelling about b_s similiar (change n to m)

FIGURE 2

level circuit is the *height* of that circuit. A *good* circuit, cycle, semicycle, or loop in Γ_Q is one whose interior does not contain b_+. A circuit, cycle, semicycle, or loop in Γ_Q whose interior does contain b_+ is *bad*.

1.6. Recall the following elementary facts, cf. [Sc_1, 2.5]:

Proposition. (1) *No chord of an innermost (good) cycle in* Γ_P (Γ_Q) *is oriented.*

(2) *Any chord of an innermost (good) semicycle in* Γ_P (Γ_Q) *is a level loop.*

(3) *If an innermost (good) cycle or semicycle in* Γ_P (Γ_Q) *has an interior vertex it must have an interior source or sink.*

(4) *Any (good) loop in* Γ_P (Γ_Q) *which has interior vertices either has an interior source or sink or else there is a cycle interior to the loop.*

2. Preliminary combinatorics

2.1 **Proposition.** *If* α *is a level edge in a (good) circuit in* Γ_P (Γ_Q) *then the labels which precede and follow the labels of* α *in the label sequence are equal to each other.*

Proof. See [Sc_1, 4.2].

2.2 **Proposition.** *A (good) unicycle in* Γ_P (Γ_Q) *has interior vertices.*

Proof. See [Sc_1, 4.5].

2.3 **Proposition.** *A good semicycle in* Γ_Q *has interior vertices.*

Proof. See [Sc_1, 4.7].

2.4 **Proposition.** *A cycle in* Γ_P *has interior vertices.*

Proof. See [Sc_2, 5.6].

2.5 **Proposition.** *A good loop in* Γ_Q *has interior vertices.*

Proof. Suppose there were an innermost good level loop α based at the vertex b_j in Γ_Q without interior vertices. The two ends of α either have label 1 or label m, without loss of generality, say m. Let D be the interior of α and denote $\partial D \cap \partial N$ by α'. The arc α' is an arc in b_j with endpoints the two labels m and meets no other label. Form a new planar surface P' by attaching the band B given by a regular neighborhood of α' in ∂N, see Figure 3.

The surface P' as boundary components $a_1, \cdots, a_{m-1}$ and two other components a_+, a_-, which (up to isotopy) are meridians of a 1-handle. By Figure 4, ∂D separates a_+ from a_- in P'.

Compressing P' via D gives two planar surfaces P_+, P_- each of which has a single boundary component on a 1-handle and any other boundary component parallel to D_h. Now it is easy to construct a simple closed curve on ∂N intersecting the boundary of P_+ in a single point, and hence intersecting in one point the sphere S in M obtained by attaching a disc to each component of ∂P_+ in N.

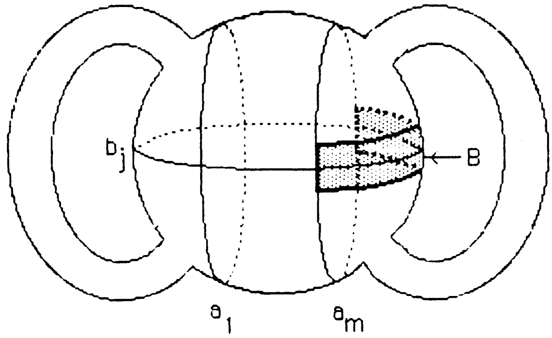


FIGURE 3

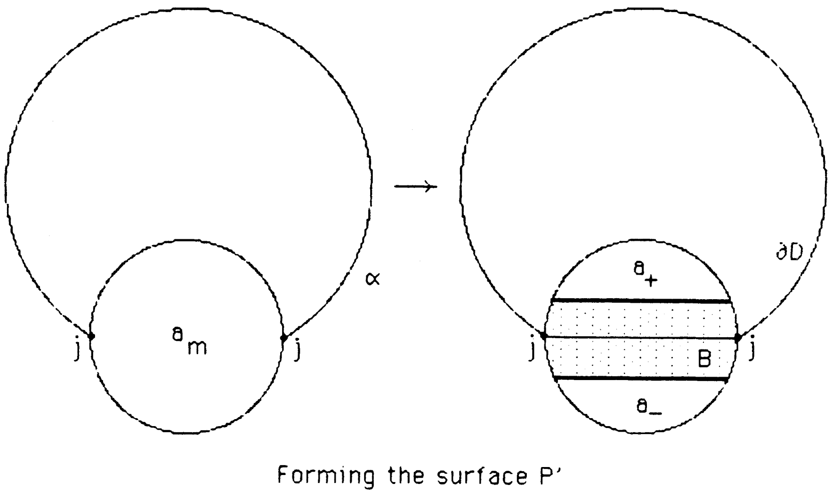


FIGURE 4

2.6 **Proposition.** *If* Γ_Q *contains a good loop then* Γ_P *contains a unicycle.*

Proof. See [Sc_1, 5.2].

