Arc-presentations of links: Monotonic simplification

by

I. A. Dynnikov (Moscow)

Abstract. In the early 90's J. Birman and W. Menasco worked out a nice technique for studying links presented in the form of a closed braid. The technique is based on certain foliated surfaces and uses tricks similar to those that were introduced earlier by D. Bennequin. A few years later P. Cromwell adapted Birman–Menasco's method for studying so-called arc-presentations of links and established some of their basic properties. Here we further develop that technique and the theory of arc-presentations, and prove that any arc-presentation of the unknot admits a (non-strictly) monotonic simplification by elementary moves; this yields a simple algorithm for recognizing the unknot. We also show that the problem of recognizing split links and that of factorizing a composite link can be solved in a similar manner. We also define two easily checked sufficient conditions for knottedness.

Contents

INTRODUCTION

The most important problem of knot theory is the classification of knots and links. To construct an algorithm that, for any two given links, decides whether they are isotopic or not is known to be a very hard problem. A solution published in [13] uses contributions of many mathematicians including W. Haken, F. Waldhausen, W. Jaco, P. Shalen, K. Johannson, G. Hemion, W. Thurston, and S. Matveev. The previously announced solution by G. Hemion was shown to be incomplete by S. Matveev. It is not difficult to construct an algorithm that enumerates all diagrams of links, which, in conjunction with the previously mentioned algorithm, makes it possible to generate a list of diagrams in which each isotopy class of links is presented exactly once and, for any other diagram, to locate the corresponding isotopy class in the list.

The algorithm for recognizing links mentioned above is extremely complex and gives only a theoretical solution of the problem. It actually deals with three-manifolds which are link complements provided with a meridian-longitude pair in each connected component of the boundary. Up to now, this algorithm cannot be used in practice because of the huge number of operations needed for its implementation, even for links with a small number of crossings.

At the same time, a theoretical solution of the problem will be given if, for some computable function $f(n)$ (say, $\exp(\exp(\exp(n!)))$), one shows that any two diagrams of the same link type that have $\leq n$ crossings can be obtained from each other by $\leq f(n)$ Reidemeister moves. However, an explicit formula for such a function is still unknown.

It seems to be widely understood that the following problems, which are important ingredients of the general classification problem, are simpler from the practical point of view:

1) recognizing the trivial knot;
2) recognizing a split link and presenting it as the distant union of two non-empty links;
3) recognizing a composite non-split link and presenting it as the connected sum of two non-trivial links.

We shall call the aggregate of these three problems the *decomposition problem*, since a solution of all three allows one to express any given link in terms of prime non-split links and the unknot by using the connected sum and distant union operations. Such a decomposition is known to be essentially unique, but to solve the general classification problem, one must also construct an algorithm for comparing isotopy classes of two prime non-split links that are presented by their diagrams.

The decomposition problem was solved by W. Haken [10] and H. Schubert [18] by using Haken's method of normal surfaces, which plays a very important rôle in modern three-dimensional topology. (In particular, normal surfaces are used in the algorithm for recognizing Haken manifolds mentioned above and in the Rubinstein–Thompson algorithm for recognizing the three-sphere).

Finding normal surfaces seems to be an exponentially hard problem. At least, the known realizations are that hard, both theoretically and in practice. So, any algorithm that uses Haken's technique has almost no chance to be implementable in a reasonable time.

In recent years, a few attempts have been made to find an algorithm for recognizing the trivial knot by using some *monotonic simplification*. By the latter we mean the following. One chooses a way of presenting knots by diagrams of certain type and introduces the notion of *complexity* $c(D)$ of a diagram. For the chosen type of presentation, one also fixes a set of *elementary moves* that do not alter the topological type of a knot. Then, for a given diagram D of a knot, one searches for a sequence of diagrams

$$(1) \qquad\qquad D_0 = D, D_1, D_2, \ldots$$

such that, for each i, the diagram D_{i+1} is obtained from D_i by an elementary move and we have $c(D_{i+1}) \leq c(D_i)$ (or $c(D_{i+1}) < c(D_i)$, in which case we shall speak of a *strictly monotonic simplification*). If there exists an algorithm that, for any D, produces such a sequence $D_0 = D, D_1, \ldots, D_N = D'$ in which the last diagram D' cannot be simplified any more, we shall say that D is *monotonically simplifiable* to D'. Such an algorithm certainly exists if, for any n, there are only finitely many diagrams D such that $c(D) < n$ (which will be the case for arc-presentations) because, in this case, we can find *all* monotonic simplification sequences in finite time by using an exhaustive search. If, for any diagram D of the unknot, the diagram D' obtained from D by a simplification algorithm is always the *trivial diagram*, we shall speak about the recognition of the unknot by monotonic simplification.

It is well known that, in the case of ordinary planar diagrams with Reidemeister moves as elementary moves and the crossing number as the measure of complexity, recognizing the unknot by monotonic simplification is not possible because there exist diagrams of the unknot that cannot be simplified to the trivial circle by using only Reidemeister moves not increasing the crossing number.

According to A. E. Hatcher's solution of the Smale conjecture [11], there is no topological obstruction to the existence of a flow ω on the space of knots such that any unknotted curve will evolve under ω to a round circle. Such a flow may be the gradient flow of some energy function, which plays the rôle of the measure of complexity. An appropriate discretization of the flow may

lead to a strictly monotonic simplification algorithm. Some investigation of the gradient flows on the knot space and numerical experiments have been done in the literature (see, e.g., [16], [9], [12]).

A monotonic simplification procedure for spines of three-manifolds has been worked out by S. V. Matveev in [14] for recognizing three-manifolds. It can also be used for recognizing the unknot: for a given knot one first constructs a spine of the knot complement and then applies the simplification algorithm. Numerical experiments of this kind have been done by E. Fominykh.

In [8], the author introduced three-page presentations of knots. The complexity of a three-page knot is defined to be the number of vertices at the binding line of the three-page book. A simplification procedure for three-page knots has been tested on a series of examples.

In all three situations mentioned above, it has not been proved that a monotonic simplification to a trivial circle is always possible for any presentation of the unknot, but no counterexamples have been found either, though attempts have been made.

The first successful attempt to find a measure of complexity with respect to which any presentation of the unknot admits a monotonic simplification was made by J. Birman and W. Menasco in [4]. They studied presentations of a knot as the closure of a braid and showed that the number of strands in a non-trivial braid presenting the unknot can always be reduced by moves of certain types. The moves include the Markov moves that do not increase the number of strands (i.e., braid conjugations and "destabilizations", which are the Markov moves that reduce the number of strands) and so-called exchange moves, which preserve the number of strands. The ideas of [4] and the known solution of the conjugacy problem for braid groups were used by J. Birman and M. Hirsch in [2] to construct a new algorithm for recognizing the unknot. This algorithm, however, does not have the form of a monotonic simplification algorithm in our sense, and it is not clear whether the straightforward idea of monotonic simplification can be used here, since no algorithm is known to decide whether a given braid conjugacy class admits a destabilization after finitely may exchanges.

The problem of recognizing the unknot happens to be very closely related to two other problems: recognizing split links and recognizing composite links. In many cases, the same technique allows one to solve all three problems. The distant union and connected sum operations can be defined at the level of diagrams. But a split (or composite) link can usually be presented by a non-split (respectively, prime) diagram. So, the general question is this: for a given diagram D, can one apply finitely many elementary moves without increasing the complexity so that the final diagram is obtained by

the distant union and connected sum operations from trivial diagrams and diagrams of prime non-split links? J. Birman and W. Menasco have shown in [3] that the answer is positive in the case of closed braids (there is a gap in the proof of the result on composite links, but the assertion is true).

At first sight, arc-presentations, which we study in this paper, seem to have nothing in common with braids. However, the braid foliation technique that J. Birman and W. Menasco developed in a series of papers including [3] and [4] turned out to be very well adapted for studying arc-presentations of knots. The fact that the Birman–Menasco foliated surface technique can be extended to arc-presentations was discovered by P. Cromwell in [6], where he proved the additivity of the arc index under the connected sum operation. The possibility of a monotonic simplification and other algorithmic questions are not discussed in [6], but a sequence of moves preserving the complexity appears in the proof. The generalized type I move of [6] can be easily decomposed into elementary moves (Proposition 5 of this paper). In conjunction with this remark, the arguments of [6] would suffice for a proof of recognizibility of split links by monotonic simplification. As for recognizing composite links, there is a gap in the final part of the proof of the main result of [6], which is covered in this paper. We also provide some technical details that are not mentioned in [6].

In addition to this, we extend the foliated surface technique to spanning discs of arc-presentations of the unknot. Compared to the case of a splitting or factorizing sphere, spanning discs require care in considering their behaviour near the boundary.

In the earlier version of the present manuscript, the treatment of the boundary was not careful enough, and a gap occurred in the proof of the main result. The mistake was discovered by W. Menasco and A. Sikora who also suggested an idea for filling the gap. We use their suggestion in a modified form.

The main result of this paper is the following.

THEOREM 1. *The decomposition problem of arc-presentations is solvable by monotonic simplification.*

In order to make this claim less abstract in this superficial introduction, we explain it in a very elementary language. Let a *rectangular diagram* be an ordinary planar diagram D of a link satisfying the following restrictions:

1) D consists only of vertical and horizontal straight line segments, which we call *edges*;
2) at each crossing of D, the vertical arc is overcrossing and the horizontal one undercrossing;
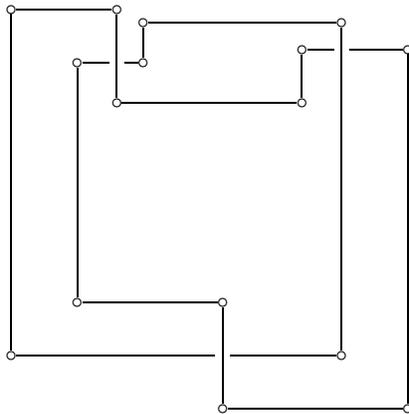3) no two edges are collinear (see Fig. 1).

Fig. 1. A rectangular diagram

Two such diagrams are called *combinatorially equivalent* if they are isotopic in the plane via an ambient isotopy h that has the form $h(x, y) = (f(x), g(y))$.

Since the type of a crossing is determined by its position, there is no need to indicate in figures which arc is overcrossing and which one undercrossing.

Clearly, any planar diagram of a link is isotopic to a rectangular diagram (see below).

The following transforms of rectangular diagrams are *elementary moves*:

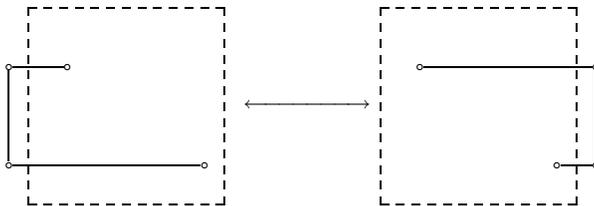0) cyclic permutation of horizontal (vertical) edges (see Fig. 2);

Fig. 2. Cyclic permutation of vertical edges

1) stabilization and destabilization (Fig. 3);
2) interchanging neighbouring edges if their pairs of endpoints do not interleave (Fig. 4).

Figures 2–4 illustrate particular cases of the moves. To obtain all the other cases, one should apply all possible rotations by $\pi k/2$, $k \in \mathbb{Z}$, and reflections in horizontal and vertical lines if necessary.

The *complexity* $c(D)$ of a rectangular diagram D is the number of vertical edges in D. The only operations changing the complexity are stabilization (which increases it by 1) and destabilization (inverse to stabilization). We
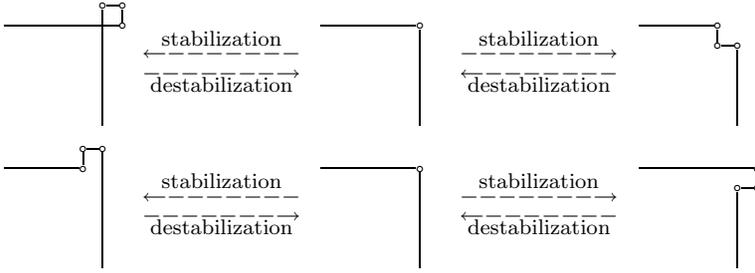
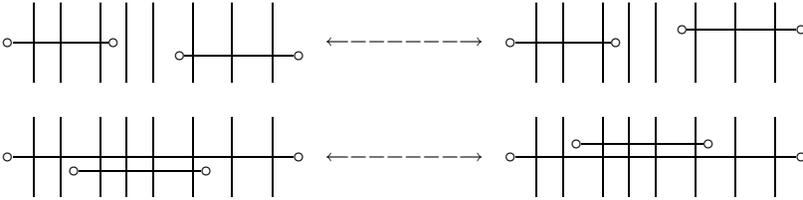Fig. 3. Stabilization and destabilization moves

Fig. 4. Interchanging horizontal edges

stress that the number of crossings is *not* taken into account in the definition of complexity. The number of crossings of a rectangular diagram of complexity n is bounded from above by $(n-1)^2/2$. The simplest possible rectangular diagram, the *trivial diagram*, is an ordinary rectangle. Its complexity is 2.

A rectangular diagram D is a *distant union* of diagrams D_1, D_2 if there is a vertical line l that does not intersect D and splits it into two non-trivial parts which are equivalent to D_1, D_2. We write $D = D_1 \sqcup D_2$ in this case.

A rectangular diagram D is a *connected sum* of diagrams D_1 and D_2 if there exists a rectangular disk $\mathcal{R}$ with vertical and horizontal sides whose boundary $\partial\mathcal{R}$ meets D in two points so that the rectangular diagrams D_1 and D_2 (in either order) are obtained from $\mathcal{R} \cap D$ and $(\mathbb{R}^2 \setminus \mathcal{R}) \cap D$ by connecting the two points of $(\partial\mathcal{R}) \cap D$ by a simple path lying outside or inside $\mathcal{R}$, respectively. In this case, we say that the diagram D is *composite* and write $D = D_1 \# D_2$. In particular, a rectangular diagram D is composite if there is a vertical or horizontal line intersecting D in exactly two points such that on both its sides there are at least four vertices of D.

As we shall see in 1.3, rectangular diagrams are just another way of thinking about arc-presentations. Thus, the assertion of Theorem 1 is actually this. For any given rectangular diagram D, one can find a sequence of elementary moves $D \mapsto D_1 \mapsto \cdots \mapsto D_N$, not including stabilization, such that the final diagram D_N is obtained by the connected sum and distant union operations from diagrams of prime non-trivial non-split links and trivial diagrams.

The idea of specifying arc-presentations by drawing the corresponding rectangular diagrams is not new. It is already present in P. Cromwell's work [6], where he calls them "loops and lines" diagrams and uses them to show that any link has an arc-presentation. A similar result is proved in the manuscript [17] for graphs embedded in $\mathbb{R}^3$ by using an appropriate generalization of the notion of rectangular diagram.

1. PRELIMINARIES

1.1. Definition of arc-presentations. We shall regard a three-dimensional sphere S^3 as the join of two circles: $S^3 = S^1 * S^1$, and use the coordinate system (φ, τ, θ) on S^3, where φ and θ are coordinates on the circles, $\varphi, \theta \in \mathbb{R}/(2\pi\mathbb{Z})$, and τ takes values in the interval $[0, 1]$. We have $(\varphi, 0, \theta_1) \sim (\varphi, 0, \theta_2)$ and $(\varphi_1, 1, \theta) \sim (\varphi_2, 1, \theta)$ for all $\varphi, \varphi_1, \varphi_2, \theta, \theta_1, \theta_2 \in S^1$. In order to distinguish between the circles $\tau = 0$ and $\tau = 1$, we shall denote them by S^1_φ and S^1_θ, respectively. The circle S^1_φ will be called the *binding circle*. We shall denote by $\mathcal{D}_t$ the open disk defined by $\theta = t$, $\tau > 0$. Such a disk will be called a *page*. We have $\partial \mathcal{D}_t = S^1_\varphi$.

We shall regard φ and θ as functions on S^3. As a rule, we use the notation s, s', s_1, $s_2, \ldots$ for points from S^1_φ and t, t', t_1, $t_2, \ldots$ for points from S^1_θ.

Let L be a link in S^3. By an *arc-presentation* of L we shall mean a link L' in S^3 isotopic to L such that the set $L' \cap S^1_\varphi$, whose elements are called *vertices*, is finite, and, for any $t \in S^1$, the intersection of L' with any page $\mathcal{D}_t$ is either empty or an open arc approaching two distinct vertices.

