

AN INVARIANT OF REGULAR ISOTOPY

LOUIS H. KAUFFMAN

ABSTRACT. This paper studies a two-variable Laurent polynomial invariant of regular isotopy for classical unoriented knots and links. This invariant is denoted L_K for a link K , and it satisfies the axioms:

1. Regularly isotopic links receive the same polynomial.

2. $L_\emptyset = 1$.

3. $L \text{ (crossing) } = aL$, $L \text{ (negative crossing) } = a^{-1}L$.

4. $L \text{ (crossing) } + L \text{ (crossing) } = z(L \text{ (parallel) } + L \text{ (parallel) })$.

Small diagrams indicate otherwise identical parts of larger diagrams.

Regular isotopy is the equivalence relation generated by the Reidemeister moves of type II and type III. Invariants of ambient isotopy are obtained from L by writhe-normalization.

I. INTRODUCTION

In this paper I introduce a two variable Laurent polynomial invariant of knots and links in three-dimensional space. The primary version of this invariant, denoted L_K , is a *regular isotopy invariant* of unoriented knots and links. Regular isotopy will be explained below.

Associated to L_K are two normalized polynomials F_K , for K oriented, and U_K for K unoriented. These are each ambient isotopy invariants of K . Each is obtained from L_K by multiplying it by a normalizing factor. Since F_K is an invariant for oriented knots and links, it may be compared with the original Jones V_K -polynomial [13] and with the homfly polynomial P_K [10]. In fact, F_K has the Jones polynomial V_K as a special case (see §3). F_K is distinct from both the Jones and the homfly polynomials. The F -polynomial is good at distinguishing knots and links from their mirror images. In this regard, it appears to do somewhat better than the homfly polynomial. The F_K polynomial was announced in [24 and 19] and has become known as the Kauffman polynomial [16, 31, 37, 40, 45].

Our approach to these invariants is based upon the theorem [41] stating that two knot or link diagrams represent ambient isotopic embeddings in three-space if and only if the diagrams are related by a sequence of the three Reidemeister

Received by the editors July 13, 1988.

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

Research partially supported by NSF Grant #DMS-8701772.

© 1990 American Mathematical Society
0002-9947/90 \$1.00 + \$.25 per page

moves (denoted I, II, and III as in Figure 1) plus topological deformations of the diagram (that do not change the graphical structure of the crossings). Invariance is proved by examining the polynomial behaviour under these moves.

A knot is an embedding of a single circle into three-space, while a link is an embedding of a collection of circles. The number of circles for a link embedding is referred to as the number of *components* of the link. A knot is a link of one component.

A link diagram may be regarded as the shadow of a projection of the link into the plane. It is assumed that the singularities of the shadow are ordinary double points produced by transversally intersecting line segments. Each intersection is drawn to indicate over and under-crossing lines. The most common form of this convention is to indicate the under-crossing line by breaking it at the intersection.

Regular isotopy is the equivalence relation on diagrams generated by the Reidemeister moves of types II and III. There are a number of motivations for this formal definition. First of all, I choose the word *regular* because a regular isotopy projects to a regular homotopy of the underlying plane curves. Secondly, an embedded band (circle \times unit interval) is susceptible to ambient isotopies that correspond to the moves II and III on the core of the band. Thus regular isotopy of the core can be used in describing the knot theory of embeddings of bands (or of framed links). Through this correspondence, the L -polynomial can be regarded as an ambient isotopy invariant of embedded bands (see §5), or as a regular isotopy invariant of link diagrams.

The paper is organized as follows: §2 defines L and F and U , and gives examples. We give a short exposition of Morwen Thistlethwaite's proof using L_K that the twist number (the writhe) of a reduced alternating projection is an invariant of ambient isotopy. This solves an old conjecture [33, 43]. §3 discusses the relationship with the original Jones polynomial. §4 discusses the braid monoid [4, 20, 22, 33, 48] in relation to our invariant. We show that two loop-free elements of the braid monoid are regularly isotopic if and only if they are related by a sequence of specific types of algebraic moves. This generalizes known results for the Artin braid group, and provides a diagrammatic context for the Birman-Wenzel algebra [4]. §6 gives our proof of the well-definedness and invariance of the L -polynomial. Finally, §7 discusses a useful reformulation of L that we have called the Dubrovnik polynomial D , and a special case of D having a state expansion analogous to the bracket model of the Jones polynomial. This section also mentions extensions of the L -polynomial to polynomial invariants of embedded graphs in three-space.

I am happy to take this opportunity to thank Massimo Ferri of the Università di Bologna, Mario Rasetti of the Politecnico di Torino and Jon Simon of the University of Iowa for their hospitality during various stages of this work. Research for this paper was partially supported by ONR Grant #N0014084-K-0099 and NSF Grant #DMS-8701772.

II. BASIC RESULTS

In this section we will use regular isotopy as the equivalence relation generated by the Reidemeister moves of type II and III (plus topological equivalence of the underlying planar shadow). See Figure 1. The reader is referred to §5 for a discussion of regular isotopy in relation to embedded bands and framed links.

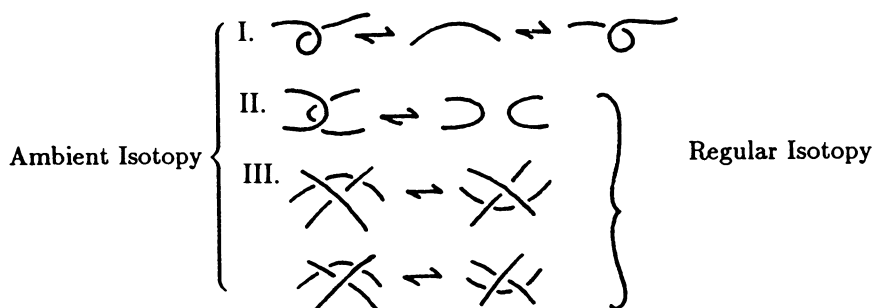


FIGURE 1. Reidemeister moves

A word about this equivalence relation (regular isotopy) is in order. First, we denote regular isotopy by \approx and ambient isotopy by \sim . ($A \approx B$ means A is regularly isotopic to B .) Second, the approach via diagrammatic moves goes back to Reidemeister and to his book *Knotentheorie* [41] where he showed that two knot diagrams are *combinatorially isotopic* if and only if one diagram can be obtained from the other by a sequence of the three move types of Figure 1. The notion of combinatorial isotopy (that Reidemeister used) is a special case of modern piecewise linear isotopy—and in fact generates piecewise linear isotopy for curves and graphs embedded in three-dimensional space. This fact was first proved in [11]. For a modern account of the equivalence of the diagrammatic relation generated by the Reidemeister moves and piecewise linear or differentiable isotopy for space curves, see [5].

Knot theory as the study of the formal diagrammatic system generated by the link diagrams and the Reidemeister moves is a long-standing approach to the subject. Alexander's original paper on his polynomial invariant of knots and links [1] used the Reidemeister moves. The remarkable fact about the new polynomial invariants (the Jones polynomial and generalizations) is that they depend crucially upon this diagrammatic approach.

There are a number of reasons for singling out Reidemeister moves II and III as a separate equivalence relation. The first Reidemeister move (I) is of a different character.

In Figure 2 I illustrate a regular isotopy that carries a curl through a crossing. This means that (by doing a few more moves of types II and III) potential type I moves can be relocated along a given component of the diagram. Thus a curl that apparently obstructs a type II or III move can be slid out of the way. This remark, in itself, does not show that one can do away with the type I move in the

theory of knots and links. This cannot be done. However, it is possible to work strictly with regular isotopy after normalizing the diagram in an appropriate way. (See [46] and §5 of this paper.)