2.7 **Proposition.** *If* Γ_Q *contains a good loop, then* Γ_Q *does not contain both a source and a sink.*

Proof. See [Sc_1, 5.4].

2.8 **Proposition.** *Good loops do not exist in* Γ_Q.

Proof. In the presence of 2.1–2.7, the path threading argument of [Sc_1, §6] and the minimality of $m + n$ show that there cannot be an innermost good semicycle in Γ_Q. So no good unicycle exists in Γ_Q. Similarly, no good level loop can exist in Γ_Q, as applying the path threading argument of [Sc_1, §6] to the interior of a good level loop in Γ_Q produces a good semicycle interior to that loop.

3. PURE LEVEL CIRCUITS IN Γ_P

3.1 **Lemma.** *A pure level circuit in Γ_P without oriented chords has interior vertices.*

3.2 *Proof of* 3.1. Suppose Ω is an innermost counterexample and that Ω has height k.

3.3 **Proposition.** Ω *has no chords.*

3.4 *Proof of* 3.3. Suppose α_0 is a level chord of Ω with label l. The ends of α_0 are on distinct vertices of Ω as a level loop would form a pure level circuit without oriented chords interior to Ω. By Figure 2, the other label l at either of these vertices is an interior label of Ω, and so must be incident to another level chord α_1, with label l. Continuing, one forms a pure level circuit without oriented chords interior to Ω.

3.5 *Continuation of* 3.2. As Ω has no chords, Ω is of height 1 or n. Without loss of generality, say n. Moreover, the interior of Ω is a disc, denote this disc by D.

Claim. ∂D lies entirely on one side of b_n.

Proof of claim. If $n > 1$, this is immediate. If $n = 1$, the claim follows from a slight modification of the argument in [Sc_1, 4.2].

Let D be the interior of Ω and let Q' be the planar surface obtained by attaching to Q the band B in ∂N which contains the "tops" of the vertices $a_1, \cdots, a_n$, see Figure 5.

The surface Q' has boundary $b_1, \cdots, b_{n-1}, b_+, b_-$, and two new boundary components b_* and $b_\#$ which are (up to isotopy) meridians of the 1-handles. Call any one of the circles b_+, b_-, b_*, or $b_\#$ a *1-handle boundary* of Q'.

Let $S_{Q'}$ be the two-sphere in S^3 obtained by capping off the planar surface Q' with discs in N. The surface $S_{Q'}$ is a Conway sphere for the composite knot K, decomposing S^3 into tangles T_1, in which D is a proper disc, and

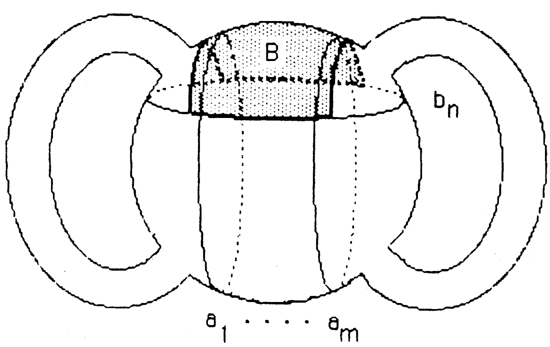


FIGURE 5

T_2 whose distinguished arcs are separated by the disc in N bounded by b_n. One of the distinguished arcs of T_2 is unknotted and the other distinquished arc is knotted according to one of the summands, say K_2, of K. It follows that the other summand K_1 of K is obtained by attaching an untangle, i.e. a trivial (3-ball, arc pair) to the tangle T_1.

3.6 **Proposition.** *The tangle* T_1 *is an untangle.*

3.7 *Proof of* 3.1 *from* 3.6. By assumption the bridge number of the summand K_1 is at least three. From 3.5 and 3.6 we conclude that the summand K_1 can be expressed as the union of two untangles. It follows that K_1 has bridge number at most two, a contradiction.

