Topology Vol. 27, No. 4, pp. 519–540, 1988
Printed in Great Britain.

0040–9383/88 $3.00+.00

A PROJECTIVE PLANE IN $\mathbb{R}^4$ WITH THREE CRITICAL POINTS IS STANDARD. STRONGLY INVERTIBLE KNOTS HAVE PROPERTY *P*

STEVEN BLEILER and MARTIN SCHARLEMANN†

(*Received in revised form* 9 *February* 1988)

LET $\mathbb{P}$ be the projective plane in $\mathbb{R}^4$ obtained by capping off the boundary of an unknotted Möbius band in $\mathbb{R}^3 \times \{0\}$ with an unknotted disk in $\mathbb{R}^3 \times [0, \infty)$. Here we show that any smoothly imbedded projective plane in $\mathbb{R}^4$ on which some projection $\mathbb{R}^4 \to \mathbb{R}$ has three non-degenerate critical points is isotopic to $\mathbb{P}$. The proof is based on a combinatorial solution to Problem 1.2B of [4]. In particular, if a band is attached to an unknot so that the result is an unknot, then the band is isotopic to the trivial half-twisted band. One consequence is that strongly invertible knots have property *P* (see [1]). Together with [2], this further implies that pretzel knots (indeed all symmetric knots) have property *P*.

The solution of 1.2B uses the techniques of [5], [6] and [1], with careful distinction made between the two sides of the planar surfaces used in those arguments. Here is the philosophy: It was pointed out in [1] that the techniques of [5] and [6] were inadequate, because a certain type of semi-cycle ([1] Fig. 3) may arise, and yet cannot be used in the reduction process because there is no control over the side of the planar surface on which the interior of the semi-cycle lies. This was easily circumvented in [5], [6] because one of the planar surfaces involved was a punctured sphere. Here the planar surfaces are punctured disks, and it is difficult to detect whether a potential reducing disk is incident to just one side of the disk. In principle, simple bookkeeping should circumvent the problem. Yet there is a crucial step in the argument ([5, 6.3]) at which all distinction between these sides seems hopelessly lost.

In 6.9 an extremely weak analogue of [5, 6.3] is recovered, however, and prompts the study of "special paths". These special paths are used, in three successive constructions of "multiflows", to produce a semi-cycle whose interior must lie on a single side of a planar surface, and hence can be used to reduce the complexity of that surface. Roughly, the goal of the three successive stages is to remove from the interior of a multiflow first all sources, then all vertices, then all apexes. During the entire process, care is taken to ensure that at the end, all edges will have the same side on the interior.

§1. THE MAIN THEOREM AND TOPOLOGICAL CONSEQUENCES

1.1 Divide a 3-ball into four quadrants by two 2-disks, D_v and D_h, one vertical and one horizontal. Label the quadrants by the points of the compass NE, NW, SW, SE. Let N be the 3-manifold obtained by attaching two 1-handles to the 3-ball, one connecting SE to NW, the other connecting SW to NE.

Supported in part by a National Science Foundation grant.

1.2 (See Fig. 1.) Let A_m be an imbedded family of simple closed curves in ∂N consisting of circles $a_1, \ldots, a_m$ (labelled south to north) parallel to ∂D_h, together with a circle α going once over each 1-handle. Let B_n be an imbedded family of n simple closed curves in ∂N consisting of circles $b_1, \ldots, b_n$ (labelled west to east) parallel to ∂D_v, together with a circle β going once over each 1-handle. Note that there is ambiguity in our definition of A_m and B_n because the number of times α and β wrap around each 1-handle as they pass over is left undefined. In particular, if this choice differs for α and β by the integers i_e and i_w in the eastern and western 1-handles respectively, there are $|i_e|$ and $|i_w|$ points of intersection in the 1-handles.

THEOREM 1.3. *Suppose that N is imbedded in an oriented 3-manifold M so that some A_m and some B_n bound imbedded planar surfaces P, Q in closure (M−N). Then some A_0 and B_0 bound imbedded disks E_P and E_Q in closure (M−N) and either:*

(a) *E_P and E_Q are disjoint and $\mathbb{RP}^3$ is a summand of M or*

(b) *There is an imbedded disk D in closure(M−N), disjoint from E_P and E_Q, such that ∂D is the union along the boundary of two arcs, one crossing a 1-handle once, and the other lying in the boundary of the 3-ball.*

THEOREM 1.4. *No 3-manifold obtained by surgery on a strongly invertible knot is simply-connected, i.e. strongly invertible knots have property P.*

Proof. See [1, 1.7]. ||

COROLLARY 1.5. (Seifert, see [3]) *Torus knots have property P.*

COROLLARY 1.6. (Takahashi [7]) *2-bridge knots have property P.*

Proof of 1.5 *and* 1.6. The torus and 2-bridge knots are strongly invertible. ||

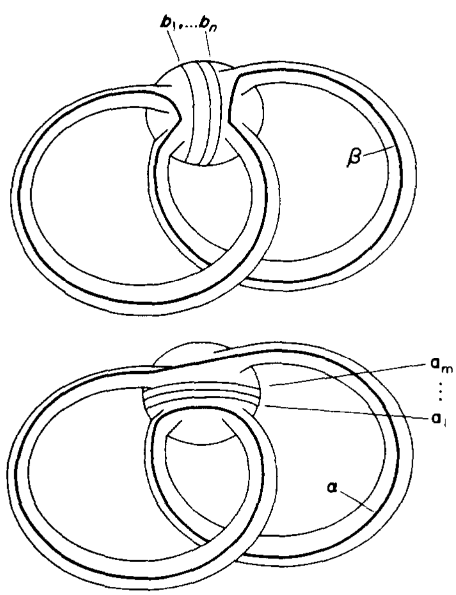


Fig. 1.

COROLLARY 1.7. *Pretzel knots have property P.*

Proof. The n-braid pretzel link $P(a_1, a_2, \ldots, a_n)$ is a knot if either all the a_i's are odd or if exactly one of the a_i's is even. In the former case, $P(a_1, a_2, \ldots, a_n)$ is periodic (with period 2) and so has property P by the recent results of Culler, Gordon, Lueke and Shalen [2]; in the latter $P(a_1, a_2, \ldots, a_n)$ is strongly invertible (see Fig. 2). ||

Oertel has pointed out that 1.7 extends readily to Montesinos knots (in which the twists in the pretzel knot are replaced by more general rational tangles) and Gordon that it extends in fact to all symmetric knots, though the latter requires the full strength of [2], not just its combinatorics.

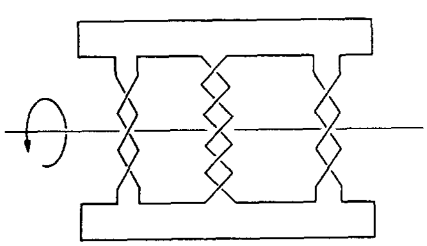
1.8 Suppose γ is a knot in S^3 and a band $b\colon I \times I$ 0. S^3 is attached so that $b^{-1}(\gamma) = I \times \partial I$. Let γ_b be the knot obtained by replacing $b(I \times \partial I)$ in γ with $b(\partial I \times I)$.

THEOREM. *Suppose γ is unknotted, and so bounds a disk E_P in S^3. If γ_b is also unknotted, then γ_b bounds an unknotted Möbius band in S^3 which contains E_P.*

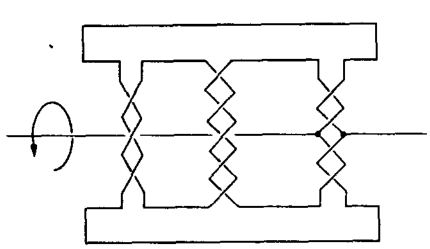
Proof. Let N be a regular neighborhood of $\gamma \cup b(I \times I)$. By general position (with $b(\{1/2\} \times I)$), E_P can be isotoped so that $b^{-1}(D) = I \times \{a_i\}$ for some finite $\{a_i\} \subset I$. Similarly, by general position with $b(I \times \{1/2\})$, there is a disk E_Q with boundary γ_b such that $b^{-1}(D) = \{b_s\}$ for some $\{b_s\} \subset I$. Then it is easy to arrange that the complements P and Q of the interior of N in E_P and E_Q respectively are surfaces satisfying all the hypotheses of 1.3.

Conclusion 1.3(a) does not apply, so 1.3(b) does. Since $m = n = 0$, $E_P \cup b(I \times I)$ is an I-bundle on a circle, say with p half-twists. The disk D ensures that the circle is unknotted. Then γ_P is a (2, p) torus knot. Since it is unknotted, $p = 1$. ||

1.9 A surface M in $\mathbb{R}^4$ has p critical points if a projection $p\colon \mathbb{R}^4 \to \mathbb{R}$ is a Morse function on M with p non-degenerate critical points.



P(3,5,3)



P(3,5,4)

Fig. 2.