REMARK 1. The term "arc-presentation" was introduced by J. Birman and W. Menasco, though the object itself appeared already in [5]. However, it received almost no attention before P. Cromwell's paper [6].

1.2. Duality of arc-presentations. Two arc-presentations, say L_1 and L_2, are not distinguished if they have the same set of vertices, and, for any $t \in S^1_\theta$, we have either $L_1 \cap \mathcal{D}_t = L_2 \cap \mathcal{D}_t = \emptyset$, or $L_1 \cap \mathcal{D}_t$ and $L_2 \cap \mathcal{D}_t$ are arcs with the same endpoints. Having agreed about this, we may assume without loss of generality that any arc $L \cap \mathcal{D}_t$ of an arc-presentation L consists of two radii of the disk $\mathcal{D}_t$. This means that L is a link consisting of segments of the form $P * Q \subset S^1_\varphi * S^1_\theta$, where $P \in S^1_\varphi$, $Q \in S^1_\theta$. Notice that, in this definition, the rôles of the circles S^1_φ and S^1_θ are the same, and L can be considered as an arc-presentation with respect to both circles. Vertices of one of the presentations are the centers of arcs of the other.

So, there is a duality operation on the set of arc-presentations, which is defined by the mapping $\xi : S^3 \to S^3$ written in our coordinate system

as $\xi(\varphi, \tau, \theta) = (\theta, 1 - \tau, \varphi)$. The rôles of arcs and vertices are interchanged under this duality.

REMARK 2. There is a triangulation $T(L)$ of S^3 associated naturally with any arc-presentation L having $n \geq 2$ vertices. $T(L)$ has n^2 simplices that have the form $I' * I''$, where $I' \subset S_{\varphi}^1$ (respectively, $I'' \subset S_{\theta}^1$) is an interval between neighbouring vertices of the arc-presentation L (respectively, $\xi(L)$). In this construction, L is a subset of the 1-skeleton of $T(L)$ containing all the vertices. The main result of this paper can be proved by using normal surfaces if one chooses $T(L)$ as the main triangulation. In this case, the reasoning becomes "self-dual" with respect to ξ. However, it seems to us that Birman–Menasco's approach, which "breaks the symmetry", is easier to use in our case.

1.3. Arc-presentations and rectangular diagrams

PROPOSITION 1. *Let L be an arc-presentation of a link*, *and D a rectangular diagram* (*see the Introduction*) *with vertices at all points $(x, y) \in [0, 2\pi) \times [0, 2\pi)$ such that $L \cap \mathcal{D}_y$ is an arc one of whose endpoints is $(x, 0, 0) \in S_{\varphi}^1$. Then D presents the same link as L does.*

The correspondence $L \mapsto D$ between arc-presentations and rectangular diagrams in $[0, 2\pi) \times [0, 2\pi)$ is one-to-one.

Proof. The definition of D makes sense, since each arc has two endpoints, and two arcs are attached to each vertex. Thus, in any vertical or horizontal straight line in the (x, y)-plane, we will have either none or two vertices. In the latter case, we connect them by a straight line segment, obtaining the rectangular diagram D.

It is easy to see that one can obtain D from L by cutting S^3 along disks $\mathcal{D}_0$ and $\xi(\mathcal{D}_0)$ and then projecting the result appropriately to the plane $\tau = 1/2$. For more explanations, see [6].

The last assertion of the proposition is obvious. ∎

We shall regard rectangular diagrams as a convenient way for depicting arc-presentations. Vertical edges of a rectangular diagram correspond to vertices, and horizontal edges to arcs of an arc-presentation. We shall often draw appropriate rectangular diagrams to illustrate properties or transforms of arc-presentations.

1.4. Arc-presentations and closed braids.
There is also an easy way to convert a rectangular diagram (an arc-presentation) to a closed braid. An arc presentation or a rectangular diagram can be endowed with an orientation as an ordinary link or an ordinary planar link diagram, respectively. For an oriented rectangular diagram D, we shall call a horizontal edge of D *negative* if its orientation is opposite to the orientation of the x-axis.

PROPOSITION 2. *Let D be an oriented rectangular diagram in $[0, 2\pi) \times [0, 2\pi)$. Replace each negative horizontal edge in D of the form $[x_1, x_2] \times y$, $x_1 < x_2$, by the union of two straight line segments $([-\varepsilon, x_1] \cup [x_2, 2\pi]) \times y$, where $\varepsilon > 0$. As before, at each crossing, we interpret the vertical arc as overcrossing and the horizontal one as undercrossing. After a small deformation of the resulting picture, we can get a planar diagram of a braid, say b_D. Then the closure of b_D is a link equivalent to the one defined by D.*

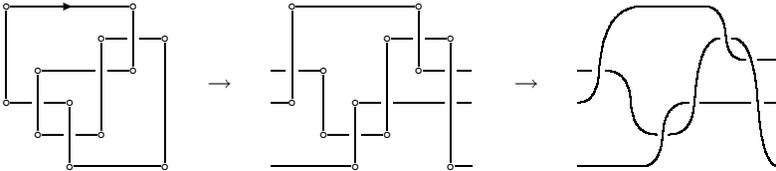*Any diagram of a braid can be obtained from some rectangular diagram by the method just described.*

Fig. 5. Converting a rectangular diagram to a braid

The idea is illustrated in Fig. 5. We skip the easy proof. See also [6].

1.5. Elementary moves. Actually, elementary moves of arc-presentations have already been introduced in the language of rectangular diagrams (see the Introduction). Here we describe them directly.

Throughout this section, we assume that each arc of any arc-presentation L consists of two radii of a disk $\mathcal{D}_t$. Thus, we have $t \in S^1_\theta \cap L$ if and only if the page $\mathcal{D}_t$ contains an arc of L.

Suppose that we have a continuous family $L(u)$ of arc presentations, i.e., such that the sets $L(u) \cap S^1_\varphi$ and $L(u) \cap S^1_\theta$ depend on u continuously. Then all the links $L(u)$ are isotopic to each other via isotopies of the form $(\varphi, \tau, \theta) \mapsto (f(\varphi), \tau, g(\theta))$. Let us see what happens to the corresponding rectangular diagrams $D(u)$ when u changes. If no point in $L(u) \cap S^1_\varphi$ and $L(u) \cap S^1_\theta$ passes through the origin $0 \in S^1$, the diagram $D(u)$ also changes continuously, and the combinatorial type of $D(u)$ stays unchanged. If one of the points in $L(u) \cap S^1_\varphi$ (respectively, $L(u) \cap S^1_\theta$) passes through the origin, the corresponding rectangular diagram is changed by a cyclic permutation of vertical (respectively, horizontal) edges. Two arc-presentations that can be included into such a continuous family are said to be *combinatorially equivalent*.

The following assertion, although quite obvious, is very important from the algorithmic point of view.

PROPOSITION 3. *For any n, there are only finitely many classes of combinatorially equivalent arc-presentations of complexity n.*

REMARK 3. One can show that the number $N(n)$ of pairwise distinct classes of combinatorially equivalent arc-presentations of complexity n satisfies the inequalities $((n-1)!)^2/(2n) < N(n) \leq ((n-1)!)^2$.

For three distinct points x_1, x_2, x_3 of S^1, we shall write $x_2 \in (x_1, x_3)$ if, when going in the positive direction of S^1 we meet the points x_i in this order: $\ldots, x_1, x_2, x_3, x_1, x_2, \ldots$. We shall write $x_2 \in [x_1, x_3]$ if we have either $x_2 \in (x_1, x_3)$ or $x_2 \in \{x_1, x_3\}$. By $\gamma_{t,s,s'}$ we shall denote an arc in $\mathcal{D}_t$ such that $\partial \gamma_{t,s,s'} = \{s, s'\} \subset S^1_\varphi$. When using this notation, we shall not distinguish between two arcs lying in the same page and having the same endpoints. Thus, we write $\gamma_{t,s,s'} \subset L$ whenever we want to say that $L \cap \mathcal{D}_t$ contains an arc with endpoints s, s'.

Let $\alpha = \gamma_{t,s_1,s_2}$ be an arc of an arc-presentation L. For sufficiently small $\varepsilon > 0$, there is no arc of L other than α in $\mathcal{D}_{t'}$ if $t' \in (t - \varepsilon, t + \varepsilon)$, and there is no vertex of L other than s_1 in $(s_1 - \varepsilon, s_1 + \varepsilon) \subset S^1_\varphi$. Replacing α by the arcs $\gamma_{t,s_1+\varepsilon_1,s_2}$ and $\gamma_{t+\varepsilon_2,s_1,s_1+\varepsilon_1}$, where $\varepsilon_{1,2} \in \{\varepsilon, -\varepsilon\}$, will be called a *stabilization move*. The inverse operation will be called a *destabilization move*. In the language of rectangular diagrams, these operations coincide with stabilization and destabilization moves of rectangular diagrams, respectively, provided that $0 \notin (s_1 - \varepsilon, s_1 + \varepsilon)$ and $0 \notin (t - \varepsilon, t + \varepsilon)$, which can always be achieved by a small perturbation of L and taking ε sufficiently small.

Notice that a (de)stabilization move is self-dual with respect to ξ.

Let α_1, α_2 be two arcs of an arc-presentation L, $\alpha_i \subset \mathcal{D}_{t_i}$, and suppose there is no arc of L in $\bigcup_{t \in (t_1, t_2)} \mathcal{D}_t$ or in $\bigcup_{t \in (t_2, t_1)} \mathcal{D}_t$. In this case, we shall say that the arcs α_1 and α_2 are *neighbouring*.

Let α_1 and α_2 be neighbouring arcs, $s_1, s_2 \in S^1_\varphi$ the endpoints of α_1, and s_3, s_4 be the endpoints of α_2, $\partial \alpha_1 \cap \partial \alpha_2 = \emptyset$. We say that the arcs α_1, α_2 are *non-interleaved* if we have either $s_1, s_2 \in (s_3, s_4)$ or $s_1, s_2 \in (s_4, s_3)$. Otherwise, they are said to be *interleaved*. In other words, arcs are interleaved if the corresponding chords of S^1 intersect each other (see Fig. 6).
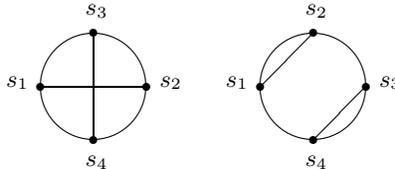
Fig. 6. Interleaved arcs (left) and non-interleaved arcs (right)

If $\alpha_1 \subset \mathcal{D}_{t_1}$, $\alpha_2 \subset \mathcal{D}_{t_2}$ are non-interleaved neighbouring arcs, we define an *arc exchange move* on L as replacing α_1, α_2 by arcs $\alpha'_1 \in \mathcal{D}_{t_2}$, $\alpha'_2 \in \mathcal{D}_{t_1}$ such that $\partial \alpha'_i = \partial \alpha_i$. This operation corresponds to interchanging horizontal

edges of a rectangular diagram, probably combined with a cyclic permutation of horizontal edges.

We define a *vertex exchange move* to be the operation dual via ξ to an arc exchange move. It corresponds to interchanging vertical edges of a rectangular diagram, probably combined with a cyclic permutation of vertical edges. We shall use the term "exchange move" for both operations, an arc exchange move and a vertex exchange move.

In what follows, stabilization, destabilization, and exchange moves will be referred to as *elementary* moves. When appropriate, we shall not distinguish between combinatorially equivalent arc-presentations and think about elementary operations as being performed on the corresponding combinatorial classes.

PROPOSITION 4. *Any link in S^3 has an arc-presentation. Two arc-presentations L_1, L_2 of the same link can be obtained from each other by finitely many elementary moves.*

Proof. We only sketch the argument, as the details are very elementary and standard. (In [6], one can find a proof that uses Markov's theorem. The latter seems to us to be much less elementary than the assertion to be proved. We prefer to use Reidemeister's theorem.)

As already mentioned, any planar diagram of a link is isotopic to a rectangular diagram. Indeed, for a given planar diagram, one can first deform it near crossings so that overcrossing arcs become vertical and undercrossing arcs horizontal. Then one can approximate the rest of the diagram by a step-line consisting of vertical and horizontal straight line segments. In the generic case, there will be no collinear segments. In order to deform such a diagram into an isotopic one, we might need to add more edges or contract some, which is achieved by (de)stabilization moves, and pass through codimension one degenerations, when two non-overlapping edges become collinear. The latter is achieved by exchange moves.

In order to prove the second assertion, one shows first that rotating a crossing by π as shown in Fig. 7 can be performed by applying finitely
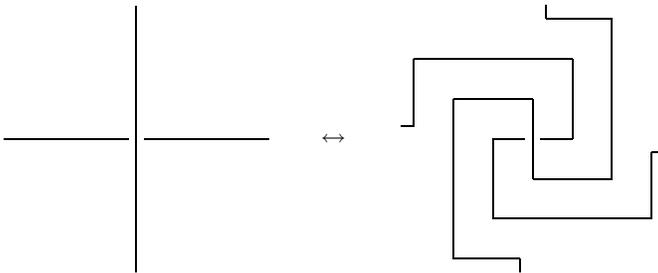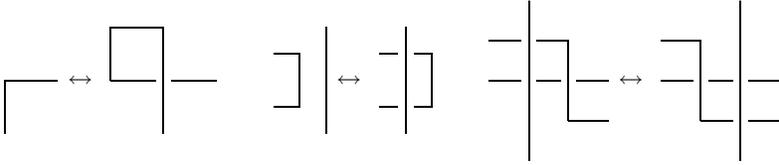

Fig. 7. Rotating a crossing

Fig. 8. A realization of the Reidemeister moves

many elementary moves. Then one does the same with Reidemeister moves, assuming that the crossings involved are already positioned as one wishes (see Fig. 8). ∎

REMARK 4. In [6], the set of allowed (de)stabilization-type moves in the formulations of an analogue of Proposition 4 is larger. We shall now see that those, more general, moves can be expressed in terms of our elementary moves.

1.6. Generalized moves. Let $\alpha \in \mathcal{D}_t$ be an arc of an arc-presentation L and $s \in S^1_\varphi$ be a point distinct from the vertices of L. Let $s_1, s_2 \in S^1_\varphi$ be the endpoints of α. Replacing α by the arcs $\gamma_{t,s_1,s}$ and $\gamma_{t+\varepsilon,s,s_2}$, where $\varepsilon > 0$ is sufficiently small, will be called a *generalized stabilization move*. (In [6], it is called a type IV move. The dual operation, which we shall not need, is called a type III move in [6].) The operation inverse to a generalized stabilization move is called a *generalized destabilization move*.

Let L be an arc-presentation, and s_1, s_2, s_3 three points of S^1_φ distinct from the vertices of L such that $s_2 \in (s_1, s_3)$. Suppose that for some $t_1, t_2 \in S^1_\theta$, the link L does not intersect any of the arcs γ_{t,s_1,s_2} if $t \in [t_1, t_2]$ and γ_{t,s_2,s_3} if $t \in [t_2, t_1]$. In this case, we shall call the operation consisting in interchanging the intervals (s_1, s_2) and (s_2, s_3) a *generalized (vertex) exchange move*. This operation can be described more formally as follows. Define a function $f : S^1_\varphi \to S^1_\varphi$ by

$$s \mapsto \begin{cases} s + s_3 - s_2 & \text{if } s \in (s_1, s_2), \\ s + s_1 - s_2 & \text{if } s \in (s_2, s_3), \\ s & \text{otherwise.} \end{cases}$$

Then the generalized exchange move just defined consists in replacing each arc of the form $\gamma_{t,s,s'}$ by the arc $\gamma_{t,f(s),f(s')}$. (In [6], this operation is called a generalized type I move.) The dual operation, a generalized arc exchange move, will not be needed here.