3.8 **Proposition.** *The distinquished arcs of* T_1 *are separated by the disc* D.

3.9 *Proof of* 3.6 *from* 3.8. It remains to show that there are no local knots in the distinquished arcs of T_1. Compressing the surface Q' via the disc D one obtains two planar surfaces Q^* and $Q^\#$. The planar surface Q^* $(Q^\#)$ in closure $(M - N)$ has exactly two boundary components which are meridians of the 1-handles (possibly of the same 1-handle) and at most $n-1$ other boundary components, each of which is parallel to ∂D_h. So the two-sphere S_{Q^*} $(S_{Q^\#})$ formed by capping of the surface Q^* $(Q^\#)$ with discs in N meets K in two points and separates S^3 into 3-balls B_1 and B_2. One of these balls, say B_1, is contained in the tangle T_1 and contains one of the two distinquished arcs (the other distinquished arc) of T_1. A local knot in either distinquished arc of T_1 would therefore contradict the minimality of $m + n$.

3.10 *Proof of* 3.8. Proposition 3.8 follows immediately from

3.11 **Proposition.** *No three of the* 1*-handle boundaries of* Q' *lie in the same component of* $Q' - \partial D$.

3.12 *Proof of* 3.11. Assume at least three of the four 1-handle boundaries of Q' lie in the same component of $Q' - \partial D$. There are two cases to consider.

Case 1. The circuit Ω is a level loop.

Proof of Case 1. By 2.8 the level loop α in Γ_Q corresponding to Ω is bad, and hence α separates b_+ from b_- in Γ_Q. As in the proof of 2.5 it follows that ∂D separates $\{b_+, b_*\}$ from $\{b_-, b_\#\}$, see Figure 6.

Case 2. The circuit Ω has more than one edge.

Proof of Case 2. It is essential to what follows to have a good picture of the family l of loops in Γ_Q corresponding to the edges of Ω in Γ_P and how their endpoints are distributed about b_n. A loop in l will be called an *l-loop* and a label on b_n incident to an l-loop an *l-label.* Both labels n on a vertex a_i of Ω in Γ_P are incident to edges of Ω, so both labels i on the vertex b_n in Γ_Q are l-labels. In particular, there is an equal number of l-labels on the front and back of b_n.

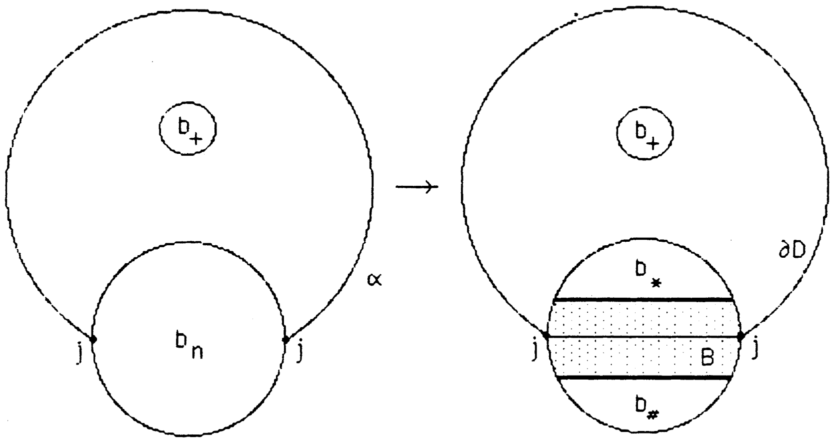


FIGURE 6

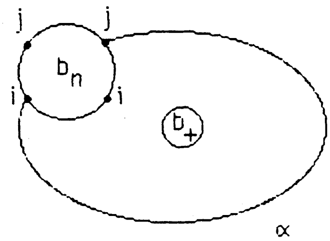


FIGURE 7

By 3.7, l is a family of bad loops based at a common vertex, so every l-loop separates b_+ from b_- in Γ_Q and the l-loops are linearly ordered by "interior".

3.13 **Definition.** A loop α in Γ_P is a *lobe* if both ends of α lie on either the front face or the back face of b_n. This is a slightly more general definition of a lobe than that of [BS, 6.1]

3.14 **Proposition.** *The inner- and outermost l-loops are lobes.*

3.15 *Proof of* 3.14. Consider a nonlobe α in l. As Ω is not a level loop, the labels i, j at the endpoints of α are not equal, say $i < j$. From Figure 7 we see that exactly one of the other l-labels i, j on b_n not incident to α is an interior label of α while the other is an exterior label.