THEOREM. *If f, g*: $\mathbb{RP}^2 \to \mathbb{R}^4$ *are imbeddings such that* $f(\mathbb{RP}^2)$ *and* $g(\mathbb{RP}^2)$ *each have three critical points, then f and g are isotopic.*

Proof. The three critical points must be a minimum, a maximum, and a single saddle. The proof now follows from 1.8 (cf. [5, 1.3]). ||

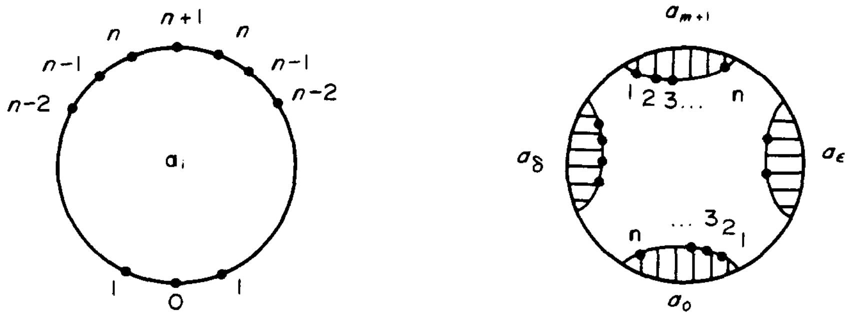
§2. PRELIMINARIES

2.1 The points of intersection of A_m and B_n can be labelled as follows: For $1 \le i \le m$, $1 \le s \le n$ there are 2 points of intersection of a_i with b_s, each of which we label (i, s). The curve α intersects the boundary of the 3-ball in two arcs, α_+ in the north and α_- in the south. For $1 \le s \le n$ label the points of intersection of b_s with α_+ and α_- by $(m+1, s)$ and $(0, s)$ respectively. The curve β intersects the boundary of the 3-ball in two arcs, β_+ in the east and β_- in the west. For $1 \le i \le m$ similarly label the points of intersection of a_i with β_+ and β_- by $(i, n+1)$ and $(i, 0)$ respectively. Points of intersection of $\alpha_\pm$ with $\beta_\pm$ are labelled one of $(0, 0)$, $(m+1, 0)$, $(0, n+1)$, or $(m+1, n+1)$ by the obvious convention. Other points of $\alpha \cap \beta$ (those lying in the 1-handles) are unlabelled.

We now proceed just as in [5], [6] and [1]: Suppose N is imbedded in an oriented 3-manifold M in such a way that A_m and some B_n bound planar surfaces P and Q respectively. Assume that we have minimized $m+n$, put P and Q in general position so that intersections consist of arcs and circles, and then (by disk-swapping and isotoping) reduced as much as possible the number of components of intersection of P and Q. We have thereby removed all circles in $P \cap Q$ except those essential in both.

Construct as follows semi-oriented graphs Γ_P and Γ_Q in 2-disks D_P and D_Q with $\partial D_P = \alpha$ and $\partial D_Q = \beta$. Let D_P and D_Q be the disks obtained by filling in disks on the boundary components a_i, $1 \le i \le m$ and b_s, $1 \le s \le n$, respectively. Regard each a_i (b_i) as a fat vertex in Γ_P (Γ_Q). Regard the two arcs α_+ and α_- (β_+ and β_-) as vertices a_{m+1} and a_0 (b_{n+1} and b_0) in Γ_P (Γ_Q) lying on ∂D_P (∂D_Q). Regard the two arc components of $\alpha(\beta)$ in the 1-handles as two vertices a_δ and a_ε (b_δ and b_ε) in Γ_P (Γ_Q) lying on ∂D_P (∂D_Q). Finally, regard the arc components of $P \cap Q$ as edges of the graphs Γ_P and Γ_Q. Each end of each edge in Γ_P and Γ_Q represents a point in $A_m \cap B_n$. The end of an edge in Γ_P (Γ_Q) not incident to a_δ or a_ε (β_δ or β_ε) is assigned the second (first) coordinate of the label above for the corresponding point in $A_m \cap B_n$. In the interior of D_P (D_Q) are m (n) vertices of Γ_P (Γ_Q), each of valence $2n+2$ $(2m+2)$. The labels around the vertices are as shown in Fig. 3. If the first or last points of intersection of $\alpha_\pm$ with B_n are with $\beta_\pm$, push the points into the 1-handles, removing labels 0 and $n+1$ at a_0 and a_{m+1} in Γ_P (and labels 0 or $m+1$ at b_0 and b_{n+1} in Γ_Q). This ensures that a_0 and a_{m+1} (b_0 and b_{n+1}) each have valence $n(m)$.

Orient edges from higher labels to lower; those edges running between identical labels are called *level*, those edges with an end on a_δ or a_ε (b_δ or b_ε) are not oriented and are called *bad edges*. A simple orientation argument shows that, since all intersections of α and β in a 1-handle must have the same sign, no bad edge is a loop. Define a *circuit* γ in Γ_P (Γ_Q) to be a closed (not necessarily imbedded) path such that for a regular neighborhood $\eta(\gamma)$ of γ there is a boundary component γ' of $\eta(\gamma)$ such that γ' is homotopic to γ in $\eta(\gamma)$ and bounds a single disk component of $D_P - \eta(\gamma)(D_Q - \eta(\gamma))$ called the *interior* of γ. This is a slight generalization of the definition used in [5], [6] and [1] which insists γ be imbedded. Here γ is obtained from the simple closed curve by pinching in the *exterior* of γ'. Further define interior vertex, chord, spoke, loop, base of a loop, cycle, semi-cycle, label sequence, interior label, sink, and source as in [5] (using this slightly generalized notion of circuit). A *good* circuit, cycle, or semi-cycle in Γ_P (Γ_Q) is one which does not contain a vertex a_δ or a_ε (b_δ or b_ε). In particular, all circuits lying in the interior of a good circuit are good.



Labelling about a_i and ∂P

Labelling about b_s similar (change n to m)

Fig. 3.

§3. PRELIMINARY TOPOLOGICAL ARGUMENTS

LEMMA 3.1. *Theorem* 1.3 *is true if* $m=0$ *or* $n=0$.

Proof. With no loss of generality assume $n=0$ so Q is a disk, in which $P\cap Q$ appears as a collection of arcs. If $P\cap Q$ is empty, then $m=n=0$ so P is also a disk and 1.3(a) applies. Otherwise, choose an outermost arc γ in Q. There are four possibilities.

(i) The ends of γ are labelled i and $i+1$ lying both on b_0 or both on b_{n+1}. Then m can be reduced by two using the outermost disk cut off from Q by γ.

(ii) One end of γ lies on b_δ or b_ε and the other is labelled 1 or m. Then m can be reduced by one using the outermost disk cut off from Q by γ.

(iii) One end of γ is in b_δ and one in b_ε. Then $m=0$, so P is also a disk. Then the union along γ of the outermost disk of Q cut off by γ and each of the components into which γ divides the disk P are the 2-disks satisfying the conclusion of 1.3(b).

The only remaining possibility is:

(iv) The ends of γ are labelled 1 and m, with one end in b_0 and one in b_{n+1}. Then the arc of β cut off by the ends of γ must pass once over a one-handle of N. Thus there are exactly two of these outermost arcs, and between them all arcs of $P\cap Q$ run parallel, with ends labelled i and $m-i+1$, for $1\le i\le m$. (See Fig. 4.) This is impossible if m is odd (apply [5, 4.2] to $i=(m+1)/2$) and allows a reduction of 2 in m if m is even. The reduction uses a component of $Q-P$ to do surgery on the union of P and the annulus on ∂N between $a_{m/2}$ and $a_{(m/2)+1}$. (See [5, 4.7, case 1]). ||

Henceforth we therefore assume $m\neq 0\neq n$.

LEMMA 3.2. Γ_P *and* Γ_Q *have no good level loops.*

Proof. A parity argument as in [5, 4.4]. ||

LEMMA 3.3. *A good cycle in* Γ_P *or* Γ_Q *has interior vertices.*

Proof. See [5, 4.8]. ||

§4. ELEMENTARY COMBINATORICS

LEMMA 4.1. *There is a good cycle in* Γ_P *or in* Γ_Q.

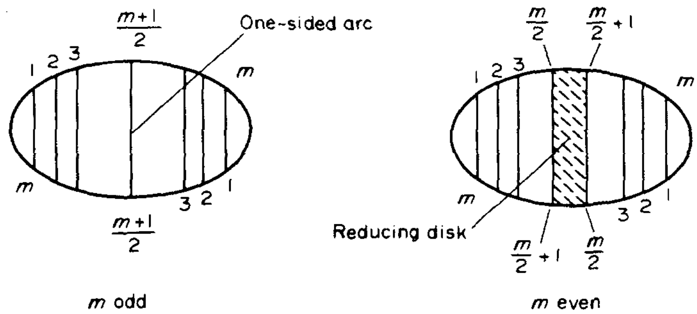


Fig. 4.

