Hosokawa, F. and Yanagawa, T.
Osaka J. Math.
2 (1965), 373-384

IS EVERY SLICE KNOT A RIBBON KNOT?

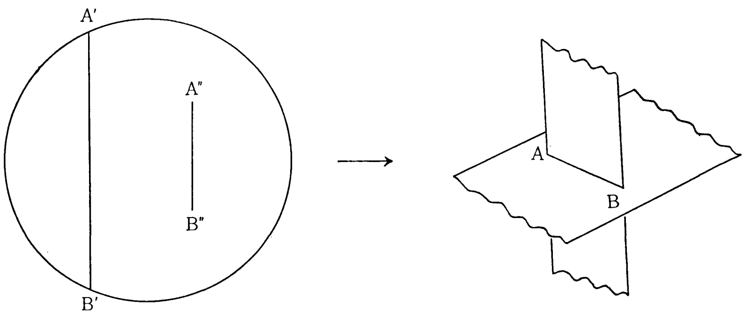
Dedicated to Professor H. Terasaka on his 60-th birthday

FUJITSUGU HOSOKAWA and TAKAAKI YANAGAWA

(Received September 4, 1965)

Let S^2 be a locally flat 2-sphere in a 4-dimensional euclidean space R^4, then the knot obtained by slicing S^2 with a hyperplane in R^4 is called *a slice knot* [1].

A singular disk that is a continuous image of the unit 2-cell in a 3-dimensional euclidean space R^3 will be called *a ribbon*, if and only if each of its singularities is of the following type:



and a knot that is the boundary[1)] of a ribbon is called *a ribbon knot.*

R. H. Fox presented a problem "Is every slice knot a ribbon knot?" in his paper [2]. The purpose of this paper is to give an affirmative answer to the problem.

In this paper we will consider everything from the semilinear point of view.

1. In this paper, we shall use *the motion picture method* of describing surfaces in a 4-space R^4.

Let κ be a slice knot in a 3-space R^3, and $H^4[t_1, t_2]$ be a subspace $R^3 \times [t_1, t_2]$ in a 4-space $R^4 = R^3 \times (-\infty, \infty)$, where (,) means open interval and [,] closed.

1) The boundary of a singular disk (a ribbon) means the image of the boundary of the inverse image.

Then, we can easily see that there exists a locally flat, non-singular 2-cell e^2 in $H^4(-1, 0]$ with boundary κ. In order to decribe e^2, we consider the intersections of e^2 with the hyperplanes $R^3_t = R^3 \times [t]$ and describe the changing of the configuration as t increases from -1 to 0. We can modify e^2 isotopically in $H^4(-1, 0]$ so that these intersections, except at a finite number of levels, consist of collections of oriented, simple closed polygons which vary continuously with t. In each exceptional level of the modified e^2 there are a finite number of critical points called *elementary critical points*, but these can be classified in three types, that is, as t increases through the critical value t_0, the configuration changes as follows;

(elliptic critical point of type I at $t = t_0$:) a small unknotted simple closed polygon appears as in figure 1,

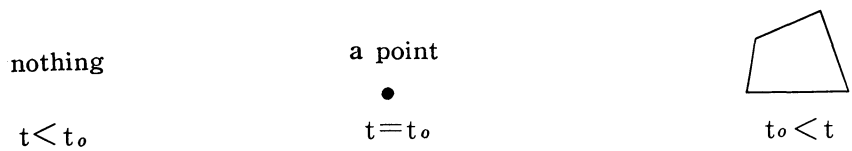


Fig. 1

(elliptic critical point of type II at $t = t_0$:) a small unknotted simple closed polygon shrinks to a point and disappears as in figure 2,

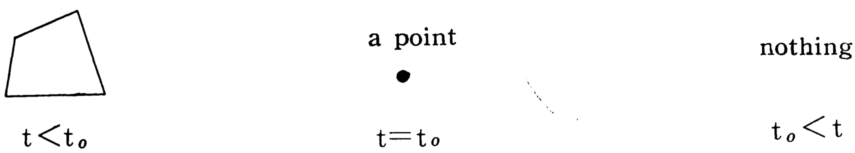


Fig. 2

(hyperbolic critical point at $t = t_0$:) two arcs approach each other and cross over as in figure 3,

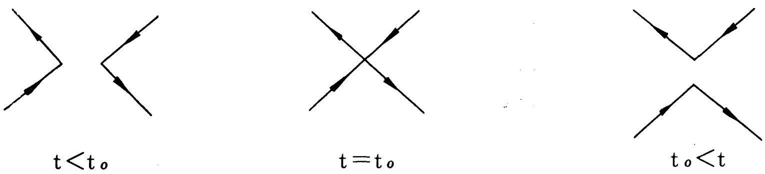


Fig. 3

Next, we will simplify the arrangement of the critical points of e^2.