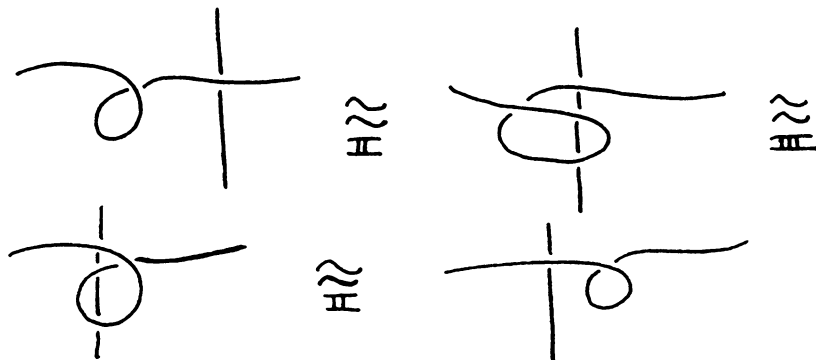


FIGURE 2

The simplest invariant of regular isotopy for oriented links K is the *writhe number* (or writhe), $w(K)$. This is the sum of the signs of all the crossings:

$$w(K) = \sum_p \varepsilon(p)$$

where the crossing signs are $+1$ or -1 as indicated in Figure 3. Note that the twist number changes by adding $+1$ or -1 under type I moves.

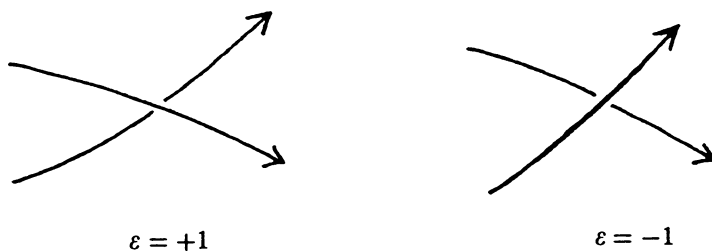


FIGURE 3. Crossing signs

The writhe can be used to normalize a regular isotopy invariant, producing a corresponding invariant of ambient isotopy:

Lemma 2.1. *Let \mathcal{R} be a ring, and a an invertible element of \mathcal{R} . Suppose that $R(K) \in \mathcal{R}$ is a regular isotopy invariant of oriented link diagrams K satisfying*

$$R(\text{positive crossing}) = aR(\text{negative crossing}),$$

$$R(\text{negative crossing}) = a^{-1}R(\text{positive crossing}).$$

Then $S(K) = a^{-w(K)} R(K)$ is an invariant of ambient isotopy, where $w(K)$ denotes the writhe of K .

Note. Here and elsewhere in this paper, small diagrams stand for parts of larger diagrams. A collection of small diagrams occurring in a single equation all share the same larger diagram. (The large diagram is only changed as indicated by the small diagrams.)

Proof. Since $w(K)$ is a regular isotopy invariant, so is $a^{-w(K)}$. Hence $S(K)$, being the product of two regular isotopy invariants is itself a regular isotopy invariant. Therefore it is only necessary to check the invariance of $S(K)$ under type I moves. Since

$$w(\text{positive crossing}) = 1 + w(\text{negative crossing})$$

and

$$w(\text{negative crossing}) = -1 + w(\text{positive crossing}),$$

this follows at once from the definition of $S(K)$.

In light of this lemma we make the following axiomatic definition of a 2-variable Laurent polynomial $L_K = L_K(a, z) \in \mathcal{R} = \mathbb{Z}[a, a^{-1}, z, z^{-1}]$:

Definition 2.2. Axioms for L_K .

0. If K is an unoriented link diagram, then $L_K \in \mathbb{Z}[a, a^{-1}, z, z^{-1}] = \mathcal{R}$ (the Laurent polynomials over the integers with commuting, independent variables a and z).

1. If K and K' are regularly isotopic diagrams, then $L_K = L_{K'}$ in \mathcal{R} .

2. The following identity holds for all quadruplets of diagrams that are identical everywhere except at the small diagrams indicated below:

$$3. \quad L \text{ (crossing) } + L \text{ (crossing) } = z(L \text{ (cup) } + L \text{ (cap) }).$$

$$L \text{ (positive crossing) } = aL,$$

$$L \text{ (negative crossing) } = a^{-1}L,$$

$$L \text{ (circle) } = 1.$$

In §6 we will prove

Theorem 2.3. The axioms of Definition 2.2 describe uniquely a well-defined Laurent polynomial invariant of regular isotopy, L_K , for unoriented link diagrams K .

By axiom 3, the restriction $Q_K(z) = L_K(1, z) \in \mathbb{Z}[z, z^{-1}]$ is an invariant of ambient isotopy. This is the polynomial Q of Brandt, Lickorish, and Millett [6] and Ho [12]. Thus our polynomial is a two-variable generalization of the Q -polynomial.

Definition 2.4. Define a polynomial $F_K \in \mathcal{R} = \mathbb{Z}[a, a^{-1}, z, z^{-1}]$ for oriented link diagrams K via the equation $F_K = a^{-w(K)} L_K$ where L_K is defined on oriented link diagrams by forgetting the orientation.

Proposition 2.5. The polynomial $F_K(a, z) \in \mathbb{Z}[a, a^{-1}, z, z^{-1}]$ is an ambient isotopy invariant of oriented links K .

Proof. This follows at once from Lemma 2.1 ($R_K = L_K$, $S_K = F_K$) and Theorem 2.3.

Remark. When speaking of regular isotopy I shall take care to speak of diagrams. When speaking of ambient isotopy, one can speak of links as embeddings of circles in three-dimensional space. Nevertheless, F_K is calculated through the use of diagrams. It is an open problem to give a definition of F_K that is not diagram dependent.

Definition 2.5.1. The *mirror image* of a diagram K , denoted $K!$, is that diagram obtained from K by reversing all of the crossings of K . (Hence it is the mirror image obtained by reflection in the plane of projection.) Orientations on K are not changed under this mirror imaging. A link K is said to be *achiral* if it is ambient isotopic to its mirror image $K!$. When this is not the case, then K is *chiral*.

In practice it is necessary to distinguish between chirality with or without orientation. For example the knot 8_{17} is known to be ambient isotopic to its mirror image with reversed orientation, but it is chiral in the oriented case.

Lemma 2.6. Let K be an (oriented) link and $K!$ its mirror image. Then the polynomials L_K and F_K undergo replacement of a by a^{-1} when K is replaced by $K!$.

$$L_{K!}(a, z) = L_K(a^{-1}, z), \quad F_{K!}(a, z) = F_K(a^{-1}, z).$$

Proof. $K!$ is obtained from K by reversing all crossings. This has the effect of interchanging a and a^{-1} in Axiom 4 of Definition 2.2 (and $w(K!) = -w(K)$). Hence any calculation of $L_{K!}(F_{K!})$ will result in an identical calculation of $L_K(F_K)$ with a replaced by a^{-1} . This completes the proof.

Remark. The upshot of this lemma is that if $F_K(a, z) \neq F_K(a^{-1}, z)$ then K is not ambient isotopic to its mirror image.

Figure 4 illustrates the calculation of L_K for the first few knots and links directly from the axioms. Note that the unlink consisting of two simple closed curves receives the value $d = z^{-1}(a + a^{-1}) + 1$. In this case L and F coincide. The value of F_T for the right-handed trefoil of Figure 4 is

$$F_T = (a^{-2} + a^{-4})z^2 + (a^{-3} + a^{-5})z + (-2a^{-2} - a^{-4}).$$

Since $F_T(a, z) \neq F_T(a^{-1}, z)$, this shows that F detects the chirality of the trefoil knot.

$$\begin{aligned}
 & L \text{ (crossing) } + L \text{ (crossing) } = z(L \text{ (cup) } + L \text{ (cap) }), \\
 (1) \quad & aL \text{ (cup) } + a^{-1}L \text{ (cap) } = z(\delta + L \text{ (cup) }), \delta = L \text{ (crossing) },
 \end{aligned}$$

$$\Rightarrow \delta = (a + a^{-1})z^{-1} - 1,$$

$$\begin{aligned}
 & L \text{ (link) } + L \text{ (link) } = z(L \text{ (link) } + L \text{ (link) }), \\
 (2) \quad & \lambda + \delta = z(a + a^{-1}), \lambda = L \text{ (link) },
 \end{aligned}$$

$$\lambda = -(a + a^{-1})z^{-1} + 1 + (a + a^{-1})z,$$

$$\begin{aligned}
 & L \text{ (link) } + L \text{ (link) } = z(L \text{ (link) } + L \text{ (link) }), \\
 (3) \quad & \tau + a = z(\lambda + a^{-2}), \tau = L \text{ (link) },
 \end{aligned}$$

$$\tau = (-2a - a^{-1}) + (1 + a^{-2})z + (a + a^{-1})z^2.$$

FIGURE 4

Remark. Note that in inductive calculations such as those indicated in Figure 4, the evaluations involving the variable a occur at the bottom of the tree of evaluations. By an evaluation tree I mean a tree of knots and links obtained by switching (S) and elimination (E and e) as illustrated in Figure 5. The operation of switching a crossing is denoted S ($S(\bowtie) = \times$) and the operations that eliminate a crossing are denoted $E(E(\bowtie)_i = \searrow, \swarrow)$ and $e(e(\bowtie)_i = \searrow, \swarrow)$. Our method (in §6) for seeing that L_K is well defined involves a specific choice of evaluation tree. (This is also called a skein decomposition.)

Definition 2.7. A diagram is *split* if it is displayed as the disjoint union of two diagrams. We write $K = K' \sqcup K''$ where K' and K'' denote the two sub-diagrams. A diagram K is a *connected sum* if it is displayed as two disjoint diagrams connected by parallel embedded arcs (up to planar isotopy) as in Figure 6. By cutting these arcs and resplicing them, a split link $\tilde{K} = K' \sqcup K''$ is obtained. We write $K = K' \# K''$ for the connected sum.

Lemma 2.8. *The following formulas hold for the polynomials L and F with respect to disjoint union and connected sum:*

$$L_{K \# K'} = L_K L_{K'}, \quad L_{K \sqcup K'} = d L_K L_{K'},$$

$$F_{K\#K'} = F_K F_{K'}, \quad F_{K\sqcup K'} = d F_K F_{K'},$$

where $d = z^{-1}(a + a^{-1}) - 1$.

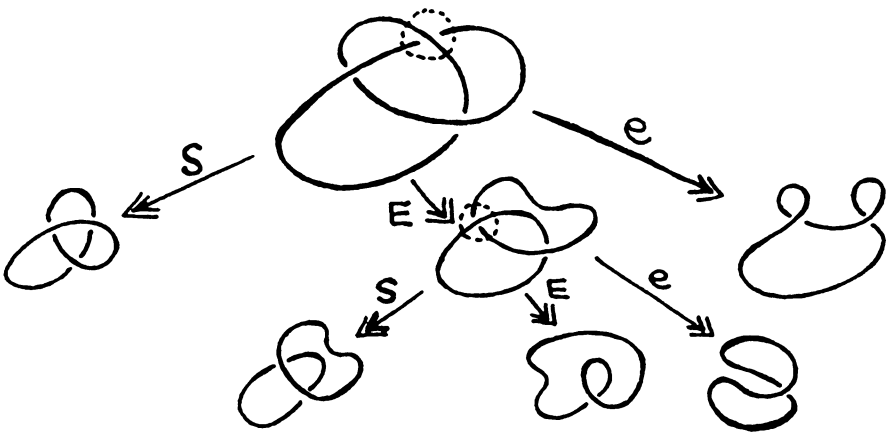


FIGURE 5. An evaluation tree for the trefoil

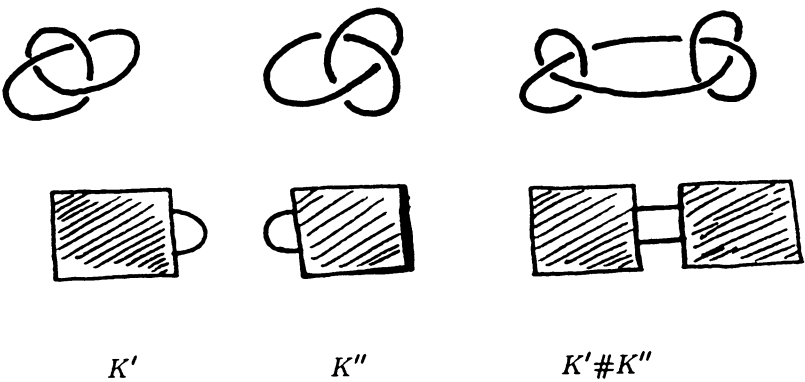


FIGURE 6

The proof is by straightforward induction and is omitted.

Figure 7 lists the values of L and the corresponding twist numbers for the first few knots. (A table for L up to 9 crossings may be found in [19].) Note that, in this figure, the highest degree term (for the variable z in L) has the form $k(a + a^{-1})z^{n-1}$ where n is the number of crossings of the knot, and k is a positive integer. This is not always the case as shown by the knot 8_{19} of Figure 8 (whose leading term in L is $(1 + a^{-2})z^6$). However, 8_{19} is also the first nonalternating knot.



$$|w = 3|$$

$$L = (a + a^{-1})z^2 + (1 + a^{-2})z^1 + (-2a - a^{-1})z^0$$



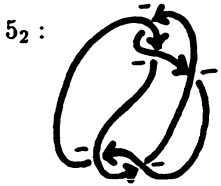
$$|w = 0|$$

$$L = (a + a^{-1})z^3 + (a^2 + 2 + a^{-2})z^2 \\ + (-a - a^{-1})z^1 + (-a^2 - 1 - a^{-2})z^0$$



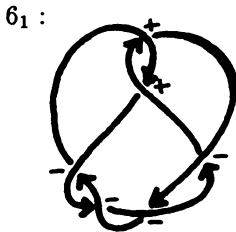
$$|w = 5|$$

$$L = (a + a^{-1})z^4 + (1 + a^{-2})z^3 \\ + (-4a - 3a^{-1} + a^{-3})z^2 \\ + (-2 - a^{-2} + a^{-4})z^1 + (3a + 2a^{-1})z^0$$



$$|w = -5|$$

$$L = (a + a^{-1})z^4 + (a^2 + 2 + a^{-2})z^3 \\ + (-2a - a^{-1} + a^{-3})z^2 \\ + (-2a^2 - 2)z^1 + (a + a^{-1} - a^{-3})z^0$$



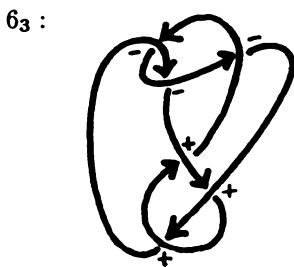
$$|w = -2|$$

$$L = (a + a^{-1})z^5 + (a^2 + 2 + a^{-2})z^4 \\ + (-3a - 2a^{-1} + a^{-3})z^3 + (-3a^2 - 4 + a^{-4})z^2 \\ + (2a + 2a^{-1})z^1 + (a^2 + 1 - a^{-4})z^0$$



$$|w = 2|$$

$$L = (a + a^{-1})z^5 + (a^2 + 3 + 2a^{-2})z^4 \\ + (-2a + 2a^{-3})z^3 + (-3a^2 - 6 - 2a^{-2} + a^{-4})z^2 \\ + (-a^{-1} - a^{-3})z^1 + (2a^2 + 2 + a^{-2})z^0$$

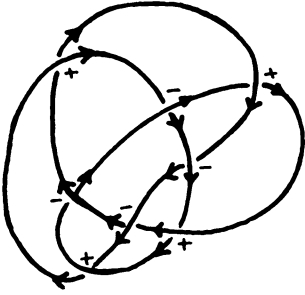


$$|w = 0| \quad |6_3 \text{ is achiral}|$$

$$L = (a + a^{-1})z^5 + (2a^2 + 4 + 2a^{-2})z^4 \\ + (a^3 + a + a^{-1} + a^{-3})z^3 \\ + (-3a^2 - 6 - 3a^{-2})z^2 \\ + (-a^3 - 2a - 2a^{-1} - a^{-3})z^1 \\ + (a^2 + 3 + a^{-2})z^0$$

FIGURE 7

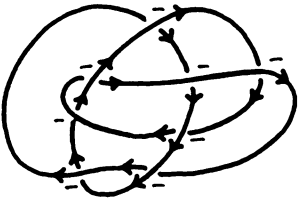
8_{18} :



$|w = 0|$ $|8_{18}$ is achiral

$$\begin{aligned} L = & (3a + 3a^{-1})z^7 + (6a^2 + 12 + 6a^{-2})z^6 \\ & + (4a^3 + 3a + 3a^{-1} + 4a^{-3})z^5 \\ & + (a^4 - 9a^2 - 20 - 9a^{-2} + a^{-4})z^4 \\ & + (-4a^3 - 9a - 9a^{-1} - 4a^{-3})z^3 \\ & + (3a^2 + 6 + 3a^{-2})z^2 + (a + a^{-1})z^1 \\ & + (a^2 + 3 + a^{-2})z^0 \end{aligned}$$

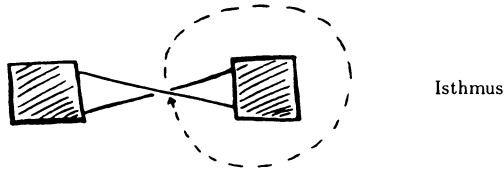
8_{19} :



$|w = -8|$

$$\begin{aligned} L = & (1 + a^{-2})z^6 + (a + a^{-1})z^5 + (-6 - 6a^{-2})z^4 \\ & + (-5a - 5a^{-1})z^3 + (10 + 10a^{-2})z^2 \\ & + (5a + 5a^{-1})z^1 + (-a^2 - 5 - 5a^{-2})z^0 \end{aligned}$$

FIGURE 8. 8_{19} is the first nonalternating knot



Isthmus

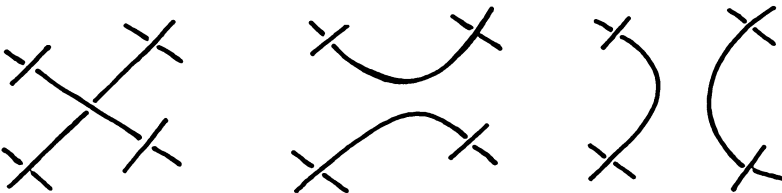


FIGURE 9. Splicing preserves alternation

That is, 8_{19} admits no projection (diagram) that is alternating (in the sense that the weave alternates in the pattern under-over-under-over \dots as one travels along any strand).

Thistlethwaite [45] proved that the top term of L has the form $k(a + a^{-1})z^{n-1}$ for reduced alternating projections. A projection is reduced if it has no isthmus as in Figure 9. From this follows the theorem.

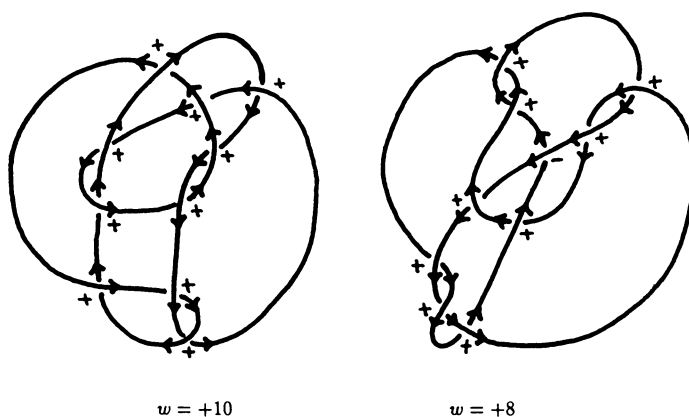


FIGURE 10. The Perko pair. Ambient isotopic, reduced projections with different writhe

Theorem 2.9. *The writhe (twist number) $w(K)$ is an ambient isotopy invariant for reduced alternating projections K . Hence any reduced alternating link of nonzero writhe is chiral.*

Proof. Let K be reduced and alternating. Then, by [45], we have

$$L_K = k(a + a^{-1})z^{n-1} + g(a, z)$$

where g has z -degree less than $n - 1$, and $k > 0$. Hence

$$F_K = a^{-w(K)} L_K = k(a^{1-w(K)} + a^{-1-w(K)})z^{n-1} + a^{-w(K)} g(a, z).$$

This shows that $1 - w(K)$ is an ambient isotopy invariant of K , since F_K is an ambient isotopy invariant. Therefore, $w(K)$ is an invariant of ambient isotopy for K reduced and alternating.

Since $w(K!) = -w(K)$ when $K!$ is the mirror image of K (Definition 2.2), the second part of the theorem follows from the invariance of $w(K)$. This completes the proof.

The invariance of the writhe was first conjectured by Tait, Little, and Kirkman [43, 33, 28] in their original compilations of the knot tables. These pioneers thought, however, that the writhe was an ambient isotopy invariant for *all* reduced projections. Their intuition was on target for alternating diagrams.

It has been known for some time that there are reduced nonalternating projections of the same knot, so that the projections differ in writhe. Figure 10 illustrates the “Perko pair”, two ten crossing knots of writhe 10 and 8 respectively, that are ambient isotopic.

The key to the induction in Thistlethwaite’s theorem is that:

1. At least one of the crossing elimination diagrams obtained from a reduced alternating diagram is itself a reduced diagram.
2. Each crossing elimination from an alternating link diagram yields an alternating diagram.

These facts can be understood from the splicing diagram of Figure 9. One then uses 1 and 2 to show, inductively, that the top term of L_K , for a reduced alternating diagram, inherits the form $T = k(a + a^{-1})z^{n-1}$.

To return to chirality, there are many achiral alternating knots and links of zero writhe. I conjectured [22] that a reduced alternating achiral projection has isomorphic graph and dual graph. The graph of a projection is the graph associated with the shaded regions in a checkerboard coloring of the diagram. Figure 11 illustrates this phenomenon for the case of the achiral knot 6_3 . In this figure, the graph and dual graph are G and G^* , respectively. The truth of this conjecture would have settled completely the matter of chirality for reduced alternating knots up to the problem: *determine when a plane graph is isomorphic with its dual graph* (the dual graph is obtained by making regions into vertices, and creating edges where regions share an edge of the original).

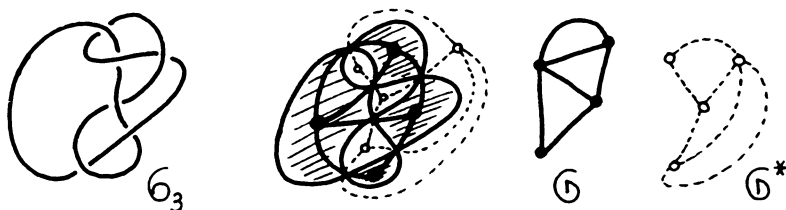


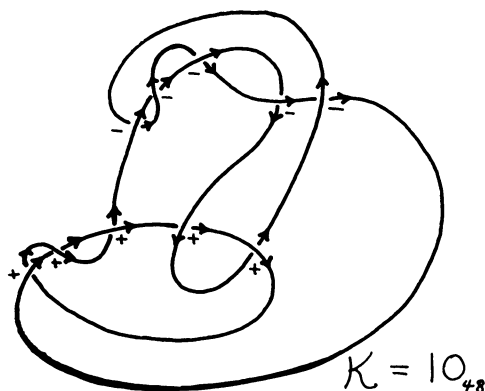
FIGURE 11

K. Murasugi has observed to the author that this conjecture is false. It seems that the graph and dual graph, while not in general isomorphic, are related by flying moves. I therefore now conjecture that (under the same hypotheses) *the cycle matroids of the graph and dual graph are isomorphic*.

Finally, the knot projection shown in Figure 12 is reduced alternating and of zero writhe. However its L -polynomial shows that it is chiral. This is the first instance of chirality detected by F and not by the homfly polynomial P [10]. So far, F detects chirality whenever the homfly polynomial detects chirality. F does better in many cases.

As expected, F does not always detect chirality. There is a small literature of calculations and examples. For example, Kanenobu and Sakuma point out [16] that the pair of knots K and K' of Figure 13 share the same polynomial ($F_K = F_{K'}$) while K is achiral but K' is chiral. It is easy to see that $F_K = F_{K'}$ because K' is a *mutant* of K (for mutants see e.g., [29]) and the same arguments for invariance under mutation that apply to the homfly and Conway polynomials also apply to F .

The polynomial U_K . Finally, I mention the normalized version of L_K for unoriented knots and links. Given an unoriented link K define the *self-writhe* $s(K)$ to be the sum of the crossing signs of *self-crossings of components of K* for any orientation of K . It is easy to see that $s(K)$ is well defined, independent



$$w(K) = 0$$

First knot where chirality is detected by F_K , but not by P_K (the Homfly polynomial).

$$\begin{aligned}
 F_K &= a^{-w(K)} L_K = L_K \quad (w(K) = 0) \\
 L_K &= (a + a^{-1})z^9 + (2a^2 + 5 + 3a^{-2})z^8 \\
 &\quad + (2a^3 + a^{-1} + 3a^{-3})z^7 \\
 &\quad + (2a^4 - 5a^2 - 20 - 11a^{-2} + 2a^{-4})z^6 \\
 &\quad + (a^5 - 3a^3 - 5a - 11a^{-1} - 9a^{-3} + a^{-5})z^5 \\
 &\quad + (-5a^4 + 9a^2 + 37 + 18a^{-2} - 5a^{-4})z^4 \\
 &\quad + (-3a^5 - a^3 + 12a + 21a^{-1} + 8a^{-3} - 3a^{-5})z^3 \\
 &\quad + (2a^4 - 11a^2 - 27 - 13a^{-2} + a^{-4})z^2 \\
 &\quad + (2a^5 - 7a - 9a^{-1} - 3a^{-3} + a^{-5})z^1 \\
 &\quad + (4a^2 + 9 + 4a^{-2})z^0 \\
 \langle K \rangle &= -A^{20} + 2A^{16} - 4A^{12} + 6A^8 - 7A^4 + 9 \\
 &\quad - A^{-20} + 2A^{-16} - 4A^{-12} + 6A^{-8} - 7A^{-4}
 \end{aligned}$$

FIGURE 12

of orientation, and that it is an invariant of regular isotopy. The polynomial U_K is defined by the formula $U_K = a^{-s(K)} L_K$ and is an ambient isotopy invariant of unoriented links. U_K and F_K differ by a power of a involving linking numbers of the components of K . I leave the exact relationship as an exercise for the reader. I am indebted to Cameron Gordon for pointing out to me this use of the self-writhe.

III. THE JONES POLYNOMIAL

In this section we show that the original Jones V -polynomial [13] is a special case of two-variable polynomial F_K .

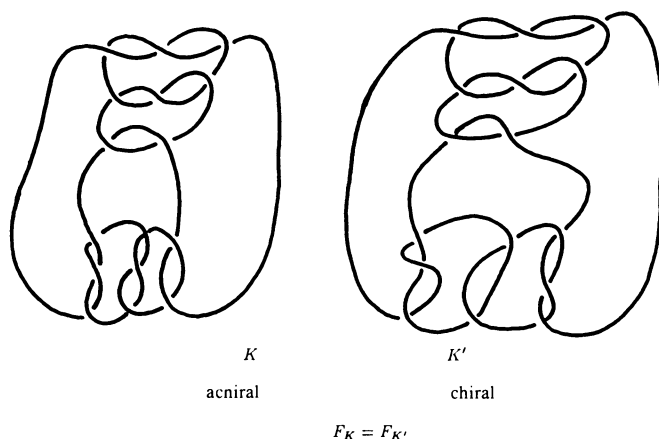


FIGURE 13

Recall that the V -polynomial is determined by the following properties:

1. For an oriented link K , $V_K(t) \in \mathbb{Z}[t, t^{-1}]$ is an ambient isotopy invariant of K .

2. $V_{\bigcirc} = 1$.

3. $t^{-1}V_{\nearrow \searrow} - tV_{\searrow \nearrow} = (\sqrt{t} - 1/\sqrt{t})V_{\rightarrow \leftarrow}$.

The V -polynomial, as originally defined by Jones, involves the trace of a representation of the Artin braid group to a von Neumann algebra. The three properties shown above are a consequence of this definition, and they can be used as a starting point both for defining and computing the polynomial. In [20] I show how to give an alternative construction of the V -polynomial, in the spirit of direct recursion and regular isotopy. It is this “bracket” model that we shall use, to show that V_K is a special case of F_K .

Before going into the details, it must be remarked that the V -polynomial was the first polynomial invariant of knots and links capable of detecting chirality.

The *bracket polynomial* [20], denoted $\langle K \rangle$, is defined for unoriented knots and links K as a regular isotopy invariant. The defining relations are:

1.

$$\begin{aligned}\langle \nearrow \searrow \rangle &= A \langle \equiv \rangle + B \langle \searrow \nearrow \rangle, \\ \langle \searrow \nearrow \rangle &= B \langle \equiv \rangle + A \langle \searrow \nearrow \rangle.\end{aligned}$$

2.

$$\langle O \sqcup K \rangle = d \langle K \rangle, \quad \langle O \rangle = 1.$$

These two statements make $\langle K \rangle$ an element of the ring $\mathbb{Z}[A, B, d]$ (commuting independent variables A , B and d). With no restriction on A , B and d the polynomial $\langle K \rangle$ is well defined on link-diagrams. The recursive definition can be reexpressed as a direct summation over *states* S of the diagram K . A state S is a diagram obtained from K by eliminating each crossing K by a *splice of type A* ($\nearrow \searrow \Rightarrow \equiv$) or a *splice of type B* ($\searrow \nearrow \Rightarrow \searrow \nearrow$). Thus S consists of a

number $(|S|)$ of disjoint simple closed curves in the plane. Let $i_A(S)$ denote the number of A -splices used to obtain S , and $i_B(S)$ the number of B -splices to obtain S . Then, as in [20], we have the formula

$$\langle K \rangle = \sum_S A^{i_A(S)} B^{i_B(S)} d^{|S|-1}$$

where the sum is taken over all 2^n states of the diagram K (with n crossings).

In this form $\langle K \rangle$ is not invariant under Reidemeister moves, however the following formula [20] shows what is needed to create a regular isotopy invariant:

$$\langle \text{crossing} \rangle = AB \langle \text{two crossings} \rangle + (ABd + A^2 + B^2) \langle \text{parallel strands} \rangle.$$

By setting $B = A^{-1}$ and $d = -A^2 - A^{-2}$, $\langle K \rangle$ becomes an invariant of the Reidemeister II move. Another observation [20] then shows that $\langle K \rangle$ is also invariant under the move of type III. Thus $\langle K \rangle$ is an invariant of regular isotopy.

Along with this, it is easy to calculate that $\langle \text{curl} \rangle = (-A^3) \langle \text{uncurl} \rangle$ and $\langle \text{twist} \rangle = (-A^{-3}) \langle \text{untwist} \rangle$. Thus the associated invariant of ambient isotopy for $\langle K \rangle$ (via Lemma 2.1) is the Laurent polynomial defined by the formula

$$f_K(A) = (-A^3)^{-w(K)} \langle K \rangle$$

where $w(K)$ is the writhe of K and $\langle K \rangle$ is defined for oriented links by forgetting the orientation.

With this background, it is easy to delineate the relationship between the Jones V -polynomial and the polynomial F_K .

Proposition 3.1. *The Jones polynomial is (up to a change of variable) the ambient isotopy invariant associated with the bracket polynomial. In particular, $V_K(t) = f_K(t^{-1/4})$ where $f_K = (-A^3)^{-w(K)} \langle K \rangle$ is defined as above.*

See [20 or 22] for a proof of Proposition 3.1. Here is the idea: Given that

$$(*) \begin{cases} \langle \text{crossing} \rangle = A \langle \text{parallel strands} \rangle + B \langle \text{two crossings} \rangle, \\ \langle \text{crossing} \rangle = B \langle \text{parallel strands} \rangle + A \langle \text{two crossings} \rangle. \end{cases}$$

Multiply the first equation by B^{-1} , the second by A^{-1} , subtract and obtain

$$B^{-1} \langle \text{crossing} \rangle - A^{-1} \langle \text{crossing} \rangle = (A/B - B/A) \langle \text{parallel strands} \rangle.$$

Write the corresponding equation for f_K by multiplying by the appropriate writhe terms and observe that the form of identity 3 for the Jones polynomial emerges after the substitution $A = t^{-1/4}$ (with $B = A^{-1}$).

Proposition 3.2. $\langle K \rangle$ is a special case of L_K and f_K is a special case of F_K . In particular, for $L_K(a, z)$ and $F_K(a, z)$ we have

$$\begin{aligned}\langle K \rangle(A) &= L_K(-A^3, A + A^{-1}), \\ f_K(A) &= F_K(-A^3, A + A^{-1}).\end{aligned}$$

Hence $V_K(t) = F_K(-t^{-3/4}, t^{-1/4} + t^{1/4})$, giving the Jones polynomial as a special case of the F -polynomial.

Proof. Adding the equations (*) for the bracket, we obtain

$$\langle \text{positive curl} \rangle + \langle \text{negative curl} \rangle = (A + B) \left(\langle \text{positive curl} \rangle + \langle \text{negative curl} \rangle \right).$$

This equation, plus the fact that (for $B = A^{-1}$, $d = -A^2 - A^{-2}$) $\langle K \rangle$ multiplies by $(-A^3)$ under a positive curl, shows that $\langle K \rangle = L_K(-A^3, A + A^{-1})$. The rest of the proposition follows at once from the remarks preceding it.

Remark. In [31] Lickorish has given a different proof that the Jones polynomial is a special case of the polynomial F_K .

In fact, the construction of the bracket polynomial is intimately related to the original construction of the Jones polynomial via von Neumann algebras, and also to the Potts model in statistical physics. We shall not treat the Potts model here. See [23]. As for the relation with braids and algebras, this will be the subject of the remainder of this section, and of the next section of the paper.

If b is a braid, let $\langle b \rangle$ denote the value of the bracket on the corresponding closed braid \bar{b} obtained by identifying input strands with output strands as shown in Figure 14.

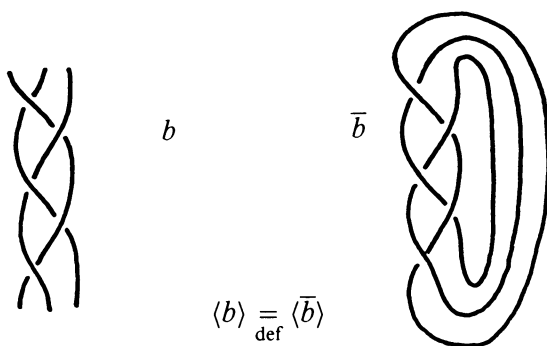


FIGURE 14

Then $\langle b \rangle = \langle \bar{b} \rangle$ and we can write the latter as a sum over the states of the closed braid diagram. These states are obtained, as described at the beginning of this section, by splicing out crossings from the diagram, $b(\bar{b})$. As illustrated in Figure 15, one form of splice (vertical) will replace a braid generator by an

identity braid, while the other (horizontal) will replace a braid generator by a “hook”. The hook is characterized by a connection between two top adjacent input stands, and two bottom adjacent output strands, as in the form $\begin{array}{c} \diagup \quad \diagdown \\ \diagdown \quad \diagup \end{array}$.

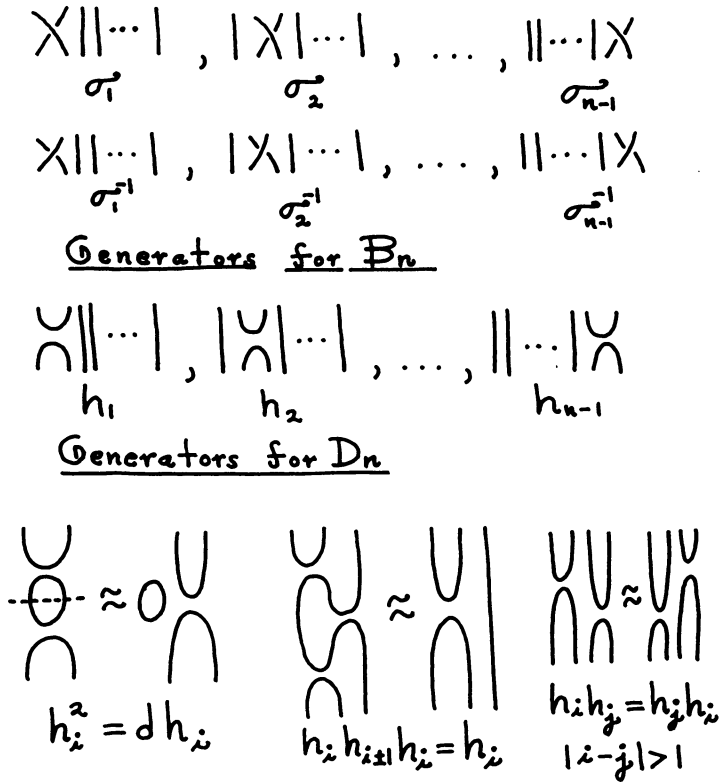


FIGURE 15. Diagram monoid relations

These hooks have an algebraic structure if multiplied like braids. Figure 15 shows the formal results of such multiplication. Each hook is an n -strand planar tangle with two paired arcs (shown as nearly touching at their corresponding maximal and minima). The remaining arcs (not at input/outputs i or $i+1$ for the hook h_i) proceed downward from top to bottom of the tangle, connecting the j th input with the j th output ($j \neq i, i+1$) without intersecting one another.

In this discussion *all tangles will be considered up to regular isotopy relative to the top and bottom of the tangle.*

With this terminology we see that the A -split of σ_i is h_i and the B -split of σ_i is 1. Conversely, the B -split of σ_i^{-1} is 1 while the A -split of σ_i^{-1} is h_i .

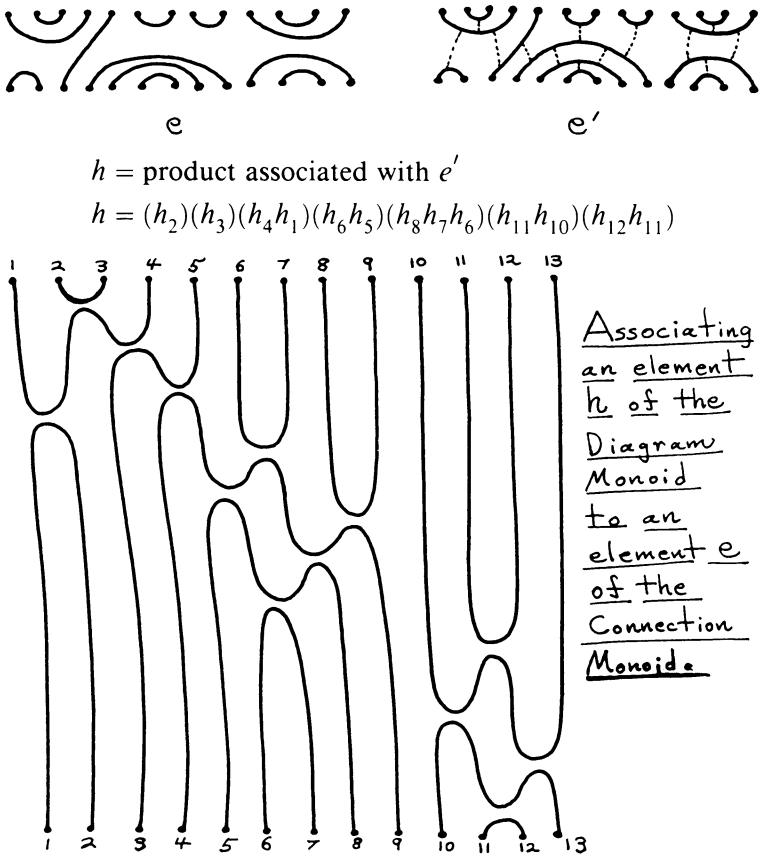


FIGURE 16

We have

$$\begin{aligned} \langle \sigma_i \rangle &= \langle \parallel \cdots \mid \times \mid \cdots \parallel \rangle \\ &= A \langle \parallel \cdots \mid \smile \mid \cdots \parallel \rangle + B \langle \parallel \cdots \mid \mid \mid \cdots \parallel \rangle, \\ \langle \sigma_i \rangle &= A \langle \sigma_i \rangle + B \langle 1 \rangle, \end{aligned}$$

and similarly,

$$\langle \sigma_i^{-1} \rangle = A \langle 1 \rangle + B \langle h_i \rangle.$$

Note that the value of the bracket on the identity braid is d^{n-1} where n is the number of braid strands. For example,

$$\langle \parallel \rangle = \langle \overline{\parallel} \rangle = \langle \bigcirc \rangle = d^2.$$

The states of \bar{b} are therefore in one-to-one correspondence with certain products of hooks in this diagram algebra. Since simple closed curves can appear in the braid states (for example on squaring h_i as shown in Figure 15) before closure (the closure of b is \bar{b}), it is natural to allow one more element δ corresponding to such a curve. All such curves can be moved out of the location

of the tangle via regular isotopy as in

$$h^2 = \bigcirc \approx 0 \bigcap = \delta h.$$

Thus we collect them and write δ^s for s curves, and $H = \delta^s H'$ where H' is a loop-free tangle (for a given product of hooks H).

A very similar structure was invented by R. Brauer in a 1937 paper [7] on representations of groups.

Definition 3.3. The (diagrammatic) *Braid Monoid* BM_n is the monoid formed by products of the generators of the Artin n -strand braid group, the hooks for n -strands, and the simple closed curve δ —under standard tangle multiplication. Two elements of the monoid are *equivalent* if they are regularly isotopic relative to the endpoints of the tangle.

A complete set of relations for the braid monoid will be determined in the next section.

In the remainder of the paper we shall refer to the submonoid $DM_n \subset BM_n$ that is generated by the hooks and the simple closed curve d as the *diagram monoid*. As remarked above, members of the diagram monoid correspond to states of elements of the braid group. Thus the bracket is computed as a sum of evaluations of elements of the diagram monoid.

Definition 3.4. Given a commutative ring \mathcal{R} and a monoid \mathcal{M} , let \mathcal{RM} denote the free additive algebra over \mathcal{R} with multiplication generated by \mathcal{M} . That is, an element of \mathcal{RM} is a formal linear combination $r_1 m_1 + \cdots + r_n m_n$ with $r_i \in \mathcal{R}$ and $m_i \in \mathcal{M}$. Multiplication follows the distributive law ($a(b+c) = ab+ac$) and $(r_i m_i)(r_j m_j) = (r_i r_j)(m_i m_j)$ where the product $r_i r_j$ is in the ring \mathcal{R} , and the product $m_i m_j$ is in the monoid \mathcal{M} .

Definition 3.5. Let $\mathcal{R} = \mathbb{Z}[A, A^{-1}]$, the ring of Laurent polynomials in A and A^{-1} . Let SD_n denote the quotient of the algebra RDM_n by the equivalence relation generated by setting $\delta = -A^2 - A^{-2}$ (δ is the loop element in DM_n). Thus for h_i in SD_n we have $h_i^2 = -(A^2 + A^{-2})h_i$. I will continue to write elements h_i of SD_n without change of notation from DM_n . The algebra SD_n is called the (n -strand) *diagram algebra*.

Proposition 3.6. Define a mapping $p: B_n \rightarrow SD_n$ from the n -strand Artin braid group to the n -strand diagram algebra by defining it on generators of the braid group by the formulas

$$p(\sigma_i) = Ah_i + A^{-1}, \quad p(\sigma_i^{-1}) = A^{-1}h_i + A.$$

Then this definition extends uniquely to a representation of the braid group to the diagram algebra.

Proof. It will suffice to show that

- (1) $p(\sigma_i)p(\sigma_i^{-1}) = 1 \in SD_n$ and that
- (2) $p(\sigma_i)p(\sigma_{i+1})p(\sigma_i) = p(\sigma_{i+1})p(\sigma_i)p(\sigma_{i+1})$ and
- (3) $p(\sigma_i)p(\sigma_j) = p(\sigma_j)p(\sigma_i)$ for $|i - j| > 1$

since these equations are the image under p of the relations in B_n [3].

We shall check these facts by noting that (1) and (3) are easy to verify directly and that (2) follows from the relation $h_i h_{i+1} h_i = h_i$ in the diagram algebra (see Figure 15).

$$(1) \quad p(\sigma_i)p(\sigma_i^{-1}) = (Ah_i + A^{-1})(A^{-1}h_i + A) = h_i^2 + 1 + (A^2 + A^{-2})h_i \\ = -(A^2 + A^{-2})h_i + 1 + (A^2 + A^{-2})h_i = 1.$$

$$(2) \quad p(\sigma_i)p(\sigma_{i+1})p(\sigma_i) = (Ah_i + A^{-1})(Ah_{i+1} + A^{-1})(Ah_i + A^{-1}) \\ = (A^2 h_i h_{i+1} + h_i + h_{i+1} + A^{-2})(Ah_i + A^{-1}) \\ = A^3 h_i h_{i+1} h_i + Ah_i^2 + Ah_{i+1} h_i + A^{-1} h_i + Ah_i h_{i+1} \\ + A^{-1} h_i + A^{-1} h_{i+1} + A^{-3} \\ = A^3 h_i + A(-A^2 - A^{-2})h_i + A(h_{i+1} h_i + h_i h_{i+1}) \\ + 2A^{-1} h_i + A^{-1} h_{i+1} + A^{-3} \\ = A(h_{i+1} h_i + h_i h_{i+1}) + A^{-1}(h_i + h_{i+1}) + A^{-3}.$$

Since this last expression is invariant under the interchange of i and $i + 1$, we conclude that $p(\sigma_i)p(\sigma_{i+1})p(\sigma_i) = p(\sigma_{i+1})p(\sigma_i)p(\sigma_{i+1})$.

(3) Since $h_i h_j = h_j h_i$ for $|i - j| > 1$ (Figure 15), we conclude that $p(\sigma_i)p(\sigma_j) = p(\sigma_j)p(\sigma_i)$ for $|i - j| > 1$.

This completes the proof of Proposition 3.6.

Remark. Since the representation $p: B_n \rightarrow SD_n$ corresponds directly to the bracket expansion, $\langle \bowtie \rangle = A \langle \times \rangle + A^{-1} \langle \cup \rangle$, it is not surprising that the algebraic details work out. From our point of view, we had *already* verified that p would represent the braid group when we proved that $\langle K \rangle$ is a regular isotopy invariant.

On the other hand, this diagram algebra, SD_n , is a formal version of the von Neumann algebra used by Jones [13] to define the V -polynomial. To see the correspondence more directly, let $\tau = 1/\delta$ where $h_i^2 = \delta h_i$. Then let $e_i = \tau h_i$ so that the h_i relations become

$$\begin{cases} e_i^2 &= e_i, \\ e_i e_{i \pm 1} e_i &= \tau e_i, \\ e_i e_j &= e_j e_i, |i - j| > 1. \end{cases}$$

These are the relations for the von Neumann algebra A_n that gives rise to the V -polynomial. The original definition of the V -polynomial involves a (generalized) trace $\text{Tr}: A_n \rightarrow \mathbb{Z}[A, A^{-1}]$ (with A replaced by $t^{-1/4}$). In our terms, this trace appears as a diagrammatic count: $\text{Tr}(H) = \langle \overline{H} \rangle$ where H is a product of

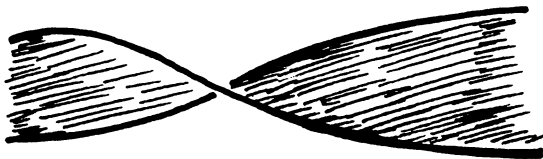
elements h_i (and $\text{Tr}(x + y) = \text{Tr}(x) + \text{Tr}(y)$ in SD_n). Thus the trace is equal to d —raised to one less than the number of circuits in the closure of the tangle corresponding to H .

We then have that (by linear extension) $\text{Tr}: SD_n \rightarrow \mathbf{Z}[A, A^{-1}]$ and that for any braid b , $\langle b \rangle = \text{Tr}(p(b))$. This gives a diagrammatic interpretation of the Jones trace, and shows how the relation with von Neumann algebras arises out of the ground of the knot theory.

Remark on the writhe. It is worth remarking here that the invariance of the writhe (Theorem 2.9) for reduced alternating links can also be seen using only the bracket. It follows from [20] that the highest and lowest degrees in f_K for K reduced and alternating are given by the formulas

$$\begin{aligned}\max \deg[K] &= -3w(K) + V + 2(W - 1), \\ \min \deg[K] &= -3w(K) - V - 2(B - 1),\end{aligned}$$

where $w(K)$ is the writhe of K , V is the number of crossings of K , W is the number of the white regions and B is the number of black regions in a checkerboard shading of K . All crossings are shaded as indicated below:



Thus we have

$$M(K) = \max \deg[K] + \min \deg[K] = -6w(K) + 2(W - B)$$

is an ambient isotopy invariant of K .

Murasugi observes [39] that the signature of a reduced alternating knot is given by the formula

$$\text{Sign}(K) = (1/2)(W - B) - (1/2)(w(K))$$

for this same shading. Putting these two formulas together he deduces that $w(K)$ and $(W - B)$ are individually invariants of ambient isotopy.

Conjecture. The full Tait conjecture [33, 43] states that ambient isotopic reduced alternating projections are related by a sequence of flypes (180 degree rotations of two-strand tangles). If the Tait conjecture is true then the *unrestricted bracket* $[K]$ of a reduced alternating knot would be an ambient isotopy invariant. The unrestricted bracket is the polynomial in three variables A, B, d computed by the bracket recursion (without any further relations on these variables). The unrestricted bracket is a flying invariant. I conjecture that it is an ambient isotopy invariant for reduced alternating diagrams.

IV. ALGEBRAS AND THE BRAID MONOID

The results of §3 extend to the context of an algebra associated to the entire braid monoid. The L and F polynomials then give rise to mappings defined

on this algebra (compare [4]). It is the purpose of this section to describe these mappings, and to give generators and relations for the corresponding algebras.

Let us begin by extending the context of the braid monoid BM_n . Recall from Definition 3.3 that the braid monoid on n strands, BM_n , is the set of regular isotopy classes of tangles generated by products of braid generators (σ_i) , hooks (h_i) and simple closed curves (δ) . The tangles are *diagrammatic tangles* in that they are given by planar diagrams with n input points and n output points. The tangles occur in the rectangular space between these rows of points, and regular isotopies are taken relative to the endpoints; these regular isotopies are confined within the rectangle.

Definition 4.1. Let \mathcal{T}_n denote all diagrammatic tangles with n inputs and n outputs (as described above). Each strand in an element $T \in \mathcal{T}_n$ proceeds from one endpoint (input or output) to another, crossing over and under other strands in the process. All tangles are taken up to regular isotopy (relative to the endpoints).