Proof. We will assume there are no good cycles and derive a contradiction. First note that if there is any level edge in Γ_P with label s, say, then the corresponding edge in Γ_Q is a (good) loop based at b_s. Hence we can assume that there are no level edges in either graph.

Call a bad edge in Γ_P (Γ_Q) which has label l on one end and other end in a_δ or a_ε (b_δ or b_ε) a bad l-edge. If there are any bad 0-edges or $(m+1)$-edges in Γ_Q or bad 0-edges or $(n+1)$-edges in Γ_P, the proof concludes as in [1, 4.1, case ii]. We henceforth assume that there are no bad 0-edges or $(m+1)$-edges in Γ_Q, nor bad 0-edges or $(n+1)$-edges in Γ_P.

Remove from consideration all bad edges in Γ_Q to get a new graph Λ on the 2-disk D_Q. Since there are no level edges and each vertex b_s, $1 \le s \le n$, has a label 0 and a label $m+1$ remaining in Λ, none of these can be a source or a sink in Λ. Thus either there is a good cycle, or one of b_0 and b_{n+1} is a source (say b_{n+1}) and one is a sink (say b_0). Double D_Q along its boundary to produce a 2-sphere Σ containing the double 2Λ of Λ. Now proceed as in [6, 6.1]. Since there are no level edges, we can extend the orientations of the edges to a vector field in the neighborhood of the 1-skeleton of Λ, which has a singularity of index ≤ 0 at each vertex b_s, $1 \le s \le n$, and index 1 at b_0 and b_{n+1}. The closure C of a component of $\Sigma - \Lambda$ has double along its boundary a surface T. The vector field naturally induces a vector field near ∂C in T whose only singularities (say there are p of them) are sources and sinks. This vector field can be extended to all of T, but must, by the Poincaré–Hopf index theorem, have in each copy of C singularities whose indices sum to $\chi(C) - p/2$.

If there is a component C for which there are no singularities in ∂C (i.e. $p = 0$) then ∂C is a cycle, so ∂C cannot come from doubling an arc in Λ. Hence C is contained in Λ, and so is a good cycle.

If $p > 0$ on each component C of $\Sigma - \Lambda$, then always $\chi(C) - p/2 \le 0$. In this case we have constructed on Σ a vector field for which every singularity has non-positive index, except b_0 and b_{n+1}, where the index is 1. Since $\chi(\Sigma) = 2$, we conclude that all inequalities must be equalities and hence, in particular, $\chi(C) - p/2 = 0$ for every component C of $\Sigma - \Lambda$ and so every component C is a disk and always $p = 2$.

Consider the component E of $D_Q - \Lambda$ which contains the segment of ∂D_Q lying between the label 1 in b_{n+1} and the label m in b_0 (see Fig. 5). From above, the double of E (along that segment) is a disk C, so E is also a disk. Since the vector field defined above has only two singularities along ∂C (in fact a source at b_{n+1} and a sink at b_0) the arc γ of ∂E lying in Λ is an oriented path with tail labelled 1 in b_{n+1} and head labelled m in b_0. (All labels of b_0 and b_{n+1} belong to edges in Λ, since there are no bad 0-edges or $(n+1)$-edges in Γ_P.).

Let e be the (bad) edge in Γ_Q whose end $*$ in b_δ is nearest the label 1 in b_{n+1}. (Such a bad edge must exist, for otherwise the heads and tails in γ would be adjacent, and the labels would be respectively always 1 and always $0 \ne m$.) The other end of e must lie in a vertex b_s lying on the oriented arc γ, since there are no loops based at b_δ. Finally, let F be the sub-disk

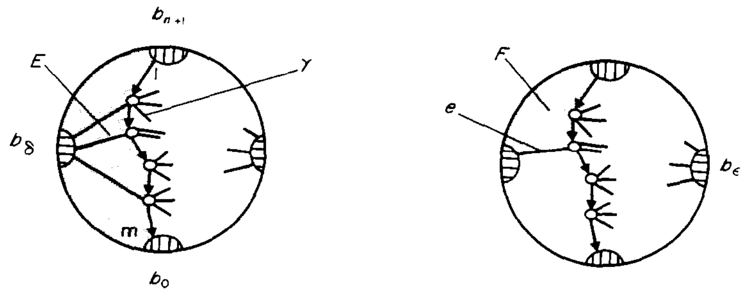


Fig. 5.

of E, cut off by e, whose boundary contains the initial segment of γ. Then F has no edges or vertices of Γ_Q in its interior. Now $*$ represents the point of intersection in the 1-handles of α with β located nearest the SE corner of the 3-ball. Push that point of intersection through the SE corner back onto the 3-ball so that it is in $\beta_+ \cap \alpha_-$. The effect is to change * to an end in β_+ labelled 0, and to change ∂F into a cycle. Then the disk F can be used as in [5, 4.8] to reduce m by one. ||

PROPOSITION 4.2. *Γ_P and Γ_Q each contain good unicycles, and sinks or sources.*

Proof. See [5, 5.2]. ||

PROPOSITION 4.3. *Neither Γ_P nor Γ_Q contains both a source and a sink.*

Proof. See [5, 5.4]. ||

CONVENTION 4.4. By renumbering, if necessary, we assume that both Γ_P and Γ_Q contain sources but no sinks.

§5. MORE TOPOLOGY: ORIENTATIONS AND NORMAL DIRECTIONS

Picture the 1-handles of N as being attached at points on the boundary of a disk $D_\perp$ in the 3-ball which is perpendicular both to D_v and D_h. The disk $D_\perp$ divides the boundary of the 3-ball from which N is constructed into two hemispheres and $\partial D_\perp \cap \alpha = \alpha_\pm$, $\partial D_\perp \cap \beta = \beta_\pm$.

Let ξ and ζ to be unit normal vector field to P and Q respectively in $M - N$, chosen so that on α_- and β_-, ξ and ζ both point into the same hemisphere into which $D_\perp$ divides the boundary of the 3-ball. We call the hemisphere the *front* of the 3-ball, and the other hemisphere (into which ξ and ζ point along α_+ and β_+) the back. The disk $D_\perp$ also divides each simple closed curve a_i and b_s ($i, s \geq 1$) into two arcs, called the front and back side of the curve, depending on the hemisphere that contains it. Our convention will be to picture $\alpha_\pm$ and $\beta_\pm$ as being pushed slightly into the front face of the 3-ball, allowing us to regard the vertices a_0, a_{m+1}, b_0, and b_{n+1} as always representing arcs on the front face. So, for example, the front face of a_0 is in fact all of a_0.

It will be convenient to have a shorthand picture which incorporates all the labelling of vertices a_i, $1 \leq i \leq m$, and of vertices b_s, $1 \leq s \leq n$, and also indicates their front face. The picture is shown in Fig. 6. The shaded region indicates the front, and the heavy dot represents the label $n+1$ in a_i or the label $m+1$ in b_s.

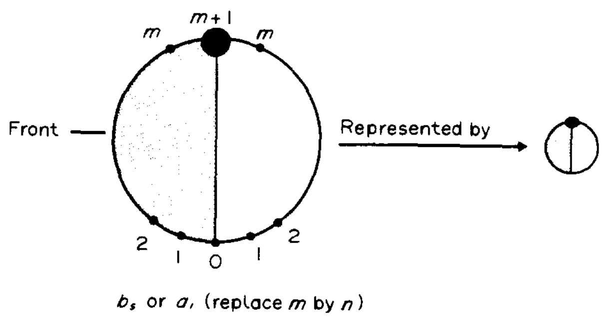


Fig. 6.

5.1 The components a_1 and a_m together bound a cylinder C in ∂N. Let η be the tangent unit vector field on C which points due south (always from a_{k+1} to a_k). Then for each $1 \le i \le m$ there is an $\varepsilon_i = \pm$ such that $\xi = \varepsilon_i \eta$ on a_i. Define $\varepsilon_0 = \varepsilon_{m+1} = -$. For $i, j \neq 0$, a_i and a_j are parallel if and only if $\varepsilon_i = \varepsilon_j$. Otherwise they are anti-parallel. Similarly define a sign δ_s for each $0 \le s \le n+1$ by comparing ζ to a westward pointing vector field υ at each b_s, $1 \le s \le n$ and setting $\delta_0 = \delta_{n+1} = -$.

5.2 Let e be an edge in Γ_Q with ends labelled i and j at vertices b_s and t_t respectively. The edge is said to be *synchronous* if either

(a) $1 \le i, j \le m$ and the labels $i+1$ at b_s and $j+1$ at b_t which are adjacent to e lie on the same side of e

(b) $i = 0\,(m+1)$, $1 \le j \le m$ and the label $j+1$ at b_t adjacent to e lies on the same (opposite) side of e as the front of b_s

(c) $i = 0$, $j = m+1$ (in which case the front of b_s and the front of b_t lie on opposite sides of e)

(d) $i = 0$, $j = 0$ (in which case the front of b_s and the front of b_t lie on the same side of e).