PROPOSITION 5. *Any generalized (de)stabilization move can be presented as the composition of one ordinary (de)stabilization move and a few exchange moves. Any generalized exchange move can be presented as the composition of ordinary exchange moves.*
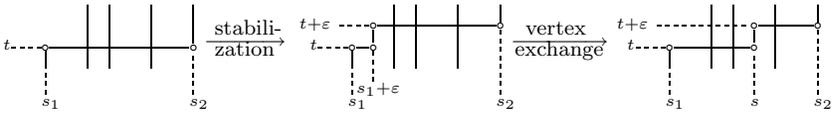
Fig. 9. Decomposition of a generalized stabilization move
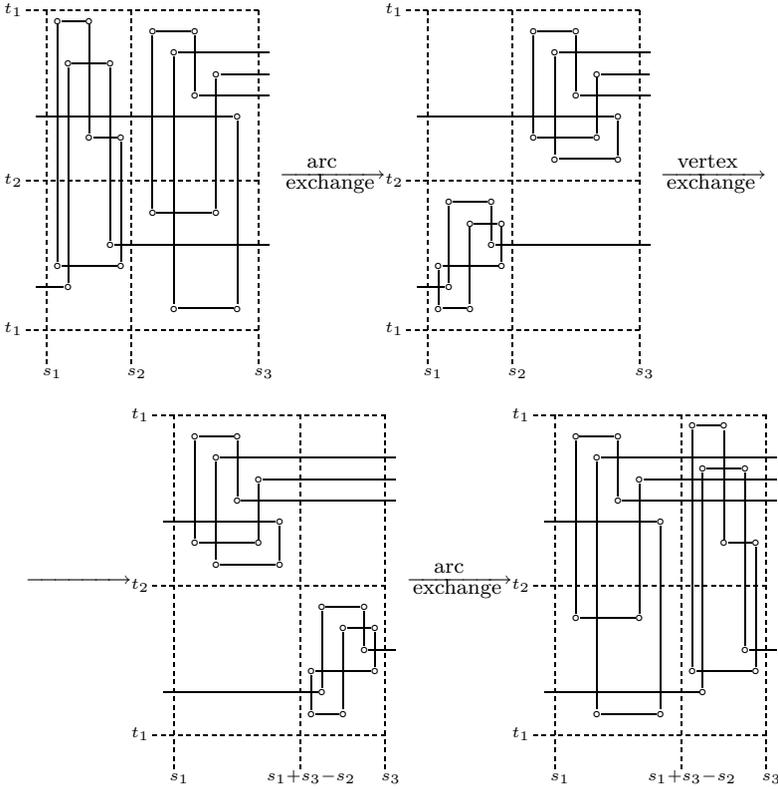
Fig. 10. Decomposition of a generalized exchange move

Proof. We illustrate the idea of the proof in Figures 9 and 10 by using corresponding rectangular diagrams. The details are easy. ∎

2. PROOF OF THE MAIN RESULT

Recall that a rectangular diagram D is said to be *trivial* if it consists of four edges. It is *split* if there is a vertical line l not intersecting the diagram such that D has edges on both sides of l.

In this section, a rectangular diagram D is said to be *composite* if there is a vertical line l intersecting D in two points such that D has horizontal

edges on both sides of l. This is a particular case of the definition given in the Introduction.

An arc-presentation whose rectangular diagram is trivial, split, or composite is called *trivial*, *split*, or *composite*, respectively.

The assertion of Theorem 1 will be a consequence of the following three propositions.

PROPOSITION 6. *If L is an arc-presentation of the unknot, then there exists a finite sequence of exchange and destabilization moves $L \mapsto L_1 \mapsto \cdots \mapsto L_n$ such that the arc-presentation L_n is trivial.*

PROPOSITION 7. *If L is an arc-presentation of a split link, then there exists a finite sequence of exchange and destabilization moves $L \mapsto L_1 \mapsto \cdots \mapsto L_n$ such that the arc-presentation L_n is split.*

PROPOSITION 8. *If L is an arc-presentation of a non-split composite link, then there exists a finite sequence of exchange and destabilization moves $L \mapsto L_1 \mapsto \cdots \mapsto L_n$ such that the arc-presentation L_n is composite.*

The proof of these three propositions follows the same scheme, only the technical details are different. So, we shall be proving all three statements simultaneously. The strategy and methods of the proof are very close to those in Birman–Menasco works [3], [4]. However, we need to concentrate on many technical details, which are specific to our case. This forces us to write out a complete proof instead of just using existing results.

To begin with, we describe the general outline of the proof.

1) The assumptions of Propositions 6–8 imply the existence of a certain surface M: a disk whose boundary is L, a two-sphere that does not intersect L and is such that there are non-trivial parts of L on both sides of M, or a two-sphere that intersects L in two points and cuts L in two non-trivial tangles. We show that such a surface M, which we call a *characteristic surface*, can be isotoped to satisfy certain restrictions. Compared to the previous works [3, 4, 6], a new thing here is only a restriction on the behaviour of the spanning disk near the boundary in the case of the unknot.

2) We consider the foliation $\mathcal{F}$ on M defined by the equation $d\theta = 0$. We define the *complexity* $c(M) \in \mathbb{Z}$ of the characteristic surface and show that, if $c(M) > 0$, then the foliation $\mathcal{F}$ contains certain patterns. In the case when M is a two-sphere we just follow [3, 4], and extend the argument to the case of the disk.

3) We show that, if $\mathcal{F}$ contains a pattern defined at the previous step, then there exists an arc-presentation L' and a surface M' such that L' is obtained from L by finitely many exchange and destabilization moves, the pair (L', M') is isotopic to (L, M), the surface M' satisfies

the restrictions introduced at Step 1, and we have either $c(L') < c(L)$
or $c(L') = c(L)$ and $c(M') < c(M)$. Similarly to [3, 4, 6], this is done
by using two tricks that are closely related to those introduced by
D. Bennequin in [1]. In the unknot case we also need one more, easier,
trick that allows us to simplify the boundary of a spanning disk.

4) We notice that, if $c(F) = 0$, then we are done, i.e., the arc-presenta-
tion L is trivial, split, or composite, respectively.

2.1. Characteristic surfaces. Let L be an arc-presentation satisfying
the assumptions of one of Propositions 6–8. From now on, we shall assume
that each arc of the form $L \cap \mathcal{D}_t$ is smooth, which was not the case under
the agreement of Section 1. Let $M \subset S^3$ be

- an embedded two-dimensional disk whose boundary is L if L is an
  arc-presentation of the unknot;
- an embedded two-sphere splitting L into two non-empty links if L is
  a split link;
- a *factorizing sphere*, i.e., an embedded sphere that meets L in two
  points, cutting L into two non-trivial tangles, if L is a composite non-
  split link.

In this case, we call M a *characteristic surface* for L.

We say that a characteristic surface M is *admissible* if

- the surface M is smooth everywhere up to the boundary, except at
  $(\partial M) \cap S^1_\varphi$;
- $M \setminus \partial M$ intersects the binding circle S^1_φ transversely in finitely many
  points;
- the foliation $\mathcal{F}$ on $M \setminus S^1_\varphi$ defined by $d\theta = 0$ has only finitely many
  singularities, which are points of tangency of M with pages $\mathcal{D}_t$;
- all singularities of $\mathcal{F}$ are of Morse type, i.e., local extrema (Fig. 11b) or
  saddle critical points (Fig. 11c) of the (multivalued) function $\theta|_{M \setminus S^1_\varphi}$;
- near any point $p \in (\partial M) \cap S^1_\varphi$, the foliation $\mathcal{F}$ is radial (see Fig. 11d);
- there is at most one point $p \in (\partial M) \cap S^1_\varphi$ at which $|\int_\gamma d\theta| > 2\pi$, where
  $\gamma \subset M$ is a proper arc in a small neighbourhood of p such that the
  endpoints of γ lie in ∂M on different sides of p. Such a point is called
  a *winding vertex*;
- there is at most one point $p \in (\partial M) \setminus S^1_\varphi$ at which the surface M is not
  transversal to the corresponding page $\mathcal{D}_{\theta(p)}$; at the exceptional point,
  the foliation $\mathcal{F}$ must have a saddle critical point (see Fig. 11e). If such
  a saddle and a winding vertex are both present, then the winding
  vertex is an endpoint of the edge containing the saddle;

- each page $\mathcal{D}_t$ contains no more than one of the following:
  - an arc of L;
  - a singularity of $\mathcal{F}|_{M \setminus \partial M}$;
- if L is a non-split composite link, the two intersection points $L \cap M$ are not vertices of L.

The behaviour of $\mathcal{F}$ near a point from $(M \setminus \partial M) \cap S_\varphi^1$ is shown in Fig. 11a. The points of intersection of M with the binding circle, where the foliation
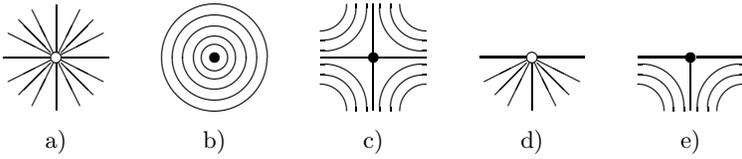

Fig. 11. Vertices and singularities of $\mathcal{F}$

$\mathcal{F}$ is not defined, will be called *vertices* of $\mathcal{F}$. Sometimes we will distinguish between *interior* and *boundary* vertices that lie in $M \setminus \partial M$ and ∂M, respectively. The singularities of $\mathcal{F}$ shown in Fig. 11b,c,e will be called a *pole*, a(n interior) *saddle*, and a *boundary saddle*, respectively.

LEMMA 1. *For any arc-presentation satisfying the assumptions of one of Propositions* 6–8, *there exists an admissible characteristic surface.*

Proof. In the case of a split link or a composite link, there is almost nothing to prove: a characteristic surface is admissible if it is in general position with respect to the binding circle, to the link, and to the foliation of $S^3 \setminus S_\varphi^1$ by disks $\mathcal{D}_t$.

In the case of unknot, the admissibility of M means something more than just general position. By a small perturbation of an arbitrary characteristic surface M, we can achieve that M satisfies almost all the restrictions. The only thing that we may not avoid in this way is that $\mathcal{F}$ may have *many* boundary saddles, whereas we allow it to have just one.

Notice that if the number of vertices of an arc-presentation of the unknot is odd, then the foliation on the spanning disk is forced to have a boundary saddle, since the parity of the number of boundary saddles should be equal to that of the number of boundary vertices. Indeed, the coorientation of ∂M defined by $\operatorname{grad} \theta$ flips at each vertex and at each boundary saddle, so in total there should be an even number of them.

In order to avoid the occurrence of many boundary saddles, we shall construct the disk M, starting from a small neighbourhood of the boundary. We construct initially a narrow ribbon $R \cong [0, 1] \times S^1$ such that $\partial R = L \cup L'$, where L' is a circle unlinked with L, and the foliation on R defined by $d\theta = 0$ has the desired behaviour near L. Then we attach a two-dimensional disk

along L', obtaining a characteristic surface M. Finally, we deform M slightly if necessary, keeping it fixed in a small neighbourhood of $L = \partial M$, so that, after deformation, $M \setminus \partial M$ will be in general position. So, the only thing to explain is how to construct the ribbon R.

Notice that constructing the ribbon R is *not* equivalent to specifying a trivial framing of L. It would have been so if L were a smooth curve.

We start by enumerating the vertices $s_1, \ldots, s_n \in S^1_\varphi$ and arcs $\alpha_1, \ldots, \alpha_n$ of L in the order that they follow in L:

$$\partial \alpha_i = \{s_i, s_{i+1}\} \quad \text{if } i = 1, \ldots, n-1, \quad \text{and} \quad \partial \alpha_n = \{s_n, s_1\}.$$

Denote by r_u the rotation of S^3 about the binding circle by the angle u: $r_u(\varphi, \tau, \theta) = (\varphi, \tau, \theta + u)$. For an arc $\alpha \subset \mathcal{D}_t$, we denote by $d_\varepsilon(\alpha)$ the disk

$$\bigcup_{u \in [0,\varepsilon]} r_u(\alpha) \quad \text{if } \varepsilon > 0, \quad \text{and} \quad \bigcup_{u \in [\varepsilon,0]} r_u(\alpha) \quad \text{if } \varepsilon < 0.$$

Take a point p in the arc α_n. Let α'_n, α''_n be the two parts of the arc α cut by p: $\partial \alpha'_n = \{s_n, p\}$, $\partial \alpha''_n = \{p, s_1\}$. Pick an $\varepsilon > 0$ which is smaller than the θ-distance between any two neighbouring arcs. Consider the union R_1 of the following disks:

$$d_\varepsilon(\alpha_1), \, d_{-\varepsilon}(\alpha_2), \, d_\varepsilon(\alpha_3), \, \ldots, \, d_{(-1)^n \varepsilon}(\alpha_{n-1}), \, d_{(-1)^{n+1}\varepsilon}(\alpha'_n), \, d_{-\varepsilon}(\alpha''_n).$$

If n is even, the last two disks form the disk $d_{-\varepsilon}(\alpha_n)$. If n is odd, then the intersection of R_1 with a neighbourhood of the point p will look as shown in Fig. 12 (left). We attach a small disk "perpendicular" to α_n along the
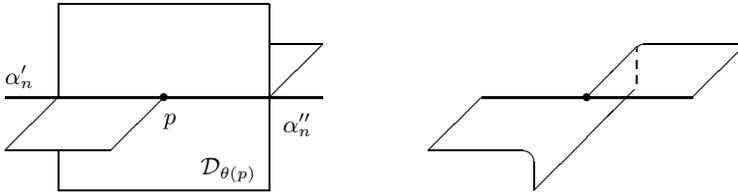
Fig. 12. The ribbon R near a boundary saddle

segment $\{r_u(p)\}_{u \in [-\varepsilon, \varepsilon]}$ and then smooth the obtained surface as shown in Fig. 12 (right), still denoting the result by R_1. The point p will be the only boundary saddle of $\mathcal{F}$.

The union R_1 of disks is only a part of the ribbon under construction. Near any vertex, R_1 looks as in Fig. 13 (left). We transform it into a ribbon as shown in Fig. 13 (middle), attaching a sector of a small disk transversal to the binding circle. If we do this near each vertex, then, possibly, the ribbon obtained will be twisted, that is, the connected components of ∂R will be linked. We can compensate this by twisting the ribbon around the binding circle "in the opposite direction" at an arbitrarily chosen vertex as shown in

Fig. 13. The ribbon R near a vertex of L

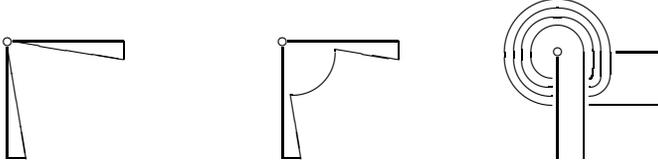Fig. 13 (right). This will create a winding vertex, so, if a boundary saddle is present, we do this at one of the boundary vertices adjacent to the saddle. ∎

2.2. Moves of the characteristic surface. We will think about an admissible characteristic surface M of an arc-presentation L as a periodic flow of planar pictures that are intersections of M with the pages $\mathcal{D}_t$, $t \in S^1$. The pages $\mathcal{D}_t$ containing singularities of $\mathcal{F}$ and arcs of L will be called *singular*, and the corresponding values of t *critical*. All the others are *regular*.

For simplicity, we assume in this subsection that $\mathcal{F}$ has no poles and no closed regular fibres. The notions and claims of this subsection will be used only when poles and closed fibres are absent.

If $\mathcal{D}_t$ is a regular page, then the intersection $\mathcal{D}_t \cap (L \cup M)$ is the union of pairwise non-intersecting arcs having endpoints on the binding circle $S^1_\varphi = \partial \mathcal{D}_t$. Notice that two arcs in $\mathcal{D}_t \cap M$ may have a common endpoint if L is an arc-presentation of the unknot (refer to the right picture in Fig. 13), but no arc forms a loop, which follows from an orientation argument. Each section $\mathcal{D}_t \cap (L \cup M)$ will be regarded up to a homeomorphism of $\overline{\mathcal{D}_t}$ fixed on $\partial \mathcal{D}_t$, which we shall express by saying that we consider the *combinatorial type* of the section $\mathcal{D}_t \cap (L \cup M)$. This combinatorial type can be specified by listing pairs of endpoints of all arcs in $\mathcal{D}_t$, like: $s_{i_1}s_{i_2}$, $s_{i_3}s_{i_4}$, $s_{i_5}s_{i_6}$, ... The notation $s_i s_j$ will be used for any arc with endpoints $s_i, s_j \in S^1_\varphi$. Notice that $s_i s_j = s_j s_i$, but $(s_i, s_j) \cap (s_j, s_i) = \emptyset$. We shall also refer to the combinatorial type of a section $\mathcal{D}_t \cap (L \cup M)$ as a *state* of (L, M).

We leave the following statement, which is standard, without proof.