Since $H^4[-1, 1] - e^2$ is arcwise connected, it is easy to modify e^2 so that elliptic critical points of type I and II are found only at $t = -1$ and $t = 1$ respectively, and the hyperbolic critical points are found at $-1 < t < 0$. If the critical points of the 2-cell e^2 are as above, we will say that the 2-cell e^2 has *a* $(-1, 1)$-*canonical form.* Therefore, we have the following

Lemma 1. *If a knot* κ *in* $R^3_0=R^3\times[0]$ *is a slice knots then there is a locally flat, non-singular 2-cell* e^2 *with boundary* κ *in* $H^4[-1, 1]$ *which has a* $(-1, 1)$-*canonical form.*

Now, we may assume that the p hyperbolic critical points of e^2 are found one by one at the values t_i, $i=1, 2, \cdots, p$, where $-1<t_1<t_2<, \cdots, <t_p<0$. Let q and r be the number of the elliptic critical points of type I and type II respectively, where p, q, r are non-negative integers satisfying $q+r-p=1$.

We will map $e^2\cap H^4[-1, 0]$ into R^3.

Let ε be a sufficiently small positive integer. Since $e^2\cap H^4[-1, t_1-\varepsilon]$ $(0<\varepsilon<t_1+1)$ consists of q disjoint, non-singular, locally flat 2-cells with boundary $e^2\cap R^3_{t_1-\varepsilon}$ which are q disjoint unlinked unknotted circles, we can map $e^2\cap H^4[-1, t_1-\varepsilon]$ to q disjoint, non-singular 2-cells in R^3 by homeomorphism h_0. Next, since $e^2\cap H^4[t_1-\varepsilon, t_2-\varepsilon]$ $(0<\varepsilon<t_2-t_1)$ consists of non-singular perforated disks with boundaries $e^2\cap R^3_{t_1-\varepsilon}$ and $e^2\cap R^3_{t_2-\varepsilon}$, $e^2\cap H^4[t_1-\varepsilon, t_2-\varepsilon]$ can be deformed to $e^2\cap R^3_{t_1-\varepsilon}$ with a band B_1 attached at the two disjoint small arcs of $e^2\cap R^3_{t_1-\varepsilon}$ at which the hyperbolic critical point appears, as in figure 4 [3]. Denote this deformation by g_1.

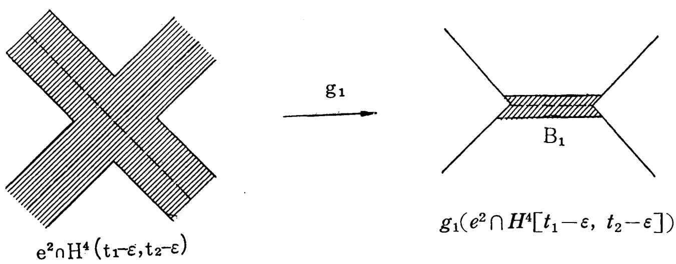


Fig. 4

We map $g_1(e^2\cap H^4[t_1-\varepsilon, t_2-\varepsilon])$ into R^3 by a homeomorphism h_1 satisfy- $h_1g_1(e^2\cap R^3_{t_1-\varepsilon})=h_0(e^2\cap R^3_{t_1-\varepsilon})$.

Next, since $e^2\cap H^4[t_2-\varepsilon, t_3-\varepsilon]$ $(0<\varepsilon<t_3-t_2)$ consists of non-singular perforated disks with boundaries $e^2\cap R^3_{t_2-\varepsilon}$ and $e^2\cap R^3_{t_3-\varepsilon}$, we see that $e^2\cap H^4[t_2-\varepsilon, t_3-\varepsilon]$ can be deformed to $e^2\cap R^3_{t_2-\varepsilon}$ with a band B_2 attached at the two disjoint small arcs of $e^2\cap R^3_{t_2-\varepsilon}$ at which the hyperbolic critical point appears. Denote this deformation by g_2. Also we map $g_2(e^2\cap H^4[t_2-\varepsilon, t_3-\varepsilon])$ into R^3 by a homeomorphism h_2 satisfying $h_2g_2(e^2\cap R_{t_2-\varepsilon})=h_1g_1(e^2\cap R^3_{t_2-\varepsilon})$.