Otherwise the edge is *asynchronous*. Define synchronous and asynchronous edges in Γ_P similarly. (See Fig. 7.)

Following is a generalization of [5, 4.2] (all level edges are synchronous).

LEMMA 5.3. *Let e be an edge in* Γ_Q (Γ_P) *with ends labelled i and j (s and t). Then e is synchronous if and only if* $\varepsilon_i = \varepsilon_j (\delta_s = \delta_t)$.

Proof. For $1 \le i, j \le m$, $\varepsilon_i = \varepsilon_j$ if and only if η and ξ agree at both ends of e, or disagree at both ends of e. The argument for $i, j = 0$ or $m+1$ is similar. ||

5.4 Let γ be a good cycle in Γ_P (Γ_Q) such that all edges are asynchronous, at most one of 0 or $n+1$ $(m+1)$ is an edge label, and neither is an interior label. Let $v_1, \dots, v_k$ be the vertices of γ and η_i and τ_i denote the labels of the head and tail of γ at v_i respectively.

LEMMA. *For each i and interior label l of* γ *at* v_i, $\eta_i < l < \tau_i$.

Proof. Since there are no interior labels 0 or $n+1\,(m+1)$ it follows from Fig. 2 that all interior labels of γ at v_i lie between η_i and τ_i. If there are no edge labels 0 or $n+1\,(m+1)$ then since the edges are asynchronous either always $\eta_i < \tau_i$ or always $\eta_i > \tau_i$. Since γ is a cycle, all edges are oriented so $\eta_{i+1} < \tau_i$ and $\eta_1 < \tau_k$. But clearly we cannot have $\eta_1 > \tau_1 > \eta_2 > \dots > \tau_k > \eta_1$.

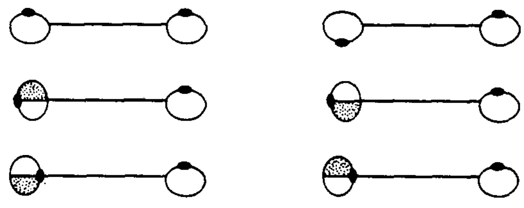


Fig. 7.

If there is an edge label 0, but not $n+1(m+1)$ then apply the same argument between successive appearances of 0. Similarly if $n+1(m+1)$ is an edge label, but 0 is not. ||

LEMMA 5.5. *Suppose e is an edge in Γ_Q with one end on the front side of a vertex b_s and the other on the front side of a vertex b_t. Then e is asynchronous if and only if $\delta_s = \delta_t$.*
Similarly for an edge in Γ_P.

Proof. Let i be the label of e on b_s and j the label on b_t. The normal bundle to e in Γ_Q and the normal bundle to the corresponding edge in Γ_P together give a normal bundle in $M-N$ for the arc of intersection of P with Q corresponding to e. Since this arc must join points of $\partial P \cap \partial Q$ of opposite sign, and the ordered pair of vector fields η, υ used in the definition of ε_i and δ_s everywhere defines the same orientation on the front face, $\varepsilon_i \delta_s \neq \varepsilon_j \delta_t$. The result follows from 5.3. ||

5.6 Without loss, choose the front side (hence ξ and ζ) so that $\varepsilon_1 = +$. WARNING: Since δ_1 may not be $+$, the roles of Γ_P and Γ_Q may no longer easily be reversed in the statements of theorems.

5.7 Following is a generalization of [5, 4.8].

LEMMA. *A good semi-cycle in Γ_Q has either*

(i) *a level spoke labelled 0*
(ii) *a level edge labelled 0 with its back side toward the interior*
(iii) *an edge or interior label $m+1$ or*
(iv) *an interior vertex.*

Proof. If not, then within it is a semi-cycle γ with neither interior vertices nor interior edges, with all edges labelled l and $l+1$, for some $l<m$, and with the front side of all its 0-labelled level edges toward the interior [5, 4.7]. If $l \neq 0$ then proceed as in [5, 4.7] to reduce m by two. If $l=0$ and the interior of γ is incident only to the front side of its vertices, or if γ is a cycle then proceed as in [5, 4.8] to reduce m by one.

So suppose there is at least one level edge of γ labelled 0 and that the interior of γ is incident to the back side of at least one vertex. Let b_t be the first such vertex in γ following such a level edge. Let e be the edge preceding b_t and f the edge succeeding b_t in γ. (See Fig. 8.) Since the interior of γ lies on the front side of all level edges labelled 0, neither e nor f can be such an edge. Hence e cannot be level with label 1 either, for the oriented edge f cannot have its head at b_t. Thus e is oriented with its head at b_t labelled 0. Our choice of b_t guarantees that the tail of e is labelled 1 at the preceding vertex b_s. But then e is synchronous (5.1(b)), yet $\varepsilon_1 \neq \varepsilon_0$ (5.6), contradicting 5.3. ||

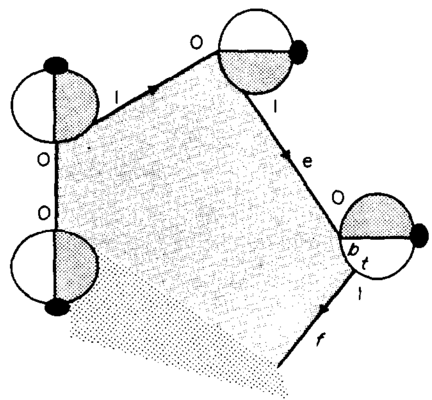


Fig. 8.

§6. SPECIAL PATHS IN Γ_Q

DEFINITION 6.1. A good cycle in Γ_P is *coherent* if each of its vertices has subscript smaller than that of each vertex in its interior. A coherent loop based at a_i is a *lobe* if the interior of the loop is incident only to the front side of a_i (See Fig. 9). A level edge in Γ_Q is coherent (*lobal*) if the corresponding loop in Γ_P is coherent (a lobe). Note that any level edge in Γ_Q labelled 0 is lobal.

6.2 Since Γ_Q contains sources (4.4), any innermost good loop in Γ_P is coherent [5, 5.4]. Let r be the highest subscript of all vertices in Γ_P which are contained in the interiors of innermost coherent cycles.

LEMMA. (a) *Any vertex of a coherent cycle in Γ_P has subscript less than r.*

(b) *Any edge incident to a_r is incident to some a_i with $i<r$.*
(c) *In Γ_Q any label r is the tail of an oriented edge.*
(d) *A level edge in Γ_Q with label greater than r is not coherent.*

Proof. (a) Contained within the coherent cycle is an innermost one. hence a vertex with subscript $\leq r$.
(b) If the edge were a loop, it would contain an innermost good loop, hence a yet further in coherent cycle.
(c) Follows from b.
(d) The level edge corresponds to a loop in Γ_P, containing an innermost good loop, hence an innermost coherent cycle, hence a vertex with subscript $\leq r$. ||

DEFINITION 6.3. The edge f is *larger* than the edge e in Γ_Q if

(i) e and f are incident to the front of the same vertex b_s at labels i and j respectively, with $i<j<r$.
(ii) neither e nor f is oriented with its tail at b_s
(iii) if e or f is level, it is coherent and
(iv) the label of the end of f not incident to b_s is also greater than the label of the corresponding end of e. If e or f is lobal, the end of f not incident to b_s need only be $\geq$ the corresponding end of e.

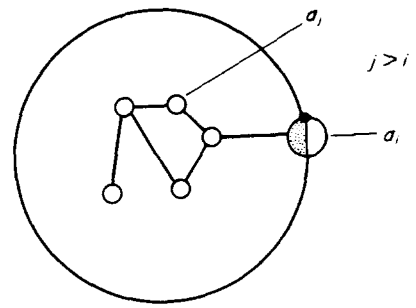


Fig. 9. (A lobe in Γ_P)

DEFINITION 6.4. A label $i \geq 1$ is *special* if any oriented edge in Γ_Q with tail labelled i is asynchronous.

DEFINITION 6.5. A *special path* is a possibly infinite, possibly non-imbedded path μ in Γ_Q through oriented edges $e_1, e_2, e_3, \ldots$ with tails labelled $\tau_1, \tau_2, \tau_3, \ldots$ respectively, such that for all $i \geq 1$:

(a) τ_i is special

(b) $0 < \tau_i \leq \tau_{i+1} < r$

(c) each end of each e_i lies on the front side of the vertex b_s to which it is incident.

(d) there is a larger edge than e_k incident to the vertex v at the head of e_k if and only if e_k is the last edge in the path. In this case the vertex v is called the *terminus* of μ, such a larger edge also incident to v is called a *terminator* of μ, and e_k is the *final edge* of μ.