PROPOSITION 9. *Let M and M' be admissible characteristic surfaces for arc-presentations L and L', respectively, such that we have $L \cap S^1_\varphi = L' \cap S^1_\varphi$, $M \cap S^1_\varphi = M' \cap S^1_\varphi$, and for all $t \in S^1_\theta$, the sections $\mathcal{D}_t \cap (L \cup M)$ and $\mathcal{D}_t \cap (L' \cup M')$ have the same combinatorial type. Then there is an isotopy f fixed on S^1_φ, preserving each page $\mathcal{D}_t$, and taking (L, M) to (L', M').*

If $\mathcal{D}_t$ is a regular page, then the combinatorial type of the section $\mathcal{D}_t \cap (L \cup M)$ does not change if t varies slightly. In addition to simple arcs, a singular page may contain one of the following:

1) a *T-joint*, which is a boundary saddle of $\mathcal{F}$ with three arcs attached;

2) an *X-joint*, which is either an interior saddle of $\mathcal{F}$ with four arcs attached or a regular fibre of $\mathcal{F}$ crossed by an arc of L (in the case of a composite link).

When t passes a critical value, the state of (L, M) changes. We shall call such a change an *event*. Below is the list of all possible events.

Event	Singularity	Case
a single arc appears and immediately disappears	an arc of L	split link, composite link
a single arc appears	an arc of L	unknot
a single arc disappears	an arc of L	unknot
an arc $s_1 s_2$ is replaced by some arc $s_2 s_3$	an arc $s_1 s_3$ of L containing a boundary saddle	unknot
two arcs $s_1 s_2$, $s_3 s_4$ are replaced by $s_1 s_4$, $s_2 s_3$	a saddle of $\mathcal{F}$	any

If an arc with endpoints $s, s' \in S_\varphi^1$ is contained in $\mathcal{D}_t \cap (L \cup M)$, where $\mathcal{D}_t$ is a singular page, but there is no arc with endpoints s, s' "just before" or "just after" the corresponding events, i.e., in the section $\mathcal{D}_{t-\varepsilon} \cap (L \cup M)$ or $\mathcal{D}_{t+\varepsilon} \cap (L \cup M)$, where ε is sufficiently small, we shall say that the arc ss' *participates* in the corresponding event.

Clearly, if we know the state of (L, M) just before an event and have a description of the event in terms of the participating arcs, then we can recover the state of (L, M) right after the event. So, in order to specify the combinatorial type of an arc-presentation L with a characteristic surface M, we must describe the state of (L, M) at just one non-critical moment and provide the ordered list of all the events that occur when t runs over the circle. Such a state with a list of events will be called a *combinatorial description* of (L, M).

We shall use combinatorial descriptions whenever we will need to change the surface M or an arc-presentation L. Of course, an arbitrary initial state equipped with an ordered list of events does not necessarily define an arc-presentation and a characteristic surface; the combinatorial data must be *consistent*. It is easy, but tedious, to list all the requirements for the data to be consistent. Here are some of them: the final state (to which we come after all the events) must coincide with the initial one, an arc that is supposed to disappear in a forthcoming event must be present just before the event, etc. In each case when we change the combinatorial description of the pair (L, M), it will be clear that the combinatorial data remain consistent.

LEMMA 2. *Let E_1, E_2 be successive events such that no arc participating in E_1 coincides or interleaves with an arc participating in E_2. If both events*

correspond to arcs of L, we assume additionally that those arcs do not have a common endpoint. Then the events can be interchanged, which will have no effect on the foliation $\mathcal{F}$ and the isotopy class of (L, M). If both singular pages corresponding to the events E_1, E_2 contain an arc of L, then the exchange of events will result in one arc exchange move on L. Otherwise, there will be no effect on the combinatorial type of L.

Proof. More precisely, the first assertion means the following. Let $t_1 = \theta(E_1)$ and $t_2 = \theta(E_2)$ be the time of the events. Then there exists an arc-presentation L' and an admissible characteristic surface M' for L' such that, for all $t \in (t_2, t_1)$, the states of (L, M) and (L', M') at time t coincide. At moments t_1, t_2, the events E_2, E_1, respectively, happen with $\mathcal{D}_t \cap (L' \cup M')$, and there is no event in between.

By saying that a transform has no effect on the foliation $\mathcal{F}$, we mean that the new foliation $\mathcal{F}'$ is equivalent to $\mathcal{F}$ via a homeomorphism $M \to M'$.

The assumptions of the lemma imply that there exists a continuous family of arcs $\gamma_{t,s_1,s_2} \subset \mathcal{D}_t$, $t \in (t_1 - \varepsilon, t_2 + \varepsilon)$, $\partial\gamma_{t,s_1,s_2} = \{s_1, s_2\} \subset S^1_\varphi$, forming an open disk d that does not intersect the surface M and the link L, such that the event E_1 occurs "on the right" of the disk d and the event E_2 occurs "on the left" of d. The latter means, in particular, that all arcs participating in E_1 (respectively, E_2) have both endpoints in the interval $[s_1, s_2]$ (respectively, in $[s_2, s_1]$). It is not forbidden that an arc participating in the events has one or both endpoints in $\{s_1, s_2\}$. If s_1 and s_2 are the endpoints of an arc participating in E_1 or E_2, the interior of the arc must be on the appropriate side of γ_{t,s_1,s_2}, where the corresponding event occurs.

Now we can define a self-homeomorphism f of $S^3 \setminus d$ preserving the foliation $d\theta = 0$ by making the time θ "go faster" on the left of d and "go slower" on the right of d, which will result in E_2 occurring before E_1. In the pages $\mathcal{D}_t$ with $t \in [t_2 + \varepsilon, t_1 - \varepsilon]$ the mapping f is assumed to be identical. The mapping f will send the surface M to another one, preserving the foliation. Clearly, the mapping f can be adjusted in a small neighbourhood of d to become a self-homeomorphism of the whole S^3.

The last two assertions of the lemma are obvious. ∎

REMARK 5. We do *not* exclude the case when an arc participating in E_1 has a common endpoint with one participating in E_2. This may occur when L is an arc-presentation of the unknot, $\partial M = L$, and the common end of two arcs is a winding vertex of $\mathcal{F}$ (refer again to Fig. 13). This is why we wrote $[s_1, s_2]$ and $[s_2, s_1]$ instead of (s_1, s_2), (s_2, s_1) in the above proof.

Whereas the transform described in Lemma 2 generalizes an arc exchange move to the case of an arc-presentation endowed with an admissible characteristic surface, the following construction is an analogue of a generalized exchange move.

LEMMA 3. *Let s_1, s_2, s_3 be three points of S_φ^1 disjoint from L and M such that $s_2 \in (s_1, s_3)$. Suppose that there exist time instances $t_1, t_2 \in S_\theta^1$ such that no arc in $\mathcal{D}_t \cap (L \cup M)$ interleaves with $s_1 s_2$ if $t \in [t_1, t_2]$ and with $s_2 s_3$ if $t \in [t_2, t_1]$. Then one can interchange all the vertices of L and $\mathcal{F}$ lying in (s_1, s_2) with those in (s_2, s_3), keeping the relative order of vertices in each of the intervals fixed, without changing the foliation $\mathcal{F}$ and the isotopy class of (L, M).*

Interchanging the vertices should be understood in the same way as described in 1.6.

Proof. Denote by $d_{t_1, t_2, s_1, s_2, s_3}$ a closed two-dimensional disk whose interior consists of the vertex s_2 and the following four disks d_1, d_2, d_3, d_4 adjacent to s_2:

$$d_1 = \bigcup_{t \in (t_1, t_2)} \gamma_{t, s_1, s_2}, \qquad d_2 = \bigcup_{t \in (t_2, t_1)} \gamma_{t, s_2, s_3},$$

and d_3 and d_4 are the two "triangles" bounded by $\gamma_{t_i, s_1, s_2}, \gamma_{t_i, s_2, s_3}, \gamma_{t_i, s_1, s_3}$, where $i = 1, 2$.

The assumptions of the lemma imply that the disk

$$d^{(1)} = d_{t_1, t_2, s_1, s_2, s_3}$$

is disjoint from L and M. Let L' and M' be the arc-presentation and the surface obtained after the transform. Then the disk

$$d^{(2)} = d_{t_2, t_1, s_1, s_1 + s_3 - s_2, s_3}$$

is disjoint from L' and M'. It is easy to see that there exists a homeomorphism $f : S^3 \setminus d^{(1)} \to S^3 \setminus d^{(2)}$ preserving each page $\mathcal{D}_t$ and interchanging the intervals (s_1, s_2) and (s_2, s_3) of the binding circle. We have $f(L) = (L')$ and $f(M) = M'$. There is a small neighbourhood U of $d^{(1)}$ that does not meet L and M. By adjusting f in U we can make f a self-homeomorphism of the whole sphere S^3. ∎

2.3. Patterns of $\mathcal{F}$ that are always present. Let L be an arc-presentation, M an admissible characteristic surface of L, and $\mathcal{F}$ the foliation on M defined by $d\theta = 0$. We define the *complexity* $c(M)$ of M as the total number of singularities of $\mathcal{F}$.

Fibres of $\mathcal{F}$ are of the following types:

- a closed circle;
- an open arc connecting two vertices;
- an open arc connecting a vertex to a saddle or a saddle to itself (we call such a fibre a *separatrix*);
- a singular point of $\mathcal{F}$.

Two regular fibres connecting the same pair of vertices will called *parallel* if they are included into a continuous family of such fibres (i.e. enclose a disk with no singularity inside). Any other two fibres connecting vertices will be called *non-parallel*.

For a vertex v of $\mathcal{F}$, we call the closure of the union of all fibres of $\mathcal{F}$ approaching v the *star* of v. By the *valence* of v we mean the number of separatrices in the star of v, which coincides with the number of saddles in the star.

An interior vertex v of $\mathcal{F}$ is said to be *bad* in three cases: 1) the star of v contains at least two non-parallel fibres each connecting v to a boundary vertex; 2) the star of v is pierced by L; 3) the star of v contains a winding vertex. If a vertex is not bad, it is said to be *good*.

Each saddle x will be endowed with an *orientation*, which is the circular order of the separatrices approaching x induced by the order of the corresponding vertices on the axis S^1_φ (see Fig. 14).
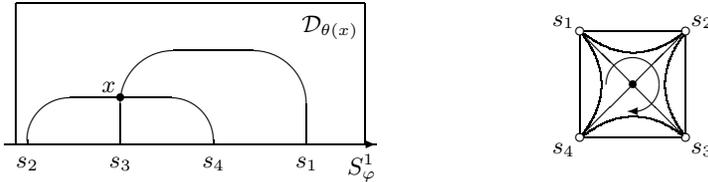

Fig. 14. Orientation of a saddle

Now we consider the star of a winding vertex in more detail.

REMARK 6. It was brought to the author's attention by Bill Menasco and Adam Sikora that the presence of a winding vertex in the star of a good two-valent vertex may produce an obstruction to applying the corresponding generalized exchange move in the proof of Lemma 6. The definitions of the characteristic surface and a bad vertex have been changed, the new Lemma 4 below has been added, Lemma 5 and the proof of Lemma 6 have been modified by introducing new cases 6–8 in order to correct the mistake in the earlier version of the paper.

Suppose that M has a winding vertex. Denote it by s^{w} and its star by σ^{w}. Let $s^{\mathrm{w}}_1, \ldots, s^{\mathrm{w}}_q$ be the vertices, and $x^{\mathrm{w}}_1, \ldots, x^{\mathrm{w}}_{q-1}$ the interior saddles in $\partial\sigma^{\mathrm{w}} \setminus s^{\mathrm{w}}$ numbered so that x^{w}_i lies between s^{w}_i and s^{w}_{i+1}, $i = 1, \ldots, q-1$. If there is a boundary saddle, we assume that it is adjacent to s^{w}_q and denote it by x^{w}_q. Thus, the valence of s^{w} is equal to

$$q' = \begin{cases} q - 1 & \text{if there is no boundary saddle,} \\ q & \text{otherwise.} \end{cases}$$

Note that if a vertex from σ^{w} is connected to s^{w} by two or more non-parallel fibres, it will appear in the sequence $s^{\mathrm{w}}_1, \ldots, s^{\mathrm{w}}_q$ more than once. So, some of the s^{w}_i may coincide.

We also introduce the *orientation* of the winding vertex, which is either
a locally θ-increasing or a locally θ-decreasing linear ordering $\prec$ of fibres
approaching s^w. (This orientation is not the one induced by the orientation
of S_θ^1, which is often used in the work of Birman and Menasco.) Take an
arbitrary page $\mathcal{D}_t$ such that at least two arcs in $\mathcal{D}_t \cap M$ have s^w as an
endpoint. Let $\alpha = \gamma_{t,s^w,s_i^w}$, $\beta = \gamma_{t,s^w,s_j^w}$, be such arcs. We orient s^w so as
to have $\alpha \prec \beta$ if and only if $s_i^w \in (s^w, s_j^w)$ (see Fig. 15). This orientation
does not depend on the choice of the page $\mathcal{D}_t$ and of the pair α, β in $\mathcal{D}_t$,
and can be described informally as follows. A small part of the surface M
near the winding vertex forms a "screw" whose axis is S_φ^1. The orientation
of s^w indicates which way this screw must rotate in order to screw up in the
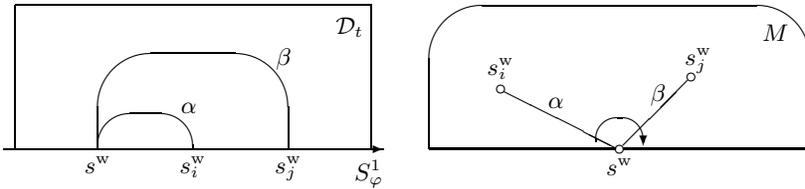direction opposite to the orientation of S_φ^1.

Fig. 15. Orientation of the winding vertex

Now we put $\delta_i = +1$ if the saddle x_i^w is coherently oriented with the
winding vertex, and $\delta_i = -1$ otherwise.

LEMMA 4. *If two successive interior vertices s_i^w and s_{i+1}^w in the star of
the winding vertex are both two-valent, then $\delta_i = +1$.*

Proof. Let s be the vertex in the stars of s_i^w and s_{i+1}^w different from s^w.
The union of these stars is shown in Fig. 16 (left). Let ψ_1 be the "θ-angle"
between the separatrices $s^w x_{i-1}^w$ and $s^w x_{i+1}^w$ at s^w, i.e., the absolute value
of the integral of $d\theta$ along an arc indicated in the picture. Similarly, let ψ_2
be the θ-angle between the separatrices $s x_{i-1}^w$ and $s x_{i+1}^w$ at s.
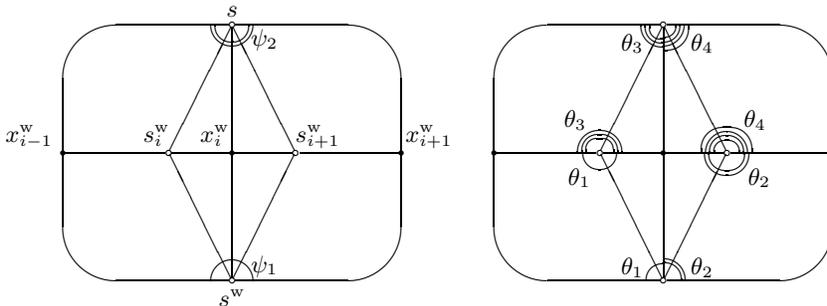

Fig. 16. Two successive two-valent vertices in σ^w ($\psi_1 = \theta_1 + \theta_2$, $\psi_2 = \theta_3 + \theta_4$, $\theta_1 + \theta_3 = \theta_2 + \theta_4 = 2\pi$)

It is not hard to see that $\psi_1 + \psi_2 = 4\pi$ (see Fig. 16). Since s is not a winding vertex, we have $\psi_2 < 2\pi$. Therefore, $\psi_1 > 2\pi$, which implies that some page $\mathcal{D}_t$ contains both arcs $\gamma_{t,s^{\mathrm{w}},s_i^{\mathrm{w}}}$ and $\gamma_{t,s^{\mathrm{w}},s_{i+1}^{\mathrm{w}}}$.

If $s_i^{\mathrm{w}} \in (s^{\mathrm{w}}, s_{i+1}^{\mathrm{w}})$, then the orientations of the saddle x_i^{w} and of the winding vertex are both clockwise (according to Fig. 16). If $s_{i+1}^{\mathrm{w}} \in (s^{\mathrm{w}}, s_i^{\mathrm{w}})$, then both orientations are counterclockwise. So, in both cases they are coherent. ∎

The following lemma is a generalization of Lemma 2 of [3]. (Although closed braids are now replaced with arc-presentations, the argument of [3] does not need to be changed to establish the claim below in the case of a split or composite link.)