Repeating these processes, we have the deformations g_i from $e^2 \cap H^4[t_i-\varepsilon, t_{i+1}-\varepsilon]$ $(0<\varepsilon<t_{i+1}-t_i)$ to $e^2 \cap R^3_{t_i-\varepsilon}$ with the attached band B_i and the homeomorphisms h_i from $g_i(e^2 \cap H^4[t_i-\varepsilon, t_{i+1}-\varepsilon])$ into R^3 satisfying $h_i g_i(e^2 \cap R^3_{t_i-\varepsilon}) = h_{i-1} g_{i-1}(e^2 \cap R^3_{t_i-\varepsilon})$ where $i=2, 3, \cdots, p$ and $t_{p+1}=0$.

Since $e_2 \cap H^4[-\varepsilon, 0]$ consists of non-singular perforated disks with boundaries $e^2 \cap R^3_{-\varepsilon}$ and $e \cap R^3_0$, we can deform $e^2 \cap H^4(-\varepsilon, 0)$ by a deformation g_{p+1} to $e_2 \cap R^3_0$, and $g_{p+1}(e^2 \cap H^4[-\varepsilon, 0])$ can be mapped into R^3 by a homeomorphism h_{p+1} satisfying $h_p g_p(e^2 \cap R^3_{-\varepsilon}) = h_{p+1} g_{p+1}(e^2 \cap R^3_{-\varepsilon}) = h_{p+1}(e^2 \cap R^3_0)$. Now, by applying $g_1, g_2, \cdots, g_{p+1}$ and $h_0, h_1, \cdots, h_{p+1}$ to $e^2 \cap H^4[-1, 0]$, we obtain a singular perforated disk composed of q disjoint 2-cells with the attached bands $B_1, \cdots, B_p$.

The boundary of the singular perforated disk now consists of κ and r disjoint unlinked unknotted circles $\alpha_1, \cdots, \alpha_r$. The circles $\alpha_1, \cdots, \alpha_r$ bound r mutually disjoint non-singular disks that do not intersect κ, and these disks correspond to the r elliptic critical points of type II. The singularities of the singular perforated disk are of the following four types:

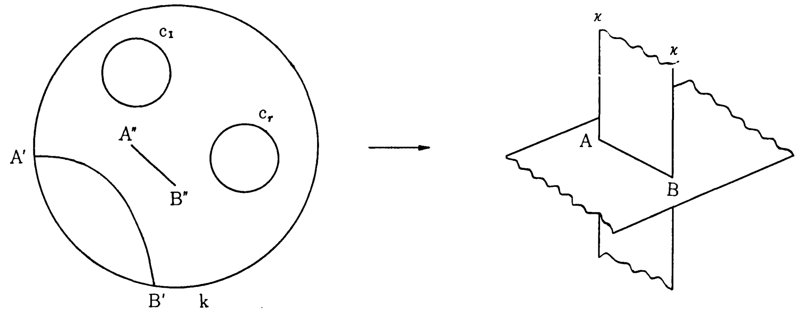


Fig. 5a

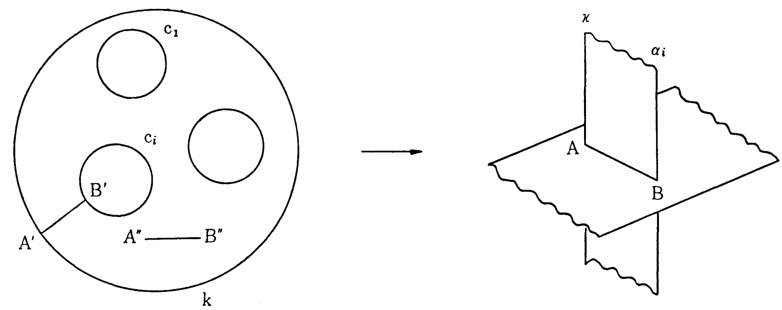


Fig. 5b

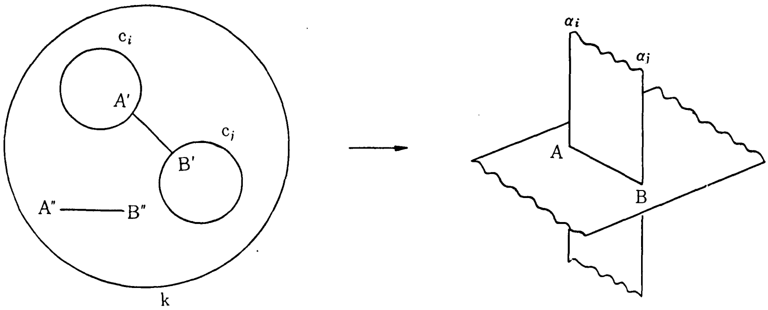


Fig. 5c

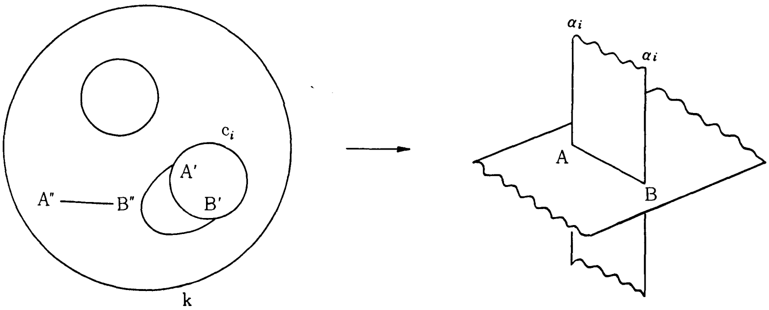


Fig. 5d

where $k, c_1, \cdots, c_r$ are inverse images of $\kappa, \alpha_1, \cdots, \alpha_r$.

A singular perforated disk in R^3 will be called *a perforated ribbon* if the singularities are all of the above four types.

From the above considerations we have

Lemma 2. *If a knot κ is a slice knot in R^3, there exists in $R^3-\kappa$ a collection of r mutually disjoint, non-singular 2-cells $\sigma_1, \sigma_2, \cdots, \sigma_r$ such that κ and the boundary circles $\alpha_1, \alpha_2, \cdots, \alpha_r$ of the 2-cells $\sigma_1, \sigma_2, \cdots, \sigma_r$ bound a perforated ribbon in R^3.*

2. Let κ be a slice knot in R^3. By lemma 2, there are non-singular disks $\sigma_1, \sigma_2, \cdots, \sigma_r$ which are disjoint to each other and do not intersect κ, and there is a perforated ribbon σ_0 such that $\partial\sigma_0=\kappa\cup\partial\sigma_1\cup\cdots\cup\partial\sigma_r$, where $\partial\sigma_i$ means the boundary α_i of σ_i $(i=1, 2, \cdots, r)$. Let $\sigma=\sigma_0\cup\sigma_1\cup\cdots\cup\sigma_r$, then σ is a singular disk with boundary κ.

Now, we will examine the singularities of σ.

The singularities of σ consist of the self-intersections of σ_0 and the intersections of σ_0 and σ_i $(i=1, 2, \cdots, r)$, for σ_i has no self-intersections and σ_i and σ_j are mutually disjoint for $i\neq j$, $i, j=1, 2, \cdots, r$.

Let $D, D_0, D_1, \cdots, D_r$ be the inverse images of $\sigma, \sigma_0 \sigma_1, \cdots, \sigma_r$, and let $k, c_1, c_2, \cdots, c_r$ be the inverse images of $\kappa, \alpha_1, \alpha_2, \cdots, \alpha_r$, that is $\partial D=k$, $\partial D_i=c_i$ $(i=1, 2, \cdots, r)$.

Now, let us consider the intersections of σ_0 and σ_i, $i\neq 0$. The singularities of σ_0 are of the four types of Fig. 5a, 5b, 5c, 5d, that is, they are double lines (the segment $\overline{AB}$ in figure 5). The two endpoints of one of the inverse images $\overline{A'B'}$, $\overline{A''B''}$ of $\overline{AB}$ are boundary point of D_0 (A', B' figure 5), and the two endpoints of the other are inner points of D_0 (A'', B'' in figure 5).