DEFINITION 6.6. A *special cycle* is a good cycle in Γ_Q of oriented edges $e_1, \ldots, e_k$ such that the infinite path $e'_1, e'_2, e'_3, \ldots$ defined by $e'_{(pk+i)} = e_i (p \geq 0,\ 1 \leq i \leq k)$ is a special path.

DEFINITION 6.7. An edge in Γ_Q is special if it belongs to some special path. An edge e with tail at a vertex v is *extremal* if it is special and, of all special edges with tail at v, the tail of e has highest label. An extremal path (cycle) is a special path (cycle) in which every edge is extremal. See Fig. 10.

LEMMA 6.8. (1) *If any initial segment of a special path is deleted, the path is still special.*

(2) *Suppose $e_1, e_2, e_3, \ldots$ is a special path μ. Then either an edge e_k is the final edge or there is a special path $e'_1, e'_2, e'_3, \ldots$ such that $e_i = e'_i$ for $i < k$ and e'_{k+1} is extremal.*

(3) *If a vertex is the tail of a special edge, then it is the initial vertex of some extremal path.*

Proof. (1) Follows from the definition 6.5.

(2) Let v be the vertex at the head of e_k. Either e_k is the final edge of μ or a special edge e_{k+1} has tail at v. In the latter case, let e'_{k+1} be that special edge with tail at v whose label τ'_{k+1} there is highest. Then e'_{k+1} is extremal and τ'_{k+1} is no smaller than τ_{k+1}, the label of the tail of e_{k+1}. Since e'_{k+1} is special, it follows from (1) that it is the first edge of a special path e'_{k+1}, e'_{k+2}, $e'_{k+3}, \ldots$. Then $e_1, \ldots, e_k, e'_{k+1}, \ldots$ satisfies (a), (c) and (d) of 6.5 trivially and satisfies (b) since $\tau'_{k+1} \geq \tau_{k+1}$.

(3) Since the vertex is the tail of a special edge, it is the tail of an extremal edge e_1 which, by (1) is the initial segment of a special path. Suppose inductively that the initial segment $e_1, \ldots, e_k$ of some special path μ beginning with e_1 has all of its edges extremal. If e_k is the final edge of μ we are done. Otherwise there is a special edge, hence an extremal edge e_{k+1},

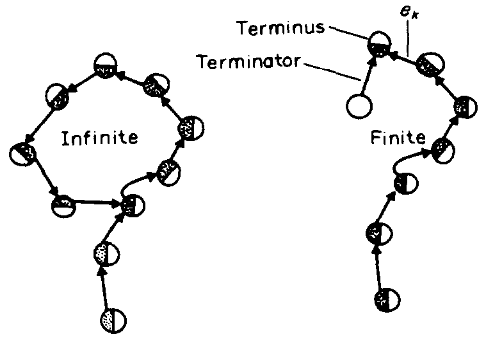


Fig. 10. (Exremal paths)

with its tail at the head of e_k. By (2) $e_1, \ldots, e_k, e_{k+1}$ is the initial segment of some special path. Continue until either reaching a final edge and hence an entire extremal path or some vertex is encountered for the second time, forming a cycle. Then the path obtained by continuing around the cycle *ad infinitum* is also extremal. ||

LEMMA 6.9. *Suppose e is a lobal edge in Γ_Q incident to vertices b_s and b_t. Then, either there is a larger edge incident to the front of b_s or b_t or there are disjoint special paths which originate at b_s and b_t.*

Proof. Suppose there is no larger edge incident to the front of b_s or b_t. With no loss of generality $s>t$.

Define an $(s-t)$ cycle γ (in Γ_P) to be a coherent cycle in which

(a) all the interior labels of γ lie between s and t.
(b) every edge is asynchronous
(c) the interior of γ is incident only to the front side of its vertices.

Since e (regarded here as a loop in Γ_P, based at a vertex a_l) is a lobe it is also an $(s-t)$ cycle.

Claim. Suppose γ is an $(s-t)$ cycle in the interior of the loop e, and i is the smallest subscript of all vertices in the interior of γ. Then the edges f_s and f_t incident to the front of a_i at the labels s and t respectively divide γ into two circuits, one of which is a coherent cycle γ'. Moreover, if $\varepsilon_i=\varepsilon_j$ for all subscripts j of vertices of γ, then γ' is an $(s-t)$ cycle.

Proof of claim. Since e is coherent, any vertex in γ or its interior has subscript no smaller than l. In particular, f_s and f_t, when viewed as edges in Γ_Q satisfy (i), (iii) and (iv) of 6.3 but by hypothesis are not larger than e, so cannot satisfy (ii). Thus the ends of f_s and f_t not incident to a_i are incident to vertices a_j with $j<i$, hence vertices of γ. (See Fig. 11).

Since all interior labels of γ lie between s and t, f_t has its head at a_i and f_s its tail. Thus f_s and f_t divide γ into two circuits, one of which is a coherent cycle. If, furthermore, $\varepsilon_i=\varepsilon_j$ for all subscripts j of vertices of γ, then by 5.5 every edge of γ, as well as f_s and f_t, are asynchronous. It follows from 5.4 that the head of f_s has label lower than the tail of the next edge of γ', so, by 5.2, the interior of γ' is incident also to a_i only on its front face, between the labels s and t. Hence γ' is an $(s-t)$ cycle.

Proof of Lemma 6.9 (*Continued*). Since any good cycle (hence (*s*–*t*) cycle) must contain interior vertices [3.3], it follows from the claim (used repeatedly) that there must be some

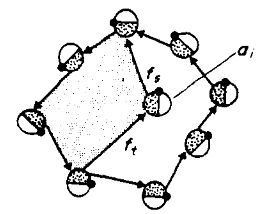


Fig. 11.

$(s-t)$ cycle γ within the loop e in Γ_P for which the interior vertex with smallest label, a_i, has $\varepsilon_i \neq \varepsilon_j$ for *every* a_j in γ. Then the label i is special (5.3 and 6.4). Let $s_0 = s$ and e_1 be the edge in Γ_P incident to the label s_0 on the front of a_i. (The edge e_1 was called f_s in the above proof of claim.) Let s_1 be the label of the other end of e_1 which, as in the claim, lies on a vertex of γ (so $t < s_1 < s$). Let e_2 be the edge incident to the front of a_i at label s_1. If the other end of e_2 is incident to an interior label of γ, denote it by s_2 $(t < s_2 < s)$ and let e_3 be the edge incident to the front of a_i at label s_2. (See Fig. 12).

Continue in this manner indefinitely, stopping only if

(i) some e_{k+1} has both ends incident to a_i (hence both to the front of a_i, since f_s and f_t separate the front labels between s and t from all the others).

(ii) some e_{k+1} has its other end at a vertex a_j with $j > i$ (so a_j is not a_i and not in γ).

This series of edges $e_1, e_2, e_3, \ldots$, when viewed in Γ_Q is a path beginning at b_s. The tails are always labelled i, so (a) and (b) of 6.5 are always satisfied. (c) is satisfied by construction. Finally, conditions (i) and (ii) imply that e_{k+1} is larger (in Γ_Q) than e_k. Thus the longest initial segment of the path $e_1, e_2, e_3, \ldots$ in Γ_Q that satisfies (d) of 6.5 is a special path.

The construction of the special path originating at b_t is similar.

All vertices of a special path have δ the same sign (6.5(a), (c) and 5.5). Since level edges are synchronous, and e is lobal, $\delta_s \neq \delta_t$ (via 5.5). Hence the paths must be disjoint. ||

§7. MULTIFLOWS

DEFINITION 7.1. A *semi-oriented path* is an imbedded path with a given orientation which is never inconsistent with any of the orientations of its edges. For α a semi-oriented path, let $\tau(\alpha)$ denote the initial vertex and $\eta(\alpha)$ denote the terminal vertex.

7.2 Let e be a lobal (or special) edge, with label (of its head) i at a vertex b_s. Define the *large* side of e to be the side on which the adjacent label $i+1$ of b_s lies. If $i = 0$, the large side will be that on which the front of b_s lies. Note that if an edge f has tail labelled r and its head on the front of b_s and on the large side of an edge e with head also on the front of b_s, then f is larger than e (and so is a terminator if e is special).

DEFINITION 7.3. A *multiflow* in Γ_Q is a good circuit which is the union of

(a) a collection $\{\sigma_1, \ldots, \sigma_k\}$ of level edges and/or vertices, called the *apexes*

(b) $2k$ semi-oriented paths $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_k$ called the *currents* such that for all $1 \leq \lambda \leq k$

(i) $\tau(\alpha_\lambda)$ and $\tau(\beta_\lambda)$ are vertices incident to (or the vertex) σ_λ.

(ii) $\eta(\alpha_\lambda) = \eta(\beta_{\lambda+1})$ and $\eta(\alpha_k) = \eta(\beta_1)$; these vertices are called the *bases* of the multiflow.