Lemma 5. *If $c(M) > 0$, then the foliation $\mathcal{F}$ has at least one of the following*:

1) *a pole*;
2) *a univalent boundary vertex*;
3) *a good two-valent interior vertex*;
4) *a good three-valent interior vertex*;
5) *two pairs of two-valent vertices, in each the vertices are connected by a fibre of $\mathcal{F}$ intersecting L*;
6) *a winding vertex and a three-valent interior vertex s_i^{w} in σ^{w}*;
7) *a winding vertex and a two-valent boundary vertex s_i^{w} in σ^{w} such that $\delta_{i-1} = \delta_i$*;
8) *a winding vertex and a four-valent interior vertex s_i^{w} in σ^{w} such that $\delta_{i-1} = \delta_i$*.

Proof. We denote by χ the Euler characteristic of M, $\chi \in \{1, 2\}$. We also set $\epsilon = 1$ if $\mathcal{F}$ has a boundary saddle, and $\epsilon = 0$ otherwise. Denote the number of interior vertices of valence k by V_k, the number of boundary vertices of valence k by V_k^{b}, the number of interior saddles by S, and the number of poles by P. By the definition of complexity, we have

$$(2) \qquad\qquad S + P + \epsilon = c(M).$$

Notice that we always have $V_1 = 0$. Indeed, the opposite would mean that there is an event E in the flow $\mathcal{D}_t \cap (L \cup M)$ corresponding to a saddle such that some arc participating in E and existing just before the event survives after the event, which is impossible. Notice also that either $V_0 > 0$ or $V_0^{\mathrm{b}} > 0$ implies $c(M) = 0$: if there is no singularity in the star of a vertex, then this star is the whole surface M, and there are exactly two vertices, either interior or boundary ones.

Thus, we may assume that $V_0 = V_0^{\mathrm{b}} = V_1 = 0$. If $P > 0$, then case 1 takes place, and we are done. So, we assume $P = 0$.

Counting the topological indices of singularities gives

$$\sum_k V_k + \frac{1}{2}\sum_k V_k^{\mathrm{b}} - S - \frac{\epsilon}{2} = \chi. \tag{3}$$

Counting the number of separatrices in two different ways gives

$$\sum_k k(V_k + V_k^{\mathrm{b}}) = 4S + 3\epsilon. \tag{4}$$

From (3) and (4) we get

$$\sum_k (4-k)V_k + \sum_k (2-k)V_k^{\mathrm{b}} = 4\chi - \epsilon. \tag{5}$$

(i) First, consider the case of a split link, which is the simplest one because, in this case, there are no boundary vertices and no bad vertices. Relation (5) can be rewritten as

$$2V_2 + V_3 = 8 + \sum_{k\geq 4}(k-4)V_k, \tag{6}$$

which implies

$$2V_2 + V_3 \geq 8. \tag{7}$$

We have either $V_2 > 0$ or $V_3 > 0$, which are cases 3 and 4 of the lemma.

(ii) The case of a composite link is similar, but now the inequality (7) does not imply that there exists a *good* three- or two-valent vertex, because some vertices are bad. There may be at most four bad vertices. So, either there exists a good three- or two-valent vertex or we have $V_2 + V_3 \leq 4$. In conjunction with (7), the latter implies $V_2 = 4$, $V_3 = 0$, i.e., all bad vertices are two-valent. This is case 5 of the lemma.

REMARK 7. In this case, we also have $V_5 = V_6 = \cdots = 0$, but not necessarily $V_4 = 0$. One can show that V_4 must be even, but it *can* be arbitrarily large. The case $V_4 > 0$ is missing in [3] and also in [6].

(iii) Now we consider the case of the unknot. Relation (5) reads

$$2V_2 + V_3 + V_1^{\mathrm{b}} = 4 - \epsilon + \sum_{k\geq 4}(k-4)V_k + \sum_{k\geq 2}(k-2)V_k^{\mathrm{b}}. \tag{8}$$

If $V_1^{\mathrm{b}} > 0$, then case 2 of the lemma takes place.

Suppose that $V_1^{\mathrm{b}} = 0$. Then it follows from (8) that $2V_2 + V_3 > 0$. If all vertices are good (and hence there is no winding vertex), we are done: either case 3 or 4 of the lemma occurs.

(iii.a) Assume that some vertices of $\mathcal{F}$ are bad, but there is no winding vertex. In each family of parallel fibres of $\mathcal{F}$ connecting a bad vertex with a boundary vertex, we select one fibre and cut the surface M along all the selected arcs (see Fig. 17). Let $d_1, \ldots, d_l$ be the resulting disks. We say that
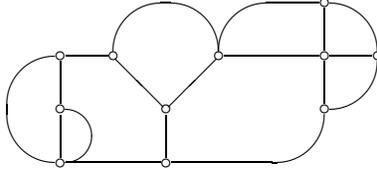
Fig. 17. Cutting the disk M

a disk d_i is *terminal* if there is exactly one bad vertex of $\mathcal{F}$ in ∂d_i. Clearly, there should be at least two terminal disks (e.g., there are six in Fig. 17). Therefore, we can find a terminal disk d_i such that ∂d_i does not contain the boundary saddle of $\mathcal{F}$. Let d_1 be that disk.

Consider the restriction $\widetilde{\mathcal{F}}$ of the foliation $\mathcal{F}$ to the disk d_1. Let $\widetilde{V}_k$ and $\widetilde{V}_k^{\mathrm{b}}$ be the number of interior and boundary k-valent vertices of $\widetilde{\mathcal{F}}$, respectively. By analogy with (8), we have

$$2\widetilde{V}_2 + \widetilde{V}_3 + \widetilde{V}_1^{\mathrm{b}} = 4 + \sum_{k \geq 4}(k-4)\widetilde{V}_k + \sum_{k \geq 2}(k-2)\widetilde{V}_k^{\mathrm{b}} \geq 4.$$

Since $V_1^{\mathrm{b}} = 0$ and the disk d_1 is terminal, we have $\widetilde{V}_1^{\mathrm{b}} \leq 3$, which implies $2\widetilde{V}_2 + \widetilde{V}_3 \geq 1$. Thus, there is at least one two- or three-valent interior vertex in d_1. By construction, this vertex is a good vertex of $\mathcal{F}$, so either case 3 or 4 takes place.

(iii.b) Now assume that there is a winding vertex, but no boundary saddle (i.e. $\epsilon = 0$). Assume also that all vertices s_i^{w} in σ^{w} are pairwise distinct and that cases 6–8 do not occur. Let V_k^{w} (respectively, V_k^{bw}) be the number of k-valent interior (respectively, boundary) vertices of $\mathcal{F}$ that lie in σ^{w}. Consider the sequence $\Delta = (\delta_1, \ldots, \delta_{q'})$ of ± 1s defined above. Let m_+ (respectively, m_-) be the number of $+1$s (respectively, -1s) in Δ, and let m_{--} be the number of subsequences $(-1, -1)$ in Δ. We have

$$(9) \qquad\qquad m_{--} \geq m_- - m_+ - 1.$$

Whenever we have a subsequence of the form $(\delta_{i-1}, \delta_i) = (-1, -1)$ in Δ, the vertex s_i^{w} cannot be

- a two-valent interior vertex, since two saddles in the star of such a vertex must be incoherently oriented (see [3]);
- a k-valent interior vertex with $2 < k < 5$ or a two-valent boundary vertex, by the assumption that cases 6–8 do not occur.

Thus the vertex s_i^{w} can only be a k-valent interior vertex with $k \geq 5$ or a k-valent boundary vertex with $k \geq 3$. Taking into account that the winding vertex itself has valence $q' = m_+ + m_- \geq 2$, we infer from (9) that

$$(10) \qquad \sum_{k \geq 5}(k-4)V_k^{\mathrm{w}} + \sum_{k \geq 3}(k-2)V_k^{\mathrm{bw}} \geq m_{--} + q' - 2 \geq 2m_- - 3.$$

It follows from Lemma 4 that

(11) $V_2^{\mathrm{w}} \le m_-.$

Combining (5), (10), (11), and recalling that $V_1^{\mathrm{b}} = 0$, we obtain

$$2(V_2 - V_2^{\mathrm{w}}) + V_3 = 4 + \sum_{k \ge 5}(k-4)V_k + \sum_{k \ge 3}(k-2)V_k^{\mathrm{b}} - 2V_2^{\mathrm{w}}$$

$$\ge 4 + \sum_{k \ge 5}(k-4)V_k^{\mathrm{w}} + \sum_{k \ge 3}(k-2)V_k^{\mathrm{bw}} - 2V_2^{\mathrm{w}} \ge 1,$$

which implies that there is a two- or three-valent interior vertex in $M \setminus \sigma^{\mathrm{w}}$. We are done unless there are bad vertices in $M \setminus \sigma^{\mathrm{w}}$. In the latter case, we apply the same cutting as in case (iii.a) and note that the presence of a bad vertex in $M \setminus \sigma^{\mathrm{w}}$ implies that one can find a terminal disk whose interior is disjoint from σ^{w}. Since the winding vertex is away from such a terminal disk, we can apply the same argument as in case (iii.a).

(iii.c) Now let both a winding vertex and a boundary saddle be present, but still assume that the vertices s_i^{w} are all distinct. Recall that, in this case, the boundary saddle lies in σ^{w} by construction. Now we have $\epsilon = 1$ and

$$2(V_2 - V_2^{\mathrm{w}}) + V_3 = 3 + \sum_{k \ge 5}(k-4)V_k + \sum_{k \ge 3}(k-2)V_k^{\mathrm{b}} - 2V_2^{\mathrm{w}} = A + B + C,$$

where

$$A = \sum_{k \ge 5}(k-4)(V_k - V_k^{\mathrm{w}}) + \sum_{k \ge 3}(k-2)(V_k^{\mathrm{b}} - V_k^{\mathrm{bw}}),$$

$$B = \sum_{k \ge 5}(k-4)V_k^{\mathrm{w}} + \sum_{k \ge 3}(k-2)V_k^{\mathrm{bw}} - 2m_- + 3,$$

$$C = 2m_- - 2V_2^{\mathrm{bw}}.$$

By the same argument as in case (iii.b), we have $B, C \ge 0$, and by construction $A \ge 0$. So, we need to show that at least one of A, B, C is non-zero.

Suppose the opposite, i.e.,

(12) $A = B = C = 0.$

If $B = 0$, then (9) must be sharp, which implies $\delta_q = -1$. The equalities $C = 0$, $\delta_q = -1$, and Lemma 4 imply that s_q^{w} is two-valent. Let s be the boundary vertex different from s^{w} adjacent to the boundary saddle. This vertex is not in σ^{w}, so its valence must be two (otherwise we would have $A > 0$). Therefore, the star of s does not contain any saddle except x_q^{w} and x_{q-1}^{w}, which implies that s_{q-1}^{w} is a boundary vertex connected to s by an arc of L (see Fig. 18). It should be at least two-valent, which implies $q \ge 3$.

Let r be the smallest integer with $2 \le r \le q-1$ such that $s_r^{\mathrm{w}} \in \partial M = L$. As we have just seen, the existence of such an r follows from (12). Let us
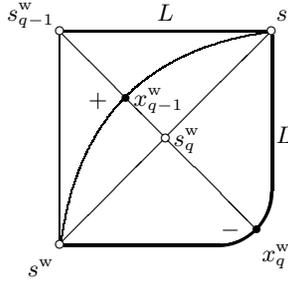
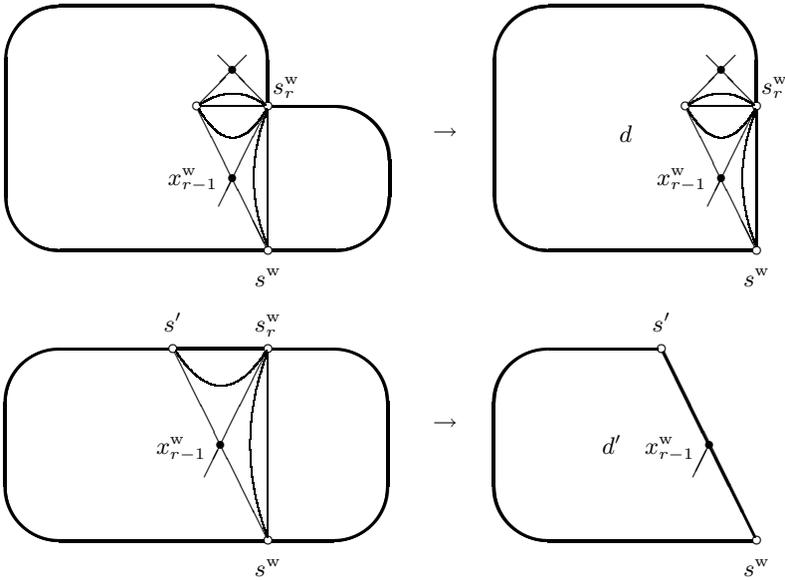Fig. 18. In case (iii.c), the equalities (12) imply $s_{q-1}^{\mathrm{w}} \in L$.

Fig. 19. The smaller disks d and d'

cut the disk M along a regular fibre connecting s^{w} and s_r^{w} into two disks. Let d be the one of the two that contains s_1^{w} (see Fig. 19).

If s^{w} becomes a univalent boundary vertex of d (i.e., $r = 2$), then its star $\sigma^{\mathrm{w}} \cap d$ in d contains no interior vertices, and $\mathcal{F}|_d$ has at most two univalent boundary vertices, so we can apply the argument of case (iii.a) to show that there is a good two- or three-valent interior vertex in d. Assume that the valence of s^{w} in d is larger than one.

If s_r^{w} is not a univalent boundary vertex of d, then we can apply the argument of case (iii.b) to d without any change.

If s_r^{w} is a univalent boundary vertex of d, let s' be the vertex adjacent to x_{r-1}^{w} but different from s^{w}, s_{r-1}^{w}, s_r^{w}. Let us cut off a "triangle" whose "hypotenuse" consists of separatrices connecting x_{r-1}^{w} to s^{w} and to s' from

the disk d. By assumption, after the cutting, s^{w} and s' become boundary vertices of valence at least two. Thus, the restriction of $\mathcal{F}$ to the smaller disk has a boundary saddle and no univalent boundary vertices. So, we can repeat our reasoning above for this disk and show that s_{r-1}^{w} is a boundary vertex, a contradiction.

(iii.d) Finally, suppose that some of the vertices s_i^{w}, $i = 1, \ldots, q$, coincide. Choose $s_i^{\mathrm{w}} = s_j^{\mathrm{w}}$ with $j - i > 0$ smallest possible. Let d be a disk enclosed by two non-parallel fibres connecting s^{w} with $s_i^{\mathrm{w}} = s_j^{\mathrm{w}}$ such that $s_k^{\mathrm{w}} \in d$ if and only if $i < k < j$. The disk d has only two boundary vertices. The vertex s_i^{w} cannot be a univalent boundary vertex of d because otherwise we would have $s_{i+1}^{\mathrm{w}} = s_{j-1}^{\mathrm{w}}$, so the argument of case (iii.b) applies to d. ∎

2.4. Simplifying the characteristic surface. This is the final part of the proof of Propositions 6–8 and Theorem 1.

LEMMA 6. *Let L be an arc-presentation satisfying the assumptions of one of Propositions 6–8, and M an admissible characteristic surface for L such that $c(M) > 0$. Then there exist an arc-presentation L' and an admissible characteristic surface M' for L' such that one of the following holds*:

1) *L' is obtained from L by finitely many exchange moves, and we have $c(M') < c(M)$;*
2) *L' is obtained from L by finitely many exchange moves and one destabilization move.*

Proof. It suffices to show how to obtain L', M' in each case listed in Lemma 5. In Cases 1, 3, 4 we just follow [3, 4, 6] and show additionally that the transformation $L \mapsto L'$ can be decomposed into elementary moves. Case 2 is specific in our situation. In Cases 5–8 our actions will be similar to those in Cases 3 and 4. In Case 5, we also consider a situation not covered by [3, 4, 6]. Two tricks play an important rôle here: changing the valence of a vertex (Cases 4 through 8) and reducing a pair of vertices one of which is two-valent (Cases 3 and 5) by using a "global exchange" from Lemma 3. These tricks are the same in nature as those that were used by Bennequin in [1] to prove an inequality relating the knot genus, the number of strands, and the crossing number of a braid representing the knot.