Hence there are members of l interior and exterior to α.

3.16 *Continuation of* 3.12. Now refer to certain intervals on b_n as "points of the compass". In particular, call that interval meeting no label and having endpoints the two labels $n(1)$ north (south) and the interval incident to the inner (outer) most component of $Q - \{l \cup b_n\}$ as east (west). Call the four complementary intervals on b_n the NW, SW, SE, or NE *quadrant* of b_n as appropriate. Define NW, SW, SE, or NE (l-) *labels* in the obvious manner.

By 2.8, no l-loop joins NW to NE or SW to SE. Hence any nonlobe in l joins NE to SW or NW to SE. However, it is impossible to draw both types of nonlobes on the planar surface P, cf. [Sc_2, 3.2]. Without loss of generality then, assume all nonlobes in l join NE to SW, and hence that any member of l with a NW or SE label is a lobe. Proposition 3.14 implies that the number of l-labels in each quadrant is nonzero. The family l is as illustrated below in Figure 8.

To see how ∂D sits on Q', add the band B and recall that on B the curve ∂D is parallel to the core of B. Figure 9 and the convention above imply that the four 1-handle vertices of Q' correspond to the four cardinal directions N, S, E, and W on b_n. By the Jordan curve theorem, two 1-handle vertices are in the same component of $Q' - \partial D$ if and only if there is an arc joining the corresponding cardinal directions in b_n which contains an even number of l-labels.

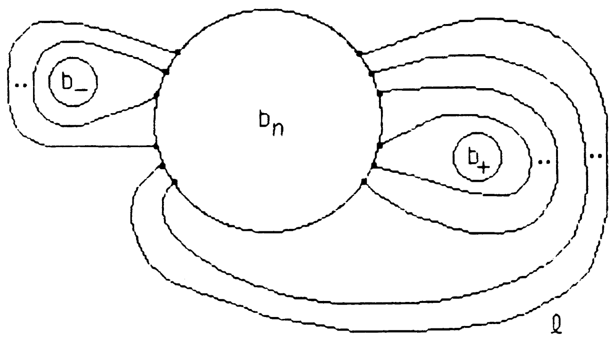


FIGURE 8

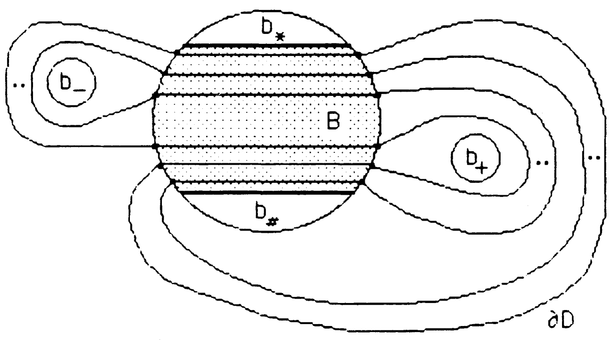


FIGURE 9

3.17 **Proposition.** *The curve ∂D separates b_+ from b_* in Q' if and only if ∂D separates b_- from $b_\#$ in Q'.*

3.18 **Corollary to 3.17.** *No 1-handle boundary in Q' is separated from the other three by ∂D.*

3.19 *Proof of* 3.17. There are an equal number of l-labels on the front and back face of b_n. As a nonlobe has one end on the front and one end on the back of b_n, the front and back face of b_n have an equal number of l-labels corresponding to nonlobes, hence equal numbers corresponding to lobes. From Figure 8, it follows that this number is exactly twice the number of NW (SE) l-labels. Hence there are an equal number of NW and SE l-labels. The curve ∂D separates b_+ from $b_\#$ (b_- from b_*) in Q' if and only if this number is odd.

3.20 *Continuation of* 3.12. By 3.18, all four of the 1-handle boundaries of Q' must lie in the same component of $Q' - \partial D$, so there must be an even number of l-labels in each quadrant. It follows that for some $k > 0$ there are $4k$ l-labels in all and so the circuit Ω has exactly $2k$ edges.

3.21 *The final contradiction.* We will contradict 3.20 by counting the edges of Ω another way. Consider traversing ∂D in Q' by beginning on and traveling in the direction of the arrow of the innermost loop in l with a NW label. Call an l-label where we pass from Q to B (B to Q) as we traverse ∂D an *exit* (*entry*) *point*, see Figure 10.