We will call the inverse image of a double line of σ_0 whose endpoints are boundary points of D_0 a *b-line* and an inverse image whose endpoints are inner points of D_0 an *i-line*.

In the case that an endpoint A of a double line $\overline{AB}$ of σ_0 is on $\alpha_i=\partial\sigma_i$, the intersection of σ_0 and σ_i must contain a double line whose endpoint is A, and we may modify σ_i so that the double line does not intersect the double line $\overline{AB}$ in a neighborhood of A (as shown in figure 6).

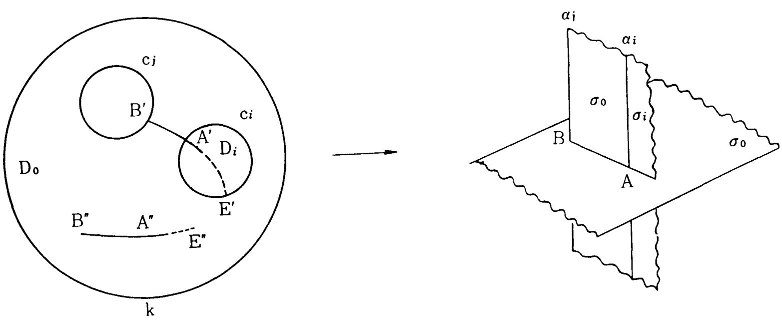


Fig. 6

At the inverse image of a double line, the b-line $\overline{A'B'}$ extends on D_i from A' to a point E' of c_i, and the i-line $\overline{A''B''}$ extends on D_0 from A'' to a point E'' whose image is identical with the image of E'. Let E be the image of E' and E'', then the following two cases can occur.

Case I. In the case that E is not a singular point of σ_0 we have $E'=E''$, because E' and E'' are on D_0 and E is not a singular point of σ_0. Then E must be a branch point of σ.

Case II. In the case that E is a singular point of σ_0 there is a double line of σ_0 having E as one of its endpoints. If the other endpoint G is on κ, G is an endpoint of a completed double line of σ. If G is not on κ but on α_i $(i=1, 2, \cdots, r)$, the double line extends again and at last either it arrives at the endpoint of a completed double line of σ or it arrives at the point B and forms a closed double line of σ or the case I occurs.

We remark that if a b-line meets an i-line, then the junction-point must be a branch point.

The remaining singular lines of σ arise from the intersection of $\sigma_0 - \partial\sigma_0$ with $\sigma_i - \partial\sigma_i$ ($i \neq 0$). Such a singular line must be a closed double curve.

We will call a double line, whose endpoints are on the boundary, *a double line of ribbon type.*

From the above considerations, the singularities of σ are of the following types[3];

i) *double lines of ribbon type.*
ii) *closed double curves.*
iii) *triple points which are crossing points of double lines.*
iv) *branch points.*

3.

(I) Branch points

Since the interiors of σ_0 and σ_i have no branch points, a branch point appears only in case I of section 2; that is, all branch points are on α_i ($i = 1, 2, \cdots, r$). As a branch point is the result of local winding of σ_i around α_i, we may modify σ_i so that there is only one double line through the branch point. Then the double lines through the points are of the following two types, and figures 7a, 7b show their inverse images.

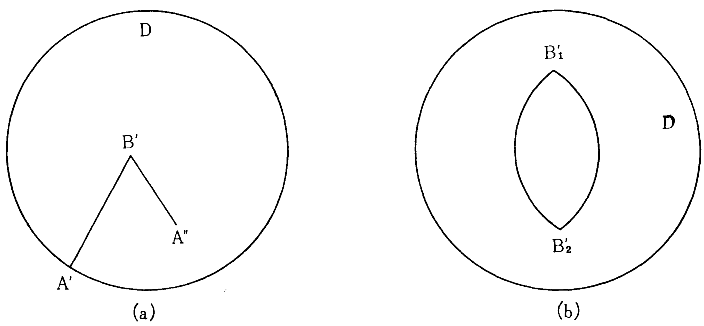


Fig. 7

We can eliminate such a branch point by cutting along a double line through the branch point, as shown in figure 8 [4].