REMARK 8. The Reeb foliation in S^3, which Bennequin considers, is different from ours and from the one used in [3, 4, 6]. Our foliation $d\theta = 0$ can be obtained from the Reeb foliation by contracting one of the solid tori bounded by the compact leaf into a circle, which becomes our circle S_φ^1. This difference between foliations turns out to be unimportant for the arguments that are used to show that the foliation $\mathcal{F}$ can be modified in a prescribed way once we know that it contains a certain pattern.

CASE 1: $\mathcal{F}$ *has a pole.* We remove all poles and closed fibres of the foliation $\mathcal{F}$. The surface obtained has new boundary components, each consisting of a saddle and a separatrix and lying in a page $\mathcal{D}_t$. For such boundary component γ, we attach a disk bounded by γ that lies in $\mathcal{D}_t$ to the surface and then deform the result slightly to obtain a smooth surface. The saddle singularity at γ disappears (see Fig. 20). The complexity of the surface is decreased, the surface remains admissible, the link L is untouched.
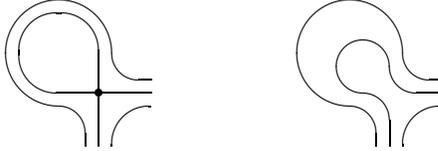

Fig. 20. Removing poles and closed fibres

In what follows we assume that there are no poles, and hence, no closed regular fibres in $\mathcal{F}$.

CASE 2: $\mathcal{F}$ *has a univalent boundary vertex.* Let s_1 be a univalent boundary vertex, s_2, s_3 the neighbouring vertices to s_1 at the knot L, α and β the two arcs of L that connect s_1 with s_2 and s_3, respectively, and x the saddle connected with s_1 by a separatrix. The saddle x may be a boundary saddle as well as an interior one.

Suppose x is a boundary saddle. We may assume without loss of generality that $x \in \beta$. The behaviour of the foliation near s_1 is shown in Fig. 21 (left). There is no topological obstruction to slide the arc α over M toward x in such a way that, at any time, α coincides with a regular fibre of $\mathcal{F}$ with endpoints s_1, s_2.

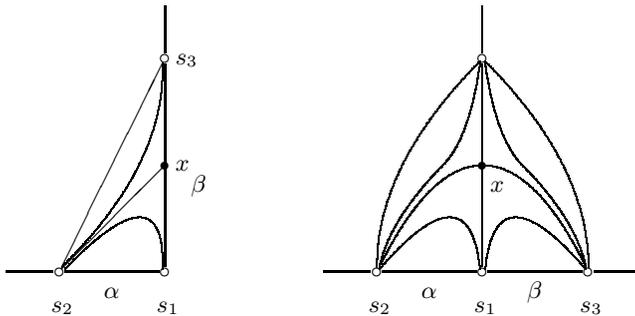Fig. 21. A univalent vertex of $\mathcal{F}$

As a result of the sliding finitely many arc exchange moves will occur. We move α until α and β become neighbouring arcs. When α and β become neighbours, we apply a generalized destabilization move.

If x is not a boundary saddle, then we do the same thing, but with both arcs α and β. The star of s_1 is shown in Fig. 21 (right). We slide α and β toward x until it becomes possible to apply a generalized destabilization move.

REMARK 9. Note that, once we have applied a destabilization move and the arc-presentation L has become simpler, we forget the surface M and search for a completely new admissible characteristic surface for the resulting arc-presentation. However, the new characteristic surface can be found by a modification of the old one, which may yield in some optimization of the simplification procedure. We do not discuss this in detail because the optimization questions are out of the scope of this paper.

CASE 3: $\mathcal{F}$ *has a good two-valent interior vertex.* Let s_1 be a good two-valent interior vertex and let x_1, x_2 be the two saddles in the star of s_1. There must also be two vertices in the star of s_1. Denote them by s_2, s_3. We may assume without loss of generality that s_2 is also an interior vertex. The star of s_1 is shown in Fig. 22. Let $t_1 = \theta(x_1)$, $t_2 = \theta(x_2)$. We may also assume that the section $\mathcal{D}_t \cap M$ contains the arc γ_{t,s_1,s_2} when $t \in (t_1, t_2)$ and the arc γ_{t,s_1,s_3} when $t \in (t_2, t_1)$, and we have $s_1 \in (s_2, s_3)$. The other cases are obtained by flipping the orientation of S_φ^1 and/or S_θ^1.
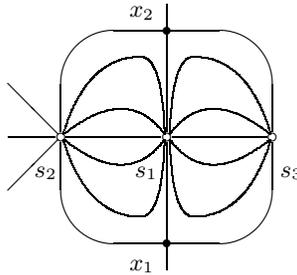

Fig. 22. A good two-valent vertex of $\mathcal{F}$

It was pointed out by Bill Menasco and Adam Sikora that it is important to know here that s_3 is not a winding vertex. That is why we consider all interior vertices in the star of a winding vertex to be bad by definition.

Pick a small $\varepsilon > 0$. No arc in $\mathcal{D}_t \cap (L \cup M)$ will be interleaved with the arc $(s_2+\varepsilon)(s_1-\varepsilon)$ if $t \in [t_1, t_2]$ and with $(s_1-\varepsilon)(s_3+\varepsilon)$ if $t \in (t_2, t_1)$. Clearly, if no arc in $\mathcal{D}_t \cap (L \cup M)$ is interleaved with an arc ss', then this remains true under a small variation of t. So, if ε is small enough, then no connected component of $\mathcal{D}_t \cap (L \cup M)$ is interleaved with the arc $(s_2+\varepsilon)(s_1-\varepsilon)$ if $t \in [t_1-\varepsilon, t_2+\varepsilon]$ and with $(s_1 - \varepsilon)(s_3 + \varepsilon)$ if $t \in [t_2 + \varepsilon, t_1 - \varepsilon]$. Thus, the arc-presentation L and the surface M satisfy the assumptions of Lemma 3, where we should

replace s_1, s_2, s_3, t_1, t_2 by $s_2 + \varepsilon, s_1 - \varepsilon, s_3 + \varepsilon, t_1 - \varepsilon, t_2 + \varepsilon$, respectively, and we can exchange the intervals $(s_2 + \varepsilon, s_1 - \varepsilon)$ and $(s_1 - \varepsilon, s_3 + \varepsilon)$.

The new arc-presentation L' is obtained from L by a generalized exchange move, which, according to Proposition 5, is the composition of ordinary exchange moves. As a result of the exchange, the vertex s_1 is moved to $s_1' = s_2 + 2\varepsilon$. There is no other vertex of L or $\mathcal{F}$ in (s_2, s_1'). The foliation of the surface is not changed, therefore, there still exists a regular fibre α connecting the vertices s_2 and s_1'. We can now isotope the surface so that the resulting surface M' is admissible and we have $|M' \cap S_\varphi^1| = |M \cap S_\varphi^1| - 2$. In order to see that there is no obstruction, we can do the following. First, we replace the interval $(s_2 - \varepsilon, s_1' + \varepsilon)$ of the binding circle by an arc parallel to α (see Fig. 23). Then we apply an isotopy that will restore the binding circle.
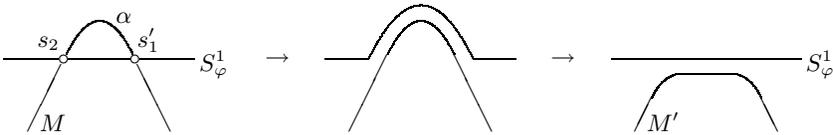

Fig. 23. Simplifying the surface M

Under this isotopy, the boundary of the surface stays unchanged. We may also assume that the surface M' is in general position (thus, it is still admissible) and there are no poles of the foliation $\mathcal{F}'$ on M' defined by $d\theta = 0$. Indeed, if there is one, we can apply the procedure described in Case 1 and get rid of all poles and closed fibres of the foliation.

In the case $P = 0$, from (2) and (3) we have

$$c(M) = \sum_k V_k + \frac{1}{2} \sum_k V_k^{\mathrm{b}} + \frac{\epsilon}{2} - \chi.$$

Since the number of boundary vertices of $\mathcal{F}'$ is the same as that of $\mathcal{F}$, and the number of interior vertices of $\mathcal{F}'$ is less than that of $\mathcal{F}$, we have $c(M') < c(M)$.

REMARK 10. The operation of removing the vertices s_2, s_1' has the following effect on the combinatorial description of M: in all pages $\mathcal{D}_t$ with $t \in (t_1, t_2)$ the arc $s_2 s_1'$ disappears, in all regular pages $\mathcal{D}_t$ with $t \in (t_2, t_1)$ a pair of arcs one of which is $s_1' s_3'$ and the other has the form $s_2 s$ is replaced by the arc $s_3' s$, the events at $t = t_1, t_2$ disappear, the events at $t \in (t_2, t_1)$ are adjusted by replacing s_2 with s_3'.

CASE 4: *$\mathcal{F}$ has a good three-valent interior vertex.* Let s_1 be a good three-valent vertex and let s_2, s_3, s_4 be the other vertices of the star of s_1 enumerated so that $s_2 \in (s_1, s_3)$, $s_4 \in (s_3, s_1)$. Let x_1 (respectively, x_2) be the saddle (from the star of s_1) connected by separatrices to s_2, s_3 (re-

spectively, s_3, s_4). The star of s_1 cannot contain a boundary saddle, since otherwise there would be two boundary vertices in the star, which would contradict the goodness of s_1. Thus, neither of the points x_1, x_2 is a boundary saddle. Let s_5 and s_6 be the vertices other than s_1, s_2, s_3, s_4 connected by separatrices to x_1 and x_2, respectively (see Fig. 24).
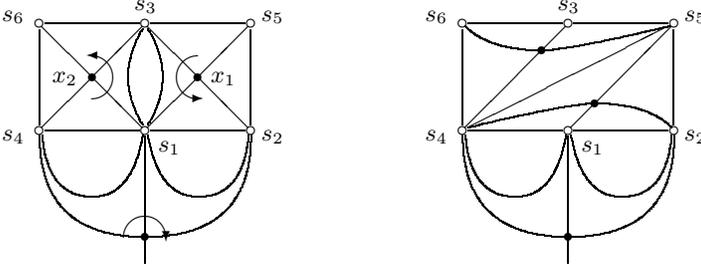

Fig. 24. Turning a three-valent vertex of $\mathcal{F}$ into a two-valent one

Let $t_1 = \theta(x_1)$ and $t_2 = \theta(x_2)$. We may assume without loss of generality that θ increases when we go from x_1 to x_2 transversely to the fibres connecting s_1 and s_3. In other words, we have $\gamma_{t,s_1,s_3} \subset M$ if $t \in (t_1, t_2)$. Let us now see what we know about the flow of the sections $\mathcal{D}_t \cap (L, M)$ when t runs over $[t_1, t_2]$. The following events occur:

- $t = t_1$: the arcs $s_1 s_2$ and $s_3 s_5$ are replaced by $s_1 s_3$ and $s_2 s_5$;
- $t \in (t_1, t_2)$: some unknown events happen, but none of them involves the arc $s_1 s_3$;
- $t = t_2$: the arcs $s_1 s_3$ and $s_4 s_6$ are replaced by $s_1 s_4$ and $s_3 s_6$.

Together with our assumptions, this implies that the following pairs of arcs are non-interleaving: $(s_1 s_2, s_3 s_5)$, $(s_1 s_3, s_2 s_5)$, $(s_1 s_3, s_4 s_6)$, $(s_1 s_4, s_3 s_6)$, which yields the following circular order of the vertices in the binding circle: $s_1, s_2, s_5, s_3, s_6, s_4$.

Since the unknown events that happen in (t_1, t_2) do not involve the arc $s_1 s_3$, all the arcs participating in them must have both endpoints either in (s_1, s_3) or in (s_3, s_1). In the first case, there is no topological obstruction to moving these events "to the future" so that they happen after t_2. In the latter case, we can move them "to the past", i.e., before t_1.

As a result, finitely many arc exchange moves are applied to L, and an isotopy preserving the foliation $\mathcal{F}$ is applied to M. The main achievement is this:

1) there is no event between t_1 and t_2;
2) all three arcs $s_1 s_2$, $s_3 s_5$ and $s_4 s_6$ are present just before time t_1;
3) all three arcs $s_2 s_5$, $s_3 s_6$ $s_1 s_4$ are present right after time t_2.

REMARK 11. There is a minor mistake in [6], in the proof of Lemma 3: it is claimed (in different terms) that the only obstruction to an isotopy making the events corresponding to x_1, x_2 successive are arcs of L. This is not necessarily so, the events in the time interval (t_1, t_2) can be arbitrary, including those corresponding to saddles. In particular, before the exchanges, the arc $s_2 s_5$, which must be present in $\mathcal{D}_{t_1+\varepsilon} \cap M$ for a small ε may be already absent in $\mathcal{D}_{t_2} \cap M$, being destroyed at some $t \in (t_1, t_2)$.

Now we replace the couple of events

$$(s_1 s_2, s_3 s_5) \mapsto (s_1 s_3, s_2 s_5) \quad \text{followed by} \quad (s_1 s_3, s_4 s_6) \mapsto (s_1 s_4, s_3 s_6)$$

by the following couple of events:

$$(s_3 s_5, s_4 s_6) \mapsto (s_3 s_6, s_4 s_5) \quad \text{followed by} \quad (s_1 s_2, s_4 s_5) \mapsto (s_1 s_4, s_2 s_5).$$

This results in an isotopy of the surface M and changing the foliation as shown in Fig. 24 (right). For more explanations and figures illustrating the change of the surface, see [3].

After the change of the foliation, the vertex s_1 becomes two-valent. It remains good because the star of s_1 becomes smaller. So, we can now proceed as in the case of a good two-valent vertex (Case 3).

CASE 5: *$\mathcal{F}$ has two pairs two-valent vertices, in each the vertices are connected by a fibre intersecting L.* This case may occur only when L is a composite link and M is a factorizing sphere. This case was not considered in full generality in [6], so here we also fill a gap in the proof given in [6] of the additivity of arc-index.

Let s_1, s_2 be a pair of two-valent vertices connected by a fibre α of $\mathcal{F}$ that intersects L. Since there are exactly two such pairs of vertices and exactly two points in $L \cap M$, the union U of the stars of s_1 and s_2 is pierced only once. The structure of $\mathcal{F}$ in U is shown in Fig. 25. The "$*$" in the figure stands for the point $L \cap U$. Let β be the arc of L that pierces U and x_1, x_2 be the saddles in U. Let $t_1 = \theta(x_1)$, $t_2 = \theta(x_2)$, $t_3 = \theta(*)$.

Let $\partial \beta = \{s_5, s_6\}$. Since the arcs $s_1 s_2$ and $s_5 s_6$ interleave, we may assume $s_5 \in (s_1, s_2)$ and $s_6 \in (s_2, s_1)$. Denote by β' the arc of L attached to s_5 but different from β. Let s_3 be the vertex from the star of s_1 other than s_2, and
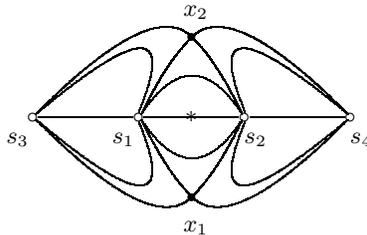

Fig. 25. A pierced region

s_4 the vertex from the star of s_2 other than s_1. We can always achieve the
following by adjusting notation: $s_1 \in (s_3, s_2)$, $s_2 \in (s_1, s_4)$, $t_3 \in (t_1, t_2)$.

Suppose that the interval (s_1, s_2) of the binding circle does not intersect
the surface M. Let A_1 (respectively, A_2) be the set of arcs in L that lie in
the sector $\theta \in (t_1, t_3)$ (respectively, $\theta \in (t_3, t_2)$) and have both endpoints
in (s_1, s_2). By using arc exchange moves, we can shift all the arcs from A_1
"to the past" $(t_1 - \varepsilon, t_1)$ and the arcs from A_2 "to the future" $(t_2, t_2 + \varepsilon)$,
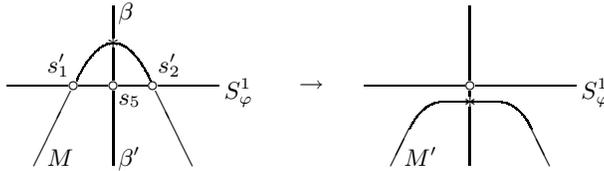keeping the relative order of arcs in each family A_i.