Map an exit (entry) point to the next exit (entry) point encountered as ∂D is traversed in the manner above to obtain a cyclic permutation of the $4k$ l-labels. As ∂D runs parallel to the core of b, an exit (entry) point moves a fixed number r of l-labels counter-clockwise (clockwise) about b_n under this permutation. From Figure 9, r equals twice the number of NW l-labels. As this latter number is even, $r = 4j$ for some $j > 0$.

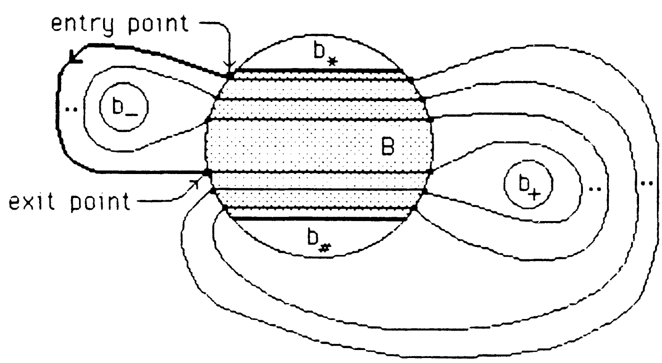


FIGURE 10

The number of edges in Ω is just the cardinality of an orbit of an exit (entry) point under this permutation. It follows that Ω has exactly

$$\frac{4k}{\text{g.c.d.}(4k, 4j)} \leq k \text{ edges}.$$

This contradicts 3.20.

4. x-CIRCUITS

4.1 **Lemma.** *There is either a cycle or a pure level circuit in Γ_P without chords and without interior vertices.*

Choose orientations for the planar surfaces P and Q, and let these induce orientations on their respective boundaries A_m and B_n. Call two components of A_m (B_n) parallel if their orientations are parallel in ∂N; if not call them antiparallel. There are two points of intersection of a_j with b_s, one on the front face and one on the back face of N. Label them respectively $f(i, s)$ and $r(i, s)$. As in [Sc$_1$, 2.2] we have the following "parity rule".

4.2 **Proposition.** *If an arc of $P \cap Q$ runs between $f(i, s)$ and $f(j, t)$, then one of the pairs a_j and a_j or b_s and b_t is parallel and the other is antiparallel. The same is true for an arc of $P \cap Q$ which runs between $r(i, s)$ and $r(j, t)$. On the other hand, if an arc of $P \cap Q$ runs between $f(i, s)$ and $r(j, t)$, then either both pairs are parallel or both are antiparallel.*

There are two classes of vertices in Γ_P, each consisting of parallel vertices. A typical vertex in each class is illustrated in Figure 11, where $f(s)$ $(r(t))$ denotes that the label s (t) lies on the front (back) face of N.

We form a new graph Γ_{P^*} from the graph Γ_P as follows: Reenumerate the labels of the vertices in the first class given above by changing $f(i)$ to i, and $r(j)$ to $2n + 1 - j$, see Figure 12.

Similarly, reenumerate the labels of the vertices of the second class by changing $r(i)$ to i, and $f(j)$ to $2n + 1 - j$, see Figure 13.

By construction, therefore, all the vertices in Γ_{P^*} are parallel. As before, orient edges from higher labels to lower; and continue to call level those edges running between identical labels. Proposition 4.2 now implies the following.