2) By a slight modification we can put σ into general position,

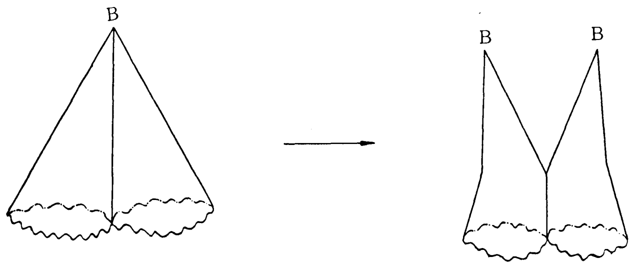


Fig. 8

[II] Triple points

A triple point is a crossing of double lines; that is, if A is a triple point then there are just three double lines through A. Let these double lines be $\overline{EG}$, $\overline{HI}$, $\overline{JK}$ in a neighborhood of A, and the inverse image of A be A', A'', A''', then the inverse image of σ in a neighborhood of A is as in figure 9.

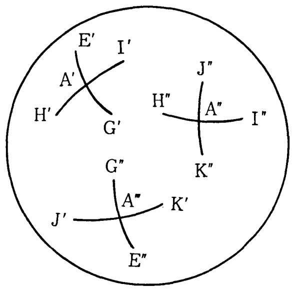


Fig. 9

Now, we will consider triple points on double lines of ribbon type. By the considerations about singularities of σ in section 2, one of the three inverse images of every triple point does not belong to i-line. By cutting σ along the image of an arc starting from a boundary point and disjoint with i-lines, we may suppose that $\overline{E'G'}$ is a subarc of a b-line, say $\overline{E'G'}$ again, and $\overline{H'I'}$ is of a b-line or of a closed double curve, and $\overline{G'A'}$ does not contain any inverse image of triple points as in figure 10-left upside.

We take such a point G_1 on $\overline{EA}$ that the double line $\overline{GG_1}$ contains no triple points except A. Now we cut off σ along $\overline{GG_1}$ as in figur 10, then the triple point A vanishes.

The cut σ has no singularities of new types and the knot type of the boundary does not change. We will denote the cut σ and its boundary by σ and κ again.

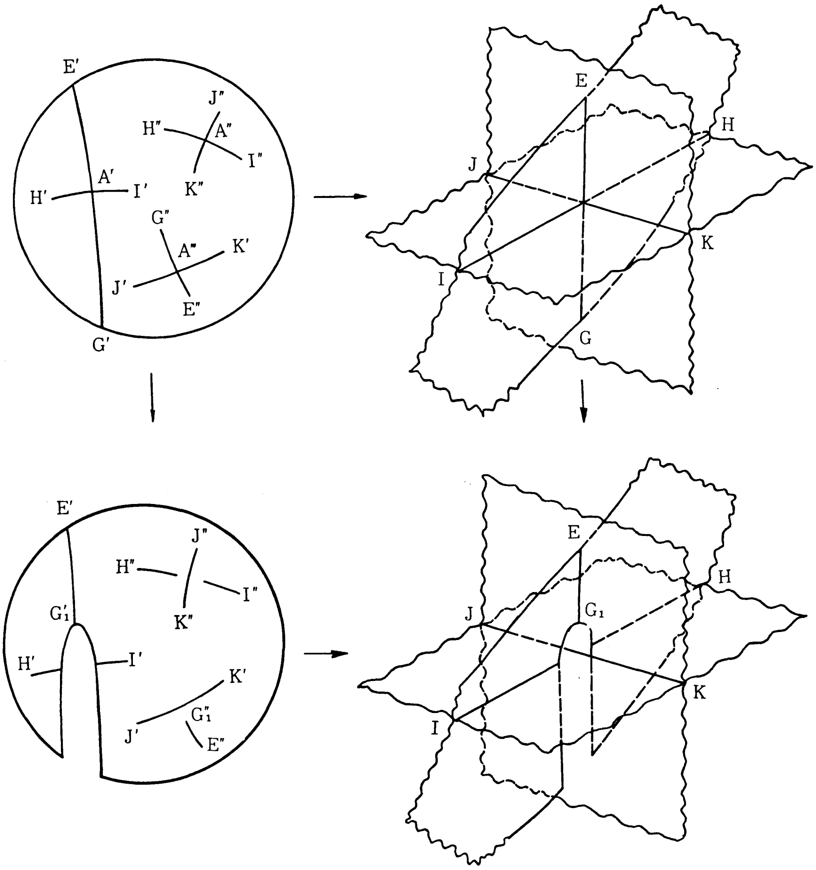


Fig. 10

Repeating these processes, *we can remove all the triple points on the double lines of ribbon type.*

Now, let the collection $\{\overline{A_1'B_1'}, \overline{A_2'B_2'}, \cdots, \overline{A_m'B_m'}\}$ be the b-lines on D; that is, the inverse images of the double lines of ribbon type whose endpoints are on k. Let d_i be a narrow band in D which contains $\overline{A_i'B_i'}$ in its interior and does not contain the inverse images of the other singularities of σ $(i=1, \cdots, m)$. By cutting D along the boundaries of $d_1, d_2, \cdots, d_m$, the cell is separated into $2m+1$ disks $d_1, d_2, \cdots, d_m$ and $d_{m+1}, d_{m+2}, \cdots, d_{2m+1}$ as in figure 11.

Let $\Delta_1, \Delta_2, \cdots, \Delta_{2m+1}$ be the images of $d_1, d_2, \cdots, d_{2m+1}$, then $\Delta_1, \cdots, \Delta_m$ are mutually disjoint, non-singular disks, and $\Delta_{m+1}, \cdots, \Delta_{2m+1}$ are singular disks. However $\Delta_{m+1} \cup \Delta_{m+2} \cup \cdots \cup \Delta_{2m+1}$ has no singularities in a neighborhood of $\partial\Delta_{m+1} \cup \partial\Delta_{m+2} \cup \cdots \cup \partial\Delta_{2m+1}$. Therefore, by a simple extension of Dehn's lemma (proof in the Appendix) we can replace $\Delta_{m+1}, \Delta_{m+2}, \cdots,$

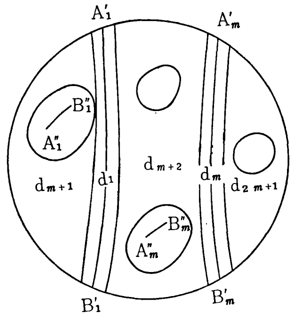


Fig. 11

Δ_{2m+1} by mutullay disjoint, non-singular disks $\vartheta_{m+1}, \vartheta_{m+2}, \cdots, \vartheta_{2m+1}$ that differ from $\Delta_{m+1} \cup \cdots \cup \Delta_{2m+1}$ only on a compact subset of $(\vartheta_{m+1} - \partial\vartheta_{m+1}) \cup \cdots \cup (\vartheta_{2m+1} - \partial\vartheta_{2m+1})$.

Now, by identifying the disks $\Delta_1, \cdots, \Delta_m, \vartheta_{m+1}, \cdots, \vartheta_{2m+1}$ along the cuts as before, we have a new singular disk with boundary κ. Let us again denote this new disk by σ.

We will now consider the singularities of the new disk σ.

Since $\Delta_1, \cdots, \Delta_m$, are mutually disjoint non-singular disks, and $\vartheta_{m+1}, \cdots, \vartheta_{2m+1}$ are mutually disjoint non-singular disks, all the singularities occur only as intersections of Δ_i and ϑ_j $(i=1, \cdots, m, j=m+1, \cdots, 2m+1)$.

After a slight modification, only double lines, triple points and branch point occur as singularities.

When a double line crosses a cut, the inverse images of the crossing point must be identical, and such a singular point must be a branch point. Furthermore every branch point arises in this way.

If there is a double line which is neither closed nor of ribbon type, the inverse images of the double line are two arcs $\overline{A'B''}$ and $\overline{A''B'}$ and both A' and B' are on κ, where A' is on the boundary of ϑ_j and B' is on the boundary of Δ_i (as in figure 12). But Δ_i does not intersect

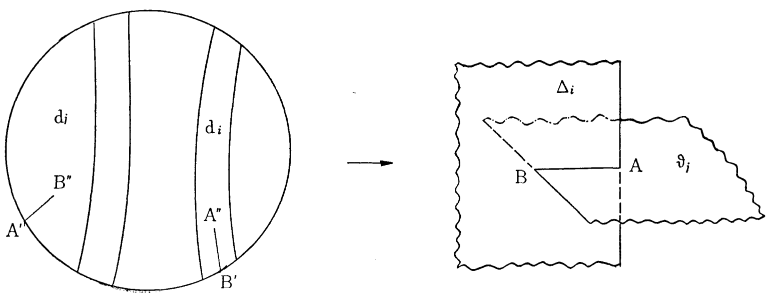


Fig. 12

$\kappa \cap \vartheta_j$, so this is a contradiction.