By convention, an imbedded semi-cycle is regarded as a multiflow with $k = 0$.

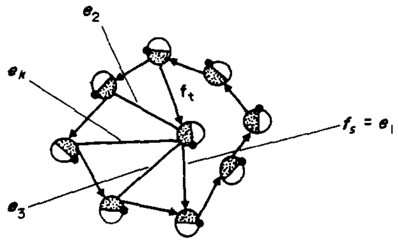


Fig. 12.

We will be examining restricted families of multiflows satisfying certain conditions on the apexes, currents, and bases. A semi-cycle will be regarded as being in the family (with $k=0$) if it satisfies the conditions on currents.

The argument proceeds as follows: First we define a type of multiflow and show that examples exist (assuming, as we have since 3.1, that $m, n \geq 1$). Then we discover properties about the interior of an innermost such multiflow γ. The quest for an innermost multiflow requires dividing multiflows by semi-oriented paths and showing that one of the two resulting multiflows still satisfies the apex, current, and base conditions.

A new class of multiflows is then defined, of which γ is an example and for which the apex, current and base restrictions are liberalized. The process is repeated on this new class.

After three such stages we arrive at a multiflow whose properties are contradictory. The contradiction (to $m, n \geq 1$) completes the proof.

DEFINITION 7.4. A type A multiflow γ is one in which

(a) All apexes σ_λ are lobal edges.

(b) If there is no larger edge than σ_λ incident to *either* end of σ_λ then each end is the origin of an extremal path $\mu \subset \gamma$.

(c) Any edge is either
  (i) lobal
  (ii) oriented with tail labelled r or
  (iii) in an extremal path $\mu \subset \gamma$ which originates at a lobal edge.

(d) The interior of γ lies on the large side of each lobal edge and each extremal edge in γ.

(e) On each vertex of a lobal edge or an extremal edge in γ, the interior of γ is incident only to the front side and the interior labels are less than r.

LEMMA 7.5. *For a type A multiflow γ:*

(i) *At least one edge of γ at each base has tail labelled r.*

(ii) *If f is an edge with tail labelled r which lies on or in the interior of γ, then f is incident to an extremal path $\mu \subset \gamma$ only at its terminus.*

(iii) *If the initial edge of an arbitrary extremal path ρ lies on or in the interior of γ, so does each edge of ρ.*

Proof. (i) Follows from 7.4(c) and (d).

(ii) If the head of f is on the extremal path this follows from 7.4 (d) and (e) and 7.2. If the tail of f (labelled r) is on the extremal path this follows from 6.5(b) and the fact that μ lies entirely in γ.

(iii) Suppose f were the first edge of ρ not to lie on or in the interior of γ. Then the tail of f lies on the front side (6.5 (c)) of a vertex b_s of γ at a label τ. Since the head of the predecessor f'

of f in ρ is also incident to the front of ρ (6.5(c)) at a label $\eta<\tau<r$ (6.5 (b)) some edge e of γ must be incident to a label between η and τ. The edge e cannot be a lobal edge or the head of an extremal edge by 7.4 (d). It cannot be the tail of an extremal edge, since f is also extremal with tail at b_s. It cannot have tail labelled r at b_s since $\tau<r$. It cannot have tail labelled r at another vertex by 7.2. This contradicts 7.4 (c) for the edge e. ||

PROPOSITION 7.6. *Γ_Q contains a type A multiflow. An innermost type A multiflow γ contains in its interior*

(a) *no extremal cycles, nor circuit in which every vertex has an interior or edge label r*
(b) *no interior labels r*
(c) *no lobal edges*
(d) *no sources.*

Proof. Via 6.2(c) there is an imbedded good cycle in Γ_Q with all tails labelled r, hence a type A multiflow (with $k=0$).

(a) A circuit in which every vertex has an interior or edge label r has in its interior a cycle with every tail labelled r (via 6.2(c)), a type A multiflow. If γ' is an extremal cycle then the front side of each edge is either always on the inside or always on the outside of γ', by 6.5(a), (b), (c). In the former case γ' is a type A multiflow (with $k=0$) and in the latter case every edge has an interior label r.

(b) Let γ be an innermost type A multiflow. If there is an interior label r, then by 6.2(c) and 7.6(a) there is an oriented imbedded path α in the interior of γ with both ends on γ and each tail of each edge of α labelled r. Then α divides γ into two multiflows, one of which γ' still satisfies 7.4(a) because it contains no new apexes. $\tau(\alpha)$ is not at an apex of γ by 7.4(e), so 7.5 (ii) shows γ' still satisfies 7.4(b). Finally, 7.4(c) is true for γ' by construction and 7.4(d) and (e) by default.

(c) Suppose f is a lobal edge in the interior of γ.

Case i. The vertices b_s and b_t at the ends of f lie in the interior of γ and there is a larger edge than f incident to b_s or b_t.

In this case construct oriented paths α' and β' with all tails labelled r, beginning at the labels r on the front side of b_s and b_t. By 7.6(a) the paths are imbedded, so eventually reach γ. Although they may not be disjoint, they can be made never to cross (indeed never to share an edge, though this fact is not used). Each complementary region of $\alpha\cup\beta$ in the interior of γ is bounded by a multiflow; let γ' be that multiflow which contains the large side of f in its interior. Then the apex f of γ' satisfies 7.4(a), (d) and (e). The proof in this case now follows as in 7.6(b).

Case ii. The vertices b_s and b_t lie in the interior of γ and there is no larger edge than f incident to b_s or b_t.

They by 6.9 and 6.8(c) there are disjoint extremal paths μ and υ in the interior of γ originating at b_s and b_t. (See Fig. 13.) By 7.5(iii) they lie entirely in γ' or its interior. If the terminus b_p of μ lies in the interior of γ, extend μ to a path μ', by always exiting at a label r, beginning at the front label r of b_p. If this extension flows into a vertex b_λ of μ (resp. υ) on the large side of the edge of $\mu(\upsilon)$ that flows into b_λ, then it terminates $\mu(\upsilon)$. Thus μ' is imbedded, for the interior of any circuit would be either a type A multiflow or would contradict 7.6(a), and μ' doesn't cross υ. Similarly extend υ to an imbedded path υ' which doesn't cross μ. By the appropriate choice of labels r used in their construction we may assure then that μ' and υ'

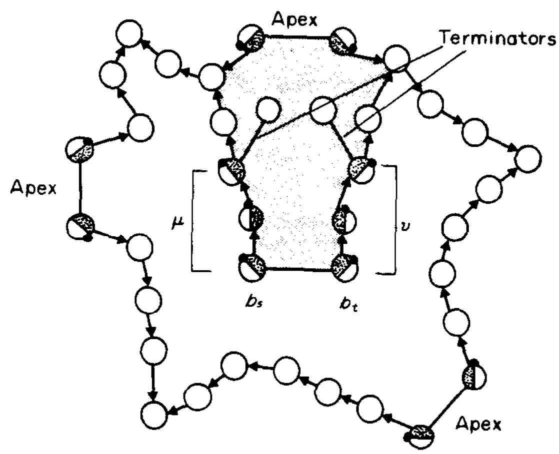


Fig. 13.

do not cross anywhere. Each complementary region of $\mu' \cup \upsilon'$ in the interior of γ is bounded by a multiflow. As in case *i*, let γ' be the boundary of the complementary region that contains the large side of f. The proof in this case now follows as in 7.6(b).

Case iii. b_s lies in γ, but b_t does not.

Proceed as in case ii if there is no larger edge than f incident to b_s or b_t. If there is a larger edge, proceed as in case (i) to construct an oriented path β' beginning at the label r on the front of b_t, and ending on γ. The union of f and β' divides γ into two multiflows; let γ' be that which contains the large side of f.

Let e_s be the edge of γ' incident to b_s. By 7.4(d) e_s cannot be lobal nor, if it is extremal, can its head be at b_s. Thus e_s is either extremal with tail at b_s or has tail labelled r at b_s or has tail labelled r, but head at b_s. In the first two cases f is an apex of γ' and the proof proceeds as in case i. In the last case, when e_s has tail labelled r, but head at b_s, then γ' is a type A multiflow in which f is part of a current satisfying 7.4(c). This current still satisfies 7.4(d) and (e) because there is no interior label r at b_s (7.6(b)).

Case iv. b_s and b_t both belong to γ.

If there is no larger edge than f incident to b_s or b_t proceed as in case ii. Suppose there is a larger edge than f incident to b_s or b_t. The edge divides γ into two multiflows; let γ' be that whose interior contains the large side of f. If either e_s or e_t has tail labelled r and head at b_s (or b_t) then γ' is a type A multiflow in which f is part of a current satisfying 7.4(c). The only other possibility, as in case iii, is that each e_s and e_t either is extremal with tail at f or has tail labelled r at f. Then γ' is a type A multiflow with f a new apex.