Fig. 26. Simplifying the surface M near a piercing point

Now, for a sufficiently small ε, we can exchange the intervals of the
binding circle: $(s_1 + \varepsilon, s_5 - \varepsilon)$ with $(s_3 - \varepsilon, s_1 + \varepsilon)$, and $(s_5 + \varepsilon, s_2 - \varepsilon)$ with
$(s_2 - \varepsilon, s_4 + \varepsilon)$, since the assumptions of Lemma 3 are satisfied. As a result of
the exchanges, the vertices s_1, s_2 are moved to $s_1' = s_5 - 2\varepsilon$ and $s_2' = s_5 + 2\varepsilon$,
respectively, becoming neighbouring to s_5. We can now reduce the number
of vertices of $\mathcal{F}$ in the same way as we did in Case 3. The intersection point
$*$ slides over L from the arc β to β' (see Fig. 26). The rest of the argument
is the same as in Case 3.

It remains to consider the situation in which M intersects the inter-
val $(s_1, s_2) \subset S_\varphi^1$. Actually, the procedure that has just been considered
may work in this case, too. One should include in A_1, respectively A_2, all
the events in (t_1, t_3), respectively (t_3, t_2), involving arcs with endpoints in
(s_1, s_2). But there is one really bad case, which indeed *can* take place: the
arc β can intersect the surface M in two points. If the second intersection
point lies in the segment of β between s_5 and $*$, we cannot exchange the
intervals. We now show that, in the case when M intersects $(s_1, s_2) \subset S_\varphi^1$,
the surface M can be simplified by using a different method.

As we have seen in the proof of Lemma 5, if none of Cases 1–4 occurs,
then four of the vertices of the foliation $\mathcal{F}$ are two-valent and all the others
are four-valent. It is not difficult to see that, in this case, the surface M can
be cut along regular non-separatrix fibres of $\mathcal{F}$ into two disks and a number
of annuli in which the foliation looks as shown in Fig. 27. In the figure, i runs
from 1 through k, where $2k$ is the total number of vertices of $\mathcal{F}$. Certainly,
we have $k \geq 3$, since the vertices s_3, s_4 do not lie in (s_1, s_2) by construction.
Notice that we have chosen the notation for s_1, s_2, s_3, s_4 to be as before; $*_1$
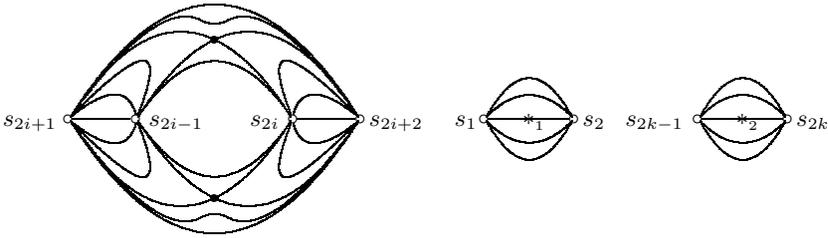and $*_2$ stand for the points of $L \cap M$.

Fig. 27. Patterns of $\mathcal{F}$ in Case 5

For a moment, we step back to Case 4 and Figure 24. What assumption did we use in the proof that the foliation can be changed as shown in Fig. 24 (right)? Examining the proof, one sees that we only needed that the region filled by arcs connecting s_1 and s_3 was not pierced by L and the arcs s_2s_5 and s_4s_6 were separated from each other in the binding circle S^1_φ by the arc s_1s_3, which is equivalent to saying that the saddles x_1 and x_2 in Figure 24 are coherently oriented. In the case of a three-valent vertex, we achieved this by appropriately numbering the vertices s_2, s_3, and s_4. But, clearly, we can apply the same procedure in the case of a vertex s_1 such that the three vertices s_2, s_3, s_4 in the star of s_1 viewed from s_1 in this order satisfy the condition $s_3 \in (s_2, s_4)$. The effect on the foliation is that the valence of s_1 is decreased by 1 and the complexity of M is unchanged.
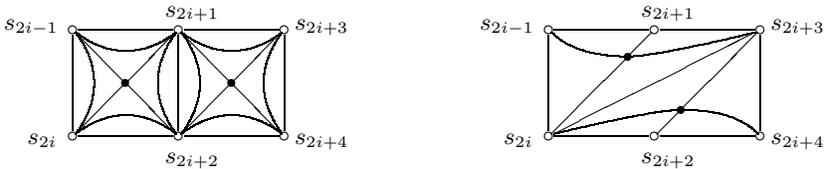

Fig. 28. Turning a four-valent vertex into a two-valent one

Now we return to the case of the foliation shown in Fig. 27. For any $i = 1, \ldots, k - 2$ the foliation $\mathcal{F}$ has two copies of the pattern shown in Fig. 28 (left). If the arcs $s_{2i-1}s_{2i}$ and $s_{2i+3}s_{2i+4}$ are separated from each other in S^1_φ by the arc $s_{2i+1}s_{2i+2}$, then we can change the foliation as shown in Fig. 28 (right) by applying the same procedure as in Case 4 to both copies. The vertices s_{2i+1} and s_{2i+2} will become two-valent, and we can simplify the surface M as in Case 3.

So, the only possibility left is that the arcs $s_{2i-1}s_{2i}$ and $s_{2i+3}s_{2i+4}$ are not separated by $s_{2i+1}s_{2i+2}$ for all $i = 1, \ldots, k - 2$. We now show that, in this case, the interval (s_1, s_2) is disjoint from M, which contradicts the assumption.

In order to help the imagination, we explain informally what the surface M looks like in the case under consideration. Let C be a simple non-closed

curve in S^3 with endpoints in S^1_φ consisting of $k-1$ arcs, each lying in a separate page $\mathcal{D}_t$. Let S be the boundary of a small neighbourhood of C. If the neighbourhood has been chosen appropriately, then the foliation $d\theta = 0$ on S looks exactly as that on M. Actually, the surface M can be isotoped to such an S without changes in the foliation and combinatorial structure of L for an appropriately chosen curve C.

Let I_j be the following interval of the binding circle:

$$I_j = \begin{cases} (s_{2j-1}, s_{2j}) & \text{if } j \text{ is odd,} \\ (s_{2j}, s_{2j-1}) & \text{if } j \text{ is even.} \end{cases}$$

We want to show that $I_i \cap I_j = \emptyset$ if $i \neq j$, which will imply $(s_1, s_2) \cap M = I_1 \cap M = \emptyset$. Recall that the vertices s_1, s_2, s_3, s_4 have been numbered so that $I_1 \cap I_2 = \emptyset$.

Suppose that we have $I_i \cap I_{i+1} = \emptyset$, $i \leq k-2$. There is a saddle of $\mathcal{F}$ that is connected by separatrices to the vertices $s_{2i+1}, s_{2i+2}, s_{2i+4}, s_{2i+3}$ in this circular order. These vertices must appear in S^1_φ in the same or opposite circular order. Since the arcs $s_{2i-1} s_{2i}$ and $s_{2i+3} s_{2i+4}$ are not separated by $s_{2i+1} s_{2i+2}$, we see that the vertices $s_{2i+1}, s_{2i+2}, s_{2i}, s_{2i-1}$ go in S^1_φ in the same circular order as the vertices $s_{2i+1}, s_{2i+2}, s_{2i+4}, s_{2i+3}$ do. Therefore, the assumption $I_i \cap I_{i+1} = \emptyset$ implies $I_{i+1} \cap I_{i+2} = \emptyset$. Since $I_1 \cap I_2 = \emptyset$, we have $I_i \cap I_{i+1} = \emptyset$ for all $i = 1, \ldots, k-1$.

Now, we show by induction that, for $i < j$, we have either $I_i \cap I_j = \emptyset$ or $I_i \subset I_j$. This is true for $i = 1$, $j = 2$. Suppose this is true for all $i < j \leq m$. We have already shown that $I_m \cap I_{m+1} = \emptyset$. Let $i \in \{1, \ldots, m-1\}$. In any regular page $\mathcal{D}_t$ one of the following is present:

1) the arc $s_{2i-1} s_{2i}$, which separates I_i from the rest of the binding circle;
2) the arcs $s_{2i-1} s_{2i+1}$ and $s_{2i} s_{2i+2}$, which separate $I_i \cup I_{i+1}$ from the rest of the circle;
3) the arcs $s_{2i-3} s_{2i-1}$ and $s_{2i-2} s_{2i}$, which separate $I_i \cup I_{i-1}$ from the rest of the circle.

In order to see this, refer to the stars of s_{2i-1} and s_{2i}. One of regular pages contains the arcs $s_{2m-1} s_{2m+1}$ and $s_{2m} s_{2m+2}$. By the induction hypothesis, the vertices s_{2m-1} and s_{2m} are outside $I_{i-1} \cup I_i \cup I_{i+1}$. Therefore, the vertices s_{2m+1} and s_{2m+2} are outside I_i. This implies $I_i \cap I_{m+1} = \emptyset$ or $I_i \subset I_{m+1}$.

By symmetry, we conclude that, for any $i > j$, we also have either $I_i \cap I_j = \emptyset$ or $I_i \subset I_j$. Therefore, $I_i \cap I_j = \emptyset$ for all $i \neq j$. This concludes the proof of the lemma in the case of a composite link.

CASE 6: $\mathcal{F}$ has a three-valent interior vertex s^{w}_i in the star of a winding vertex s^{w}. Assume for the moment that there is no boundary saddle. We apply the same trick as in Case 4, changing the valence of the vertex s^{w}_i to

two. The result depends on whether $\delta_{i-1} = \delta_i$ or not. Refer again to Fig. 24, where s_1 is the vertex s_i^{w}.

If $\delta_{i-1} = \delta_i$, then according to Fig. 24 the winding vertex must be s_3. As a result of the change of the foliation, the vertex s_1, which becomes two-valent, escapes the star of s^{w}. If it becomes good, we proceed as in Case 3. If it is still bad, we can apply the same argument as in case (iii.b) of the proof of Lemma 5 to show that there is a good two- or three-valent interior vertex or a one-valent boundary vertex somewhere else, and proceed as in Case 2, 3, or 4 above.

If $\delta_{i-1} \neq \delta_i$, then according to Fig. 24 the winding vertex must be s_2 or s_4. We may assume that it is s_4. As a result of the change of the foliation the valence of s^{w} increases, whereas the complexity of the disk M does not change. So, after finitely many such operations either all three-valent interior vertices will be removed from σ^{w} or we get a good interior vertex of valence ≤ 3 or a univalent boundary vertex.

A slight complication occurs in the case of boundary saddle present, when the three-valent vertex under consideration is s_q^{w}, and we have $\delta_{q-1} = \delta_q$. Similarly to Case 4, one can show that now the change of the foliation indicated in Fig. 29 is possible. The vertex s_q^{w} will escape the star of s^{w} and become two-valent, so we can proceed as above.

Fig. 29. The case when s_q^{w} is three-valent and x_q^{w} is a boundary saddle

CASE 7: $\mathcal{F}$ *contains a two-valent vertex* s_i^{w} *in the star of a winding vertex, and* $\delta_{i-1} = \delta_i$. By the same trick as in Case 4 we can change the valence of the vertex s_i^{w} to one, and then simplify the knot as in Case 2.

CASE 8: $\mathcal{F}$ *contains a four-valent interior vertex* s_i^{w} *in the star of a winding vertex, and* $\delta_{i-1} = \delta_i$. By changing the valence of the vertex s_i^{w} to three, we also remove it from the star of the winding vertex. So, either it becomes good and we can proceed as in Case 4, or it is still bad and then there must be a good two- or three-valent interior vertex or a univalent boundary vertex somewhere else. ∎

In order to complete the proof of Propositions 6–8, we need only consider the case $c(M) = 0$. If L is an arc-presentation of the unknot, then (2) and (3) imply $c(L) = V_2^{\mathrm{b}} = 2$.

If L is a split or composite link, we may assume without loss of generality that $L \cap S_\varphi^1 = \{0, \pi\}$. Then the line $x = \pi$ will be splitting or factorizing, respectively, for the rectangular diagram of L.

A couple of remarks are in order.

REMARK 12. The assertions of Propositions 7 and 8 can be strengthened a little by skipping destabilization moves in the formulation. Indeed, it is trivial to show that, for any sequence $L_0 \mapsto L_1 \mapsto \cdots \mapsto L_N$ of exchange and destabilization moves, there exists a sequence $L_0 \mapsto L_1' \mapsto \cdots \mapsto L_{N'-1}' \mapsto L_{N'}' = L_N$ in which the first k moves are exchanges and the last $N' - k$ ones destabilizations. If L_N is split or composite, then L_k' is also split or composite, respectively.

REMARK 13. One can easily show that, for any two sequences

$$D_0^{(1)} \mapsto D_1^{(1)} \mapsto \cdots \mapsto D_{N_1}^{(1)} \quad \text{and} \quad D_0^{(2)} \mapsto D_1^{(2)} \mapsto \cdots \mapsto D_{N_2}^{(2)}$$

of cyclic permutation, exchange, and destabilization moves of rectangular diagrams, there exists a sequence

$$D_0^{(1)} \# D_0^{(2)} \mapsto D_1 \mapsto \cdots \mapsto D_N = D_{N_1}^{(1)} \# D_{N_2}^{(2)}$$

of cyclic permutation, exchange, and destabilization moves. A similar statement is true for the distant union operation. Thus, we can obtain a *complete* decomposition of an arc-presentation by using exchange and destabilization moves.

3. APPLICATIONS

3.1. An algorithm for recognizing the unknot, recognizing split links, and link factorization. Propositions 6–8 allow us to construct a simple algorithm for recognizing the unknot, decomposing a given link into the distant union of non-split links, and factorizing a non-split link into prime links. The algorithm works as follows.

We search for all arc-presentations obtained by finite sequences of exchange moves from a given one. Since, for any n, the number of combinatorial classes of arc-presentations of complexity n is finite, the process will terminate in finitely many steps. Let $L_1, \ldots, L_N$ be the arc-presentations found. If some L_i admits a destabilization move $L_i \mapsto L'$, we replace L by L' and proceed as before. In finitely many steps, we get a list of arc-presentations none of which can be simplified any more. Denote them again by $L_1, \ldots, L_N$ and their rectangular diagrams by $D_1, \ldots, D_N$. For each $i = 1, \ldots, N$ we find a maximal decomposition

$$D_i = (D_{i1} \# \cdots \# D_{ip}) \sqcup \cdots \sqcup (D_{iq} \# \cdots \# D_{ir}) \sqcup D_{i,r+1} \sqcup \cdots \sqcup D_{i,N_i},$$

where all diagrams $D_{i1}, \ldots, D_{ir}$ are non-trivial. We choose i so as to have N_i maximal possible. Then all diagrams $D_{i1}, \ldots, D_{ir}$ will represent non-trivial links, and all $D_{i,r+1}, \ldots, D_{iN_i}$ prime non-split links or trivial knots.

REMARK 14. It is natural to ask how fast the algorithm just described is. Unfortunately, the answer is not optimistic in the sense that the algorithm is too hard to be implemented in practice: we cannot provide an estimation for the running time much better than n^{2n}, where n is the complexity of the given diagram. However, we think that the algorithm can be improved considerably. We describe this point very briefly.

We say that an arc-presentation L is *simplifiable* if there exists a sequence of exchange moves $L \mapsto L_1 \mapsto \cdots \mapsto L_k$ such that a destabilization move can be applied to L_k. So, the problem is actually this: how to detect that a given arc-presentation is simplifiable and find the corresponding sequence of exchanges as fast as possible? Also, we need a fast method to decide whether a given non-simplifiable arc-presentation can be transformed into a split or composite one by using exchange moves. There is a hope that a solution much better than an exhaustive search might exist. So far, this problem has not been investigated.

3.2. An upper bound on the crossing number needed for untangling. We shall use the notation $c_\times(D)$ for the crossing number of a planar diagram D. We shall also denote by $c_{\mathrm{arc}}(D)$ the complexity of the arc-presentation corresponding to a rectangular diagram D. We introduce two different notations, because a rectangular diagram can be viewed in both ways, as an arc-presentation and as an ordinary planar diagram.

THEOREM 2. *Let D_0 be an ordinary planar diagram of either a trivial knot, a split link, or a composite link. Then there exists a sequence of Reidemeister moves $D_0 \mapsto D_1 \mapsto \cdots \mapsto D_N$ such that the diagram D_N is trivial, split, or composite, respectively, and we have $c_\times(D_i) \leq 2(c_\times(D_0)+1)^2$ for all i.*

Proof. It is shown in [7] that there is an algorithm for converting a planar diagram D_0 to an arc-presentation that has no more than $c_\times(D_0) + 2$ arcs provided that D_0 satisfies certain restrictions. By using the same method, one can show that the estimation $2c_\times(D_0) + 2$ works for *any* connected diagram D_0. Let $D^{(1)}$ be the corresponding rectangular diagram. Following the lines of the conversion procedure, one sees that there is a sequence of Reidemeister moves $D_0 \mapsto D_1 \mapsto \cdots \mapsto D_{i_1} = D^{(1)}$ such that $c_\times(D_i) \leq c_\times(D^{(1)})$ for all $i = 1, \ldots, i_1$.

Notice that a cyclic permutation of horizontal (vertical) edges of a rectangular diagram can be decomposed into one stabilization move, a few exchange moves, and one destabilization move (see Fig. 30). Therefore, we can
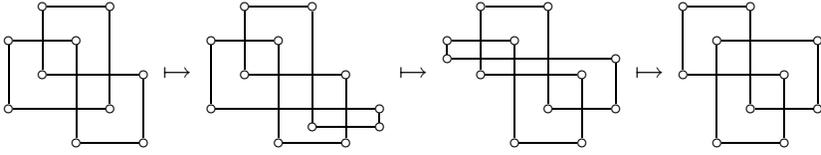
Fig. 30. Cyclic permutation via stabilization, exchanges, and destabilization

find a sequence of (de)stabilization and exchange moves $D^{(1)} \mapsto \cdots \mapsto D^{(m)}$ such that the final diagram $D^{(m)}$ is trivial, split, or composite, respectively, and, for all $j = 1, \ldots, m$, we have $c_{\mathrm{arc}}(D^{(j)}) \leq 2c_{\times}(D_0) + 3$.

It is easy to show that, for any rectangular diagram D, we have

$$c_{\times}(D) \leq (c_{\mathrm{arc}}(D) - 1)^2/2.$$

It remains to notice that, for any exchange or (de)stabilization move $D^{(j)} \mapsto D^{(j+1)}$, there exists a finite sequence of Reidemeister moves

$$D^{(j)} \mapsto D_{i_j+1} \mapsto D_{i_j+2} \mapsto \cdots \mapsto D_{i_{j+1}} = D^{(j+1)}$$

such that $c_{\times}(D_{i_j+l}) \leq \max(c_{\times}(D^{(j)}), c_{\times}(D^{(j+1)}))$ for $1 \leq l \leq i_{j+1} - i_j$. ∎

3.3. Results of Birman and Menasco on closed braids. The results of [3] and [4] on closed braid representatives of links can be deduced from the main result of this paper, which is not a surprise because we used a modification of a technique that was originally worked out by Birman and Menasco for studying closed braids.

First of all, we briefly recall what the elementary moves for braids are:

1) *conjugation*
$$b \mapsto cbc^{-1}, \quad b, c \in B_n;$$

2) *stabilization*
$$b \mapsto b\sigma_n^{\pm 1} \in B_{n+1}, \quad b \in B_n \subset B_{n+1};$$

3) *destabilization*
$$b\sigma_n^{\pm 1} \mapsto b, \quad b \in B_n \subset B_{n+1};$$

4) *exchange*
$$b_1\sigma_n^{\pm 1}b_2\sigma_n^{\mp 1} \mapsto b_1\sigma_n^{\mp 1}b_2\sigma_n^{\pm 1}, \quad b_1, b_2 \in B_n \subset B_{n+1}.$$

Moves 1)–3) are Markov's moves; the exchange moves were introduced by J. Birman and W. Menasco. In [3], [4] a more general transform is called an exchange move. However, one can easily show that the more general, "multistrand", exchange move can be decomposed into a finite number of "elementary" exchange moves and braid isotopies.

An exchange move can be decomposed into a finite sequence of Markov's moves, but not always within the class of braids with the same or smaller number of strands. The main result of [3] and [4] is the following.

THEOREM 3 (Birman and Menasco). *If b_0 is a braid whose closure is a trivial knot, then there exists a finite sequence $b_0 \mapsto b_1 \mapsto \cdots \mapsto b_N$ of conjugations, destabilizations, and exchanges such that b_N is a trivial braid on one strand.*

If b_0 is a braid whose closure is either a split link or a composite link, there exists a finite sequence $b_0 \mapsto b_1 \mapsto \cdots \mapsto b_N$ of conjugations and exchanges such that b_N is a split braid or a composite braid, respectively.

For the definition of split and composite braids, see [3].

Proof. We give a sketch only, inviting the reader to fill in the details.

The construction of 1.4 applied to a split or a composite rectangular diagram D gives a braid b_D which is conjugate to a split or composite braid, respectively.

By a straightforward check one establishes the following:

- if $D \mapsto D'$ is an exchange of vertical edges, then $b_D = b_{D'}$;
- if $D \mapsto D'$ is a cyclic permutation of edges, then either $b_D \mapsto b_{D'}$ is a conjugation or we have $b_D = b_{D'}$;
- if $D \mapsto D'$ is an exchange of the upper two horizontal edges, then $b_D \mapsto b_{D'}$ is either a conjugation or an exchange, depending on the orientation of the edges and their relative position (it is assumed that we enumerate the braid strands from bottom to top);
- if $D \mapsto D'$ is a destabilization move involving the upper two horizontal edges of D, then either $b_D \mapsto b_{D'}$ is a destabilization move or we have $b_D = b_{D'}$, depending on the orientation of the edges.

Let b be a braid whose closure is a trivial knot. According to Proposition 2 there exists a rectangular diagram D such that $b = b_D$. From Proposition 6 we know that there exists a sequence $D \mapsto D_1 \mapsto \cdots \mapsto D_N$ of exchanges, cyclic permutations, and destabilizations which takes the diagram D to a trivial diagram D_N.

Each exchange move $D_i \mapsto D_{i+1}$ can be decomposed into a number of cyclic permutations of horizontal edges and an exchange move that involves the upper two edges. Similarly, each destabilization $D_i \mapsto D_{i+1}$ can be decomposed into a number of cyclic permutations and a destabilization involving the upper two horizontal edges. Therefore, the braid b_{D_N}, which is the unit of B_1, can be obtained from the initial braid b by finitely many exchanges, conjugations, and destabilizations.

For split and composite braids, the proof is similar. One should use Propositions 7 and 8, respectively, and also Remark 12 (to show that destabilizations are not needed in this case). ∎

REMARK 15. Notice that, as a by-product, we obtained a new proof of Markov's theorem, which is now a corollary to Proposition 4.

3.4. Two tests for knottedness. How does one usually conclude that a given knot is truly knotted? The standard way to prove that a given knot is non-trivial is to compute some knot invariant like the Alexander polynomial or Jones polynomial. Such computations seem to be hard for large knots. The algorithms known up to now for computing polynomial invariants take at least exponential time in the complexity of a link diagram.

However, there is (at least) one well known exception from the above mentioned situation: if an alternating planar diagram of a knot has no separating vertex, then it presents a non-trivial knot or a non-split link (see [15]). To check that a given diagram is alternating and has no separating vertex is very easy.

Here, we provide two sufficient conditions for knottedness in terms of arc-presentations (rectangular diagrams). Verification of the conditions requires not more than quadratic time in the complexity of the diagram provided that the diagram is encoded in a reasonable way.

From Propositions 6–8 one concludes the following

COROLLARY 1. *If no exchange and no destabilization move can be applied to a non-trivial arc-presentation L, then L is a non-trivial non-split prime link.*

Proof. We need only comment on why L is prime, i.e., non-composite. The point is that, if L is composite as a link, then Proposition 8 implies that L is composite as an arc-presentation. This contradicts the assumption, since a composite arc-presentation always admits exchange moves. ∎

The assumption of Corollary 1 means that L does not have "trivial" arcs, i.e., arcs connecting two neighbouring vertices, and any two neighbouring arcs (vertices) of L interleave. In terms of the corresponding rectangular diagram, this means that any two neighbouring vertical (horizontal) edges interleave provided that we regard the rightmost and leftmost (top and bottom) edges also as neighbouring. Diagrams of this kind will be said to be *rigid*.

There are numerous examples of rigid diagrams. Any torus knot or link can be presented by a rigid diagram. A rigid diagram of the Whitehead link is displayed on the left of Fig. 31. Notice that, in contrast to the case of alternating planar diagrams, rigid and non-rigid rectangular diagrams of the same complexity can present equivalent prime links. For example, the diagram of the Whitehead link on the right of Fig. 31 is not rigid, though it has the same complexity as the one on the left.

In Fig. 32 we show a rigid diagram of a two-component link with zero multivariable Alexander polynomial and zero linking number. (The latter two invariants are often used to establish non-splitness. The example was constructed in response to a question by Józef Przytycki, who asked the
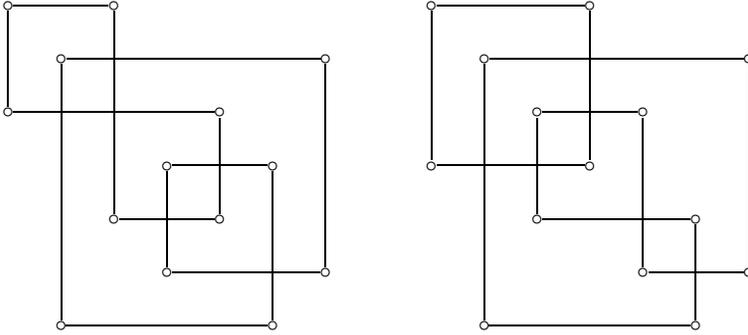
Fig. 31. Two rectangular diagrams of the Whitehead link: a rigid diagram (left) and a non-rigid diagram (right)

author whether rigid diagrams always have non-zero Alexander polynomial. So, the answer is 'no'. Recall that the non-triviality of alternating links was proved in [15] by showing that they have non-trivial Jones polynomial.)
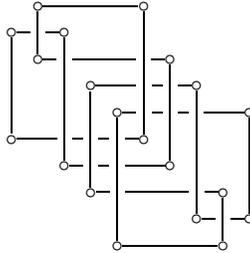

Fig. 32. A rigid diagram of a link with zero Alexander polynomial

The second test for knottedness uses an analogue of the writhe. Let us denote by f_ε^+ and f_ε^- the self-homeomorphisms of S^3 defined by

$$f_\varepsilon^\pm(\varphi, \tau, \theta) = (\varphi - \varepsilon, \tau, \theta \pm \varepsilon).$$

Let L be an arc-presentation of an *oriented* link. Pick an $\varepsilon > 0$ less than the φ-distance between any two vertices of L and the θ-distance between any two arcs of L. Then the link $f_\varepsilon^\pm(L)$ will be disjoint from L.

We define the *upper* (respectively, *lower*) *writhe* of L to be the linking number of L with $f_\varepsilon^+(L)$ (respectively, $f_\varepsilon^-(L)$). We denote the upper and lower writhes of L by $w_+(L)$ and $w_-(L)$, respectively. By the *writhe* of L we shall mean the couple $(w_-(L), w_+(L))$.

The numbers $w_\pm(L)$ can be easily computed from the knowledge of a rectangular diagram D of L. We divide the set of crossings and the set of turns of the diagram D into two types, which we call *positive* and *negative*, according to Figures 33 and 34.

Fig. 33. Positive crossings (left) and negative crossings (right)

Fig. 34. Positive turns (left) and negative turns (right)

We define the *writhe* $w(D)$ of a rectangular link diagram D in the standard way: $w(D)$ is the number of positive crossings less the number of negative crossings in D. Denote by $n_+(D)$ and $n_-(D)$, respectively, the numbers of positive and negative turns in D. The following three facts are established by a straightforward check.

PROPOSITION 10. *Let L be an arc-presentation of an oriented link, and D the corresponding rectangular diagram. Then*

$$(13) \qquad w_+(L) = w(D) + \frac{n_+(D)}{2}, \qquad w_-(L) = w(D) - \frac{n_-(D)}{2},$$

which implies, in particular, that

$$(14) \qquad\qquad w_+(L) - w_-(L) = c(L).$$

PROPOSITION 11. *The writhe of an arc-presentation does not change under exchange moves.*

Equivalently, the right-hand side of any of the formulae (13) is not changed under cyclic permutation of vertical or horizontal edges and interchange of non-interleaved neighbouring vertical or horizontal edges of D.

PROPOSITION 12. *Let L be an arc-presentation of an oriented link, and $L \mapsto L'$ a stabilization move. Then either*

$$w_+(L') = w_+(L) + 1, \qquad w_-(L') = w_-(L'),$$

or

$$w_-(L') = w_-(L) - 1, \qquad w_+(L') = w_+(L).$$

For the trivial arc-presentation L of the unknot, we have $w_+(L) = 1$, $w_-(L) = -1$. For the distant union $L_1 \sqcup L_2$ of two arc-presentations, we have $w_\pm(L_1 \sqcup L_2) = w_\pm(L_1) + w_\pm(L_2)$. In conjunction with Propositions 6, 7, 11, and 12, this implies the following.

COROLLARY 2. *If L is an arc-presentation of the trivial link of k components, then*

$$(15) \qquad\qquad w_-(L) \leq -k, \qquad k \leq w_+(L).$$

It happens very often that $w_-(L) \geq 0$ or $w_+(L) \leq 0$, which immediately implies that link L is non-trivial. For example, the torus link of type (p, q) has an arc-presentation $L_{p,q}$ whose arcs are

$$\gamma_{-\frac{2\pi k}{p+q}, \frac{2\pi k}{p+q}, \frac{2\pi(k+p)}{p+q}},$$

where $k = 1, \ldots, p + q$. A simple calculation gives

$$w_-(L_{p,q}) = pq - p - q, \qquad w_+(L_{p,q}) = pq.$$

If both p and q are large, then the interval (w_-, w_+) stays far away from 0, and this remains true after a certain number of modifications of the diagram of $L_{p,q}$.

References

[1] D. Bennequin, *Entrelacements et équations de Pfaff*, Astérisque 107-108 (1983), 87–161.

[2] J. Birman and M. Hirsch, *A new algorithm for recognizing the unknot*, Geom. Topol. 2 (1998), 175–220.

[3] J. Birman and W. Menasco, *Studying links via closed braids IV*: *Composite links and split links*, Invent. Math. 102 (1990), 115–139.

[4] —, —, *Studying links via closed braids V*: *Closed braid representatives of the unlink*, Trans. Amer. Math. Soc. 329 (1992), 585–606.

[5] H. Brunn, *Über verknotete Kurven*, in: Mathematiker-Kongress Zürich 1897, Leipzig, 1898, 256–259.

[6] P. Cromwell, *Embedding knots and links in an open book I*: *Basic properties*, Topology Appl. 64 (1995), 37–58

[7] P. R. Cromwell and I. J. Nutt, *Embedding knots and links in an open book II*: *bounds on arc index*, Math. Proc. Cambridge Philos. Soc. 119 (1996), 309–319.

[8] I. A. Dynnikov, *A new way to represent links. One-dimensional formalism and untangling technology*, Acta Appl. Math. 69 (2001), 243–283.

[9] M. Freedman, Zh.-X. He and Zh. Wang, *On the Möbius energy of knots and unknots*, Ann. Math. (2) 139 (1994), 1–50.

[10] W. Haken, *Theorie der Normalflächen. Ein Isotopiekriterium für der Kreisknoten*, Acta Math. 105 (1961), 245–375.

[11] A. E. Hatcher, *A proof of the Smale conjecture* $\mathrm{Diff}(S^3) \simeq O(4)$, Ann. of Math. 117 (1983), 553–607.

[12]   R. B. Kusner and J. M. Sullivan, *Möbius energies of knots and links*, *surfaces and submanifolds*, in: Geometric Topology (Athens, GA, 1993), AMS/IP Stud. Adv. Math., Amer. Math. Soc., Providence, RI, 1997, 570–604.

[13]   S. V. Matveev, *Algorithmic Topology and Classification of 3-Manifolds*, Springer 2003.

[14]   —, *Computer Recognition of Three-Manifolds*, Experimental Math. 7 (1998), 153–161.

[15]   K. Murasugi, *Jones polynomial and classical conjectures in knot theory*, Topology 26 (1987), 187–194.

[16]   J. O'Hara, *Energy of a knot*, Topology 30 (1991), 241–247.

[17]   V. Pinciu, *Positionings of essential surfaces in link complements*, PhD thesis, SUNY Buffalo, 1998.

[18]   H. Schubert, *Bestimmung der Primfaktorzerlegung von Verkettungen*, Math. Z. 76 (1961), 116–148.

Department of Mechanics and Mathematics
Moscow State University
Moscow 119992 GSP-2, Russia
E-mail: dynnikov@mech.math.msu.su