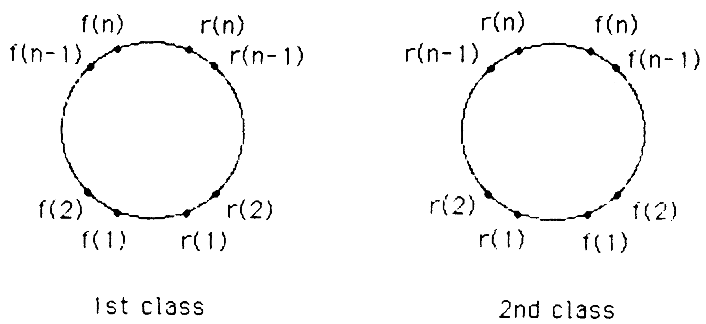


FIGURE 11

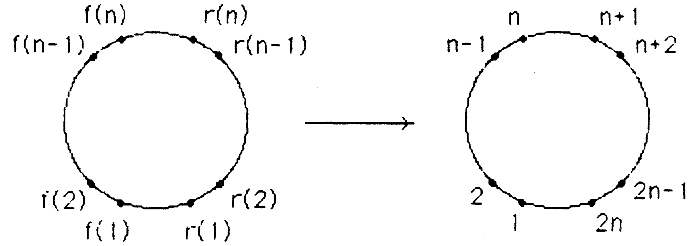


FIGURE 12

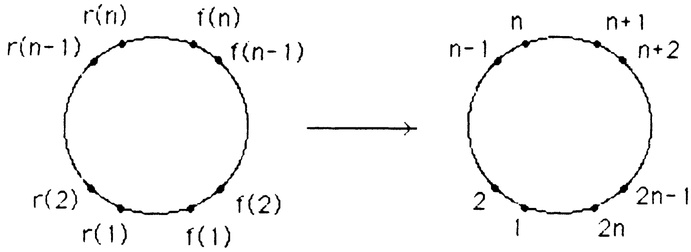


FIGURE 13

4.3 **Proposition.** *There are no level edges in* Γ_{P^*}.

For $x \in \{1, 2, \dots, 2n\}$ define an x-circuit to be a circuit which can be traversed in a direction so that the initial point of every edge has label x.

4.4 **Proposition.** *In* Γ_{P^*} *there is an* x*-circuit without interior vertices and without chords.*

4.5 *Proof of* 4.4. (cf. [CLGS, 2.6.2]). Choose an $x \in \{1, 2, \dots, 2n\}$ and construct a path beginning at some vertex v, always choosing the label x as the initial point of each edge; by 4.3 this path can be followed until a vertex is repeated, forming an x-circuit. Choose an innermost such circuit (varying over all x), and denote it by Ω and suppose that Ω has interior vertices or chords.

If Ω has no interior labels but has interior vertices, then for any $y \in [1, 2, \dots, 2n\}$ it is possible to find a y-circuit interior to Ω. So suppose Ω has interior labels. Choose a vertex v in Ω, and choose the label on v which is adjacent to x and is either an interior label of Ω or a label of an edge in the circuit. This label is either $x-1$ or $x+1 \pmod{2n}$, without loss of generality, suppose it is $x-1$. Since all the vertices of Γ_{P^*} are parallel, for each vertex of Ω the label $x-1$ is either an interior label or a label of an edge in the circuit. The circuit Ω has interior labels, so there is at least one vertex u in Ω for which the label $x-1$ is an interior label. Beginning at u, we can construct a path in which every edge has initial label $x-1$; when a vertex is repeated we obtain an $(x-1)$-circuit, call it Ω'.

The circuit Ω' must be interior to Ω as each vertex in Ω or in its interior either has the label $x-1$ as an end of an edge of Ω, or as an end of an interior edge. The circuit $\Omega' \neq \Omega$ as the edge of Ω incident to u with label at u different from x cannot be part of Ω' as its label at u is also different from the interior label $x-1$. This contradicts the fact that Ω was innermost, and 4.4 follows.

4.6 *Proof of* 4.1. The x-circuit in Γ_{P^*} given by 4.4 has a consistent orientation of its edges so that the initial point of every edge has label x and the terminal point of every edge has label $x-1$. It follows that in Γ_P this circuit is either a cycle or a pure level circuit.

4.7 *Proof of* 1.3. A contradiction exists between 2.4, 3.1, and 4.1.

5. Concluding remarks

5.1. A ribbon disc for K^2, the connected sum of a knot K with its mirror image, is obtained by spinning an appropriately knotted arc in R^3_+. If K has bridge number b, this ribbon disc has $b-1$ saddle points. Theorem 1.3 shows that the ribbon number of K^2 is $b-1$ for $b<4$. So it seems appropriate to conjecture that the ribbon number of K^2 always equals the bridge number of K minus 1.

References

[BS] S. A. Bleiler and M. G. Scharlemann, *A projective plane in R^4 with three critical points is standard*, MSRI preprint 07112-85; 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.

[Sc_1] M. G. Scharlemann, *Smooth spheres in R^4 with four critical points are standard*, Invent. Math. **79** (1985), 125–141.

[Sc_2] ——, *Unknotting number one knots are prime*, Invent. Math. **82** (1985), 37–55.

Department of Mathematics, Portland State University, Portland, Oregon 97207-0751

Department of Mathematics, University of California at Santa Barbara, Santa Barbara, California 93106