(d) Let b_s be a source in the interior of γ. Then the edge labelled 0 must be level, hence lobal. This contradicts (c). ||

§8. TYPE B MULTIFLOWS

DEFINITION 8.1. *A type B multiflow is a multiflow γ, contained in an innermost type A multiflow, for which:*

(a) All apexes are either
   (i) lobal edges

(ii) vertices to which an interior edge is incident, either coherent and level or with head at the apex

(iii) the initial vertex for an extremal path $\mu \subset \gamma$.

(b) Let f be a lobal edge which is an apex. If there is no larger edge incident to either end of f, then each end is the origin of an extremal path $\mu \subset \gamma$. If there is a larger edge incident to an end of f, at least one such edge lies in the interior of γ.

(c) Each edge is either lobal or oriented, with labels $\leq r$.

(d) The interior of γ lies on the large side of each lobal edge and each edge in an extremal path μ.

(e) At each vertex of an apex, a lobal edge, or an extremal path μ, the interior of γ is incident only to the front side, and the interior labels are less than r.

(f) A terminator for each extremal path μ is in γ or in its interior.

PROPOSITION 8.2. *An innermost type A multiflow γ_A is type B. An innermost type B multiflow has no*

(a) *interior label $\geq r$*
(b) *vertex in its interior*
(c) *interior oriented terminator of an extremal path μ in γ.*
(d) *interior oriented edge with head at an apex.*

Proof. γ_A automatically satisfies (a)–(e). To verify f, let b_s be the terminator of an extremal path in γ_A. The edge e of γ incident to b_s but not in μ cannot be extremal or lobal, by 7.4(d) and so has tail labelled r. If the tail of e is on b_s, then a terminator, which has head at b_s with label $<r$, must be in the interior of γ. If the head of e is on b_s then e itself is a terminator.

(a) First note that it suffices to show that there are no interior labels r. For it then follows from 8.1(c) that at any vertex b_s on which there is an interior label greater than r, the edges of γ incident to b_s both would have label r there. Thus b_s would be an apex, contradicting 8.1(e).

If there is an interior label r at a vertex b_s of γ, then construct an imbedded path α in the interior of γ such that $\tau(\alpha)=b_s$, $\eta(\alpha)$ is a vertex b_η of γ, and each edge of α is oriented with tail labelled r. Denote by e the first edge of α and by f the last edge. (Perhaps $e=f$.)

Then α divides γ into two multiflows, both of which continue to satisfy 8.1(f) at b_η when b_η is in an extremal path, since f would be a terminator.

If b_s is an extremal path μ of γ, but is not its terminus, then let γ' be the multiflow containing the terminal segment of μ. The vertex b_s is the only apex of γ' which is not an apex for γ, and it satisfies 8.1(a)(iii). By construction γ' still satisfies 8.1(b) and (c) and satisfies 8.1(d) and (e) because γ did. (See Fig. 14).

If b_s is the terminus of an extremal path μ of γ, choose γ' so that a terminator for μ is on γ' or in its interior (8.1(f)). If γ' contains μ then it has no new apex and continues to satisfy 8.1(f) by construction, and 8.1(a)–(e) because γ did. If γ' does not contain μ then b_s is a new apex of type 8.1(a)(ii) still satisfying 8.1(e) since b_s was on an extremal path. 8.1(b)–(d) and 8.1(f) are true for γ' since they were for γ.

If b_s is not on an extremal path, nor at an apex, choose γ' so that it contains the initial segment of the current on which it lies (if it is at a base, choose either current). Then γ' contains no new apex, and satisfies 8.1 because γ did.

If b_s is in an apex, but not an extremal path, it is in an apex of type 8.1(a)(i) or (ii). Furthermore, if it is in an apex of type 8.1(a)(i), then it is a vertex of a lobal edge to which a larger edge is incident. Choose γ' so it continues to contain this larger edge in its interior. Then γ' has an apex at b_s either of type 8.1(a)(i) satisfying 8.1(b) or of type 8.1(a)(ii).

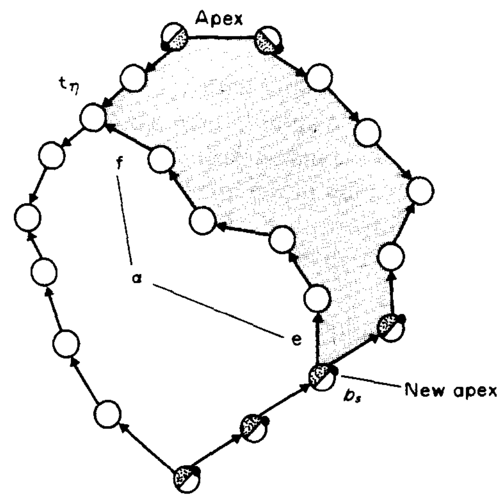


Fig. 14.

If b_s is an apex of type 8.1(a)(ii) then choose γ' so it continues to contain the interior edge, incident to b_s, which is either coherent level or has head at b_s (and is thus not e itself). Then b_s is still an apex in γ' of type 8.1(a)(ii).

(b) There are no sources in the interior, hence no cycles, by 7.6(d). Hence if there is an interior vertex there is an oriented edge e with its tail at a vertex b_s in γ and its head at an interior vertex b_t. Beginning with a label r on b_t, extend e to an imbedded oriented path α in the interior of γ such that $\tau(\alpha)=b_s$, $\eta(\alpha)$ is a vertex b_η in γ, and every edge but e has tail labelled r. The proof is now a word for word repeat of (a). (Only the fact that the last edge f of α has tail labelled r was used.)

(c) Suppose f is an oriented terminator of an extremal path μ in γ with head at b_η in μ and tail at b_s. By 8.2(b) we can take b_s to be in γ, so f divides γ into two multiflows. Both of them continue to satisfy 8.1(f) at b_η since f is a terminator of μ. The proof is now a word for word repeat of 8.2(a) (That the last edge f of the path α has tail labelled r was used there only to show that 8.1(f) was satisfied at b_η).

(d) Suppose f is an oriented edge with its head at a vertex b_η in an apex and its tail at b_s. By 8.2(a) we can take b_s to be in γ, so f divides γ into two multiflows, neither of which any longer has an apex containing b_η. The proof is now a word for word repeat of 8.2(a) (Since b_η is an apex of γ it is not a terminus). ||

§9. TYPE C MULTIFLOWS

Definition 9.1. A type C multiflow γ is a multiflow, contained in an innermost type B multiflow, in which:

(a) All apexes are either

(i) vertices to which an interior coherent level edge is incident.

(ii) the initial vertex for an extremal path $\mu \subset \gamma$.

(b) Each current is a semi-oriented path in which any non-lobal level edge has some predecessor which is oriented.

(c) The interior of γ lies on the large side of each lobal edge and each edge of an extremal path μ.

(d) On each vertex of a lobal edge, apex, or extremal path in γ, the interior of γ is incident only to the front side.

(e) The terminus of each extremal path is either a base of γ with terminator on γ, or there is a coherent level terminator in the interior of γ.

LEMMA 9.2. *An innermost type B multiflow γ is type C.*

Proof. Suppose f is an apex of type 8.1(a)(i). If there is a larger edge than f incident to one of its ends, b_s, then it is coherent level by 8.2(d) so b_s can be viewed as an apex of type 9.1(a)(i) with f the first (lobal) edge of a current. If there is no larger edge than f incident to either end of f, then by 8.1(b), b_s is similarly an apex of type 9.1(a)(ii).

Suppose a vertex b_s of γ is an apex of type 8.1(a)(ii). Then by 8.2(d) b_s is an apex of type 9.1(a)(i).

A vertex of type 8.1(a)(iii) is already of type 9.1(a)(ii).

Thus 9.1(a) is true for γ.

9.1(b) is true by default (8.1(c)).

9.1(c) is 8.1(d).

9.1(d) is 8.1(c).

9.1(e) follows from 8.1(f) and 8.2(c), since an oriented terminator must have its head at the terminus.

9.1(f) is 8.2(e). ||

PROPOSITION 9.3. *For an innermost type C multiflow γ:*

(a) *All apexes are of type 9.1(a)(ii).*

(b) *The terminus of each extremal path in γ is a base at which the other incident edge of γ is a terminator.*

(c) *There is at least one apex.*

Proof. (a) Suppose b_η is a type 9.1(a)(i) apex with f an interior coherent level edge incident to b_η. By 9.1(d) f is incident to the front of b_η. By 8.2(b) the other end of f is a vertex b_s of γ. Furthermore, since f is not lobal (7.6(c)), f cannot be incident to the front of b_s, so b_s is not an apex and is not on a lobal edge or extremal path (9.1(d)). (See Fig. 15).