If there is a triple point, it must be contained in three different disks of the collection $\Delta_1, \cdots, \Delta_m, \vartheta_{m+1}, \cdots \vartheta_{2m+1}$. But this is impossible.

Since we can eliminate the branch points as in the beginning of this section, we can modify[3)] σ so that the singularities are all of the following type;

1) *double lines of ribbon type*
2) *closed double curves.*

As there exist no triple points, all the closed double curves are simple and mutually disjoint.

Therefore we can easily remove these simple closed double curves by the well-known cut-and-exchange method.

Thus we can obtain a ribbon modified from σ with the boundary κ. Therefore we have

Theorem. *Every slice knot is a ribbon knot.*

H. Terasaka proved that *the Alexander polynomial of a ribbon knot is of the form* $\pm t^m f(t)\cdot f(t^{-1})$, [7]. Therefore it follows from the above Theorem that *the Alexander polynomial of a slice knot is of the form* $\pm t^m f(t) f(t^{-1})$, [8].

Appendix

Lemma. *If $D_1, D_2, \cdots, D_r$ is a set of normal, canonical Dehn-disks such that $\partial D_i \cap D_j = \phi$ $(i \neq j, i, j = 1, 2, \cdots, r)$ in a 3-manifold M, then there exists a set of mutually disjoint, non-singular disks $\vartheta_1, \vartheta_2, \cdots, \vartheta_r$ such that ϑ_i is identical with D_i in a sufficiently small neighborhood of ∂D_i $(i=1, 2, \cdots, r)$.*

Proof. By making use of Dehn's lemma repeatedly, we have a set of non-singular disks $D_1', D_2', \cdots, D_r'$ such that D_i' is identical with D_i in a small neighborhood of ∂D_i $(i=1, 2, \cdots, r)$.

Since D_1' is a non-singular disk and $D_1' \cap \partial D_i' = \phi$ $i=2, \cdots, r$ there is a 3-cell V_1 containing D_1' in its interior and contained in $M - \partial D_2' \cup \cdots \cup \partial D_r'$. Then we can construct such a homeomorphism φ_1 that φ_1 is the identity in $\overline{M - V_1}$ and that $\varphi_1(D_1') \subset M - D_2' \cup \cdots \cup D_r'$.

Now the set of non-singular disks $D_1', \varphi_1^{-1}(D_2'), \cdots, \varphi_1^{-1}(D_r')$ has the following properties;

(i) $D_1' \cap \varphi_1^{-1}(D_i') = \phi$, and

(ii) $\varphi_1^{-1}(D_i')$ is identical with D_i' in a small neighborhood of $\partial D_i'$

3) This modification has also to eliminate any multiple points that may appear on $\partial \Delta_i \cap \kappa$ $(i=1, \cdots, m)$.

$(i=2, \cdots, r)$.

Repeating this process $r-1$ times produces a set of disks $\vartheta_1, \vartheta_2, \cdots, \vartheta_r$ having the required property.

Bibliograpy

[1] R. H. Fox: *A quick trip through knot theory*, Top. of 3-manifolds and related topics, edited by M. K. Fort Jr., Prentice Hall, 1962.

[2] R. H. Fox: *Some problems in Knot theory*, ibid.

[3] J. Milnor: Morse theory, Ann. of Math. Studies No. 51, 1963.

[4] C. D. Papakyriakopoulos: *On Dehn's lemma and asphericity of knots*, Ann. of Math. **66** (1957), 1-26.

[5] A. Shapiro and J. H. C. Whitehead: *A proof and extension of Dehn's lemma*, Bull. Amer. Math. Soc. **64** (1958), 174-178.

[6] T. Homma: *On Dehn's lemma for S^3*, Yokohama Math. J. **5** (1957), 223-244.

[7] H. Terasaka: *On null-equivalent knots*, Osaka Math. J. **11** (1959), 95-113.

[8] R. H. Fox and J. Milnor: *Singularities of 2-spheres in 4-space and equivalence of knots*, unpublished. Cf. Bull. Amer. Math. Soc. **63** (1957), 406.