The edge f divides γ into two multiflows; choose γ' to be that containing the initial segment on which b_s lies, and let e be the edge of γ' incident to b_s. Since e is not lobal, either it

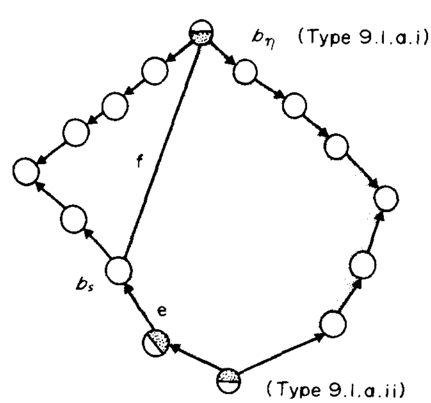


Fig. 15.

or a predecessor in the current is oriented (9.1(b)). In the multiflow γ' we then regard f as lying on the same current as e, so b_η is no longer an apex. Then γ' continues to satisfy 9.1(a), (c), and (d) since γ did. It satisfies 9.1(b) by construction and 9.1(e) since neither end of γ is a terminus for an extremal path.

(b) Let b_η be the terminus of an extremal path μ in γ. If b_η is not a base, or if a terminator does not lie in γ, then by 9.1(e) there is a coherent level terminator f in the interior of γ. As in (a), the other end of f is a vertex b_s in γ which is not an apex, not a lobal edge, and not on an extremal path of γ.

The edge f divides γ into two multiflows; choose γ' to be that containing the initial segment of the current on which b_s lies. Then 9.1(a)–(d) are true for γ' just as in 9.3(a). Finally, 9.1.3 is true for γ' because, if μ lies in γ' then b_η is a base and f a terminator.

(c) By 9.1(c), adjacent edges of γ cannot be lobal, so by 9.1(b) there is an oriented edge in γ. If γ has no apex, it is a semi-cycle. This contradicts 5.7 (via 7.6(c), 9.1(d), 8.2(a), 8.2(b)).

9.4. From 9.3(a) we see that each apex of an innermost type C multiflow γ is the initial vertex of an extremal path, and from 9.3(b) that the extremal path is the entire current. From 9.1(c) we know that two extremal paths cannot meet at a base. We conclude that each base is the terminus of exactly one extremal path in γ. Hence we can assume that in the multiflow γ the currents α_λ are all extremal paths. In view of this, we will henceforth denote them μ_λ.

LEMMA 9.5. *For f an interior edge of γ an innermost type C multiflow:*

(a) *If f is oriented, the head lies in an extremal path μ_λ.*
(b) *If f is level, at least one end lies on a vertex in an extremal path μ_λ other than the base.*

Proof. (a) Suppose f is an oriented edge with its head at a vertex b_t in a current β_λ, where b_t is not an apex or base. The tail of f is a vertex b_s of γ (8.2(b)), so f divides γ into multiflows. Let γ' be the one which contains the next edge of β_λ.

If b_s lies on a vertex of some β_υ then it follows from the definition of a multiflow that the terminal segment of β_λ and the initial segment of β_υ lie on the same side of f, and hence are both in γ'. (If $\lambda=\upsilon$ then γ' is f together with the segment of β_λ which lies between b_t and b_s.) It is easy to see that γ' is still a type C multiflow in which the initial segment of β_υ and the terminal segment of β_λ have been joined into a single current by f.

If b_s lies on a vertex of an extremal path μ_υ then it follows from the definition of a multiflow that the terminal segments of β_λ and μ_υ both lie on the same side of f, hence in γ'. It is easy to check that γ' is still a type C multiflow with new apex at b_s.

(b) $\{\mu_\lambda\}$ and $\{\beta_\lambda\}$ intersect only on the apexes and bases. Hence b_s is a vertex in $\{\mu_\lambda\}$ other than a base if and only if it is not a vertex of $\{\beta_\lambda\}$ except perhaps an apex. Suppose on the contrary, then, that a level edge f in the interior of γ has ends at b_t in β_λ and b_s in β_υ but neither vertex is an apex.

The immediate predecessors e_t for b_t in β_λ and e_s for b_s in β_υ lie on opposite sides of f, by 7.3. Since level edges are synchronous, 9.1(c) and 9.1(d) imply that e_s and e_t can't both be lobal. Similarly b_s and b_t can't both be bases (9.1(c)), and if b_s (b_t) is a base then e_s (e_t) is not lobal (9.1(c)).

So take e_s to be a non-lobal edge, with b_s a base if either b_s or b_t is. Then b_t is not a base and some predecessor of e_s is oriented (9.1(b)). Let γ' be the multiflow lying on the same side of f as e_s. Then γ' is a type C multiflow in which the initial segment of β_υ and the terminal segment of β_λ have been joined into a single current by f. ||

Let q be the largest label of any tail appearing in any of the $\{\mu_1, \dots, \mu_k\}$ in γ, say in μ_k.

LEMMA 9.6. (a) *The first edge in β_1 has both ends labelled less than q.*

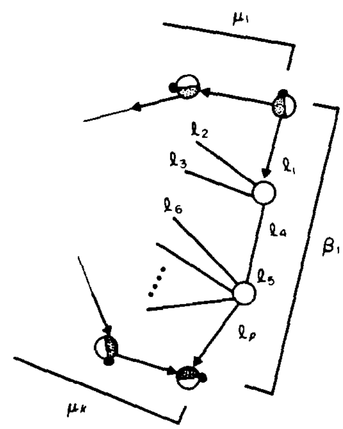


Fig. 16.

(b) *The last edge in* β_1 *has tail (or both ends, if level) greater than* q.

(c) *Any interior label of* γ *on a vertex* b_s *in* $\{\mu_\lambda\}$ *other than a base is less than* q.

Proof. (a) The first edge in μ_1 has tail with special label $\leq q$ at the apex σ_1 by 6.5(a). Then the edge is asynchronous, so the result follows from 9.1(c) and (d).

(b) By 6.5(b) the last edge in μ_k has tail labelled q. Since the last edge in β_1 can't be lobal (9.1(c)) the result follows from 9.3(b).

(c) This follows from 6.5(b) and (c). ||

Let $l_1, \ldots, l_p$ be the edge and interior labels of β_1 read in order along β_1 beginning with the label $l_1 < q$ of 9.6(a) and ending with the label $l_p > q$ of 9.6(b). (See Fig. 16).

LEMMA 9.7. *Let* l_j *be the last label for which* $l_j < q$. *Then* $l_{j+1} = q$ *and* l_{j+1} *is an interior label of* γ.

Proof. Since any oriented edge of β_1 points from an l_i to l_{i+1}, and $l_{j+1} \geq q > l_j$, l_j and l_{j+1} cannot be the ends of an edge lying in β_1. Therefore l_{j+1} and l_j are adjacent on a common vertex. Since $l_{j+1} \geq q$, it follows that $l_{j+1} = q$ and so $j+1 < p$. Similarly, if l_{j+1} and l_{j+2} are the ends of an edge in β_1, it must be a level edge. But this, too, is impossible since a level edge is synchronous, forcing $l_{j+3} = l_j < q$ and contradicting the choice of j. Hence l_{j+1} is an interior label. ||

9.8 *Proof* of 1.3. Following 3.1 we have assumed $m, n \geq 1$. Here is a contradiction: Let f be the edge in the interior of an innermost type C multiflow γ with end the label l_{j+1} of 9.7. Then the label is q and lies on a vertex b_s not in any μ_λ. By 9.5(a), f is not oriented with head at b_s. By 9.5(b) and 9.6(c) f is not level. Hence f is oriented with tail at b_s. Then f is asynchronous (6.5(a)), so $l_{j+2} < q$ (9.1(c)). This contradicts the choice of j. ||

REFERENCES

1. S. BLEILER and M. SCHARLEMANN: Property P, strongly invertible knots, and a problem of J. Martin, *Math. Ann.* **273** (1986), 215–225.
2. M. CULLER, C. GORDON, J. LUEKE, and P. SHALEN: Dehn surgery on knots, *Ann. of Math.* **125** (1987), 237–300.
3. J. HEMPEL: A simply-connected 3-manifold is S^3 if it is the sum of a solid torus and the complement of a torus knot, *Proc. Am. Math. Soc.* **15** (1964), 154–158.

4. R. KIRBY: Problems in low-dimensional topology, *Proc. Symposium Pure Mathematics* **32** (1978), 273–312.
5. M. SCHARLEMANN: Smooth spheres in $\mathbb{R}^4$ with four critical points are standard, *Invent. Math.* **79** (1985), 125–141.
6. M. SCHARLEMANN: Unknotting number one knots are prime, *Invent. Math.* **79** (1985), 121–141.
7. M. TAKAHASHI: Two-bridge knots have property *P*, *Mem. Am. Math. Soc.*, Vol. **29**, No. 239 (1981).

Department of Mathematics
Portland State University
Portland
OR 97212
U.S.A.

Department of Mathematics
University of California
Santa Barbara
CA 93106
U.S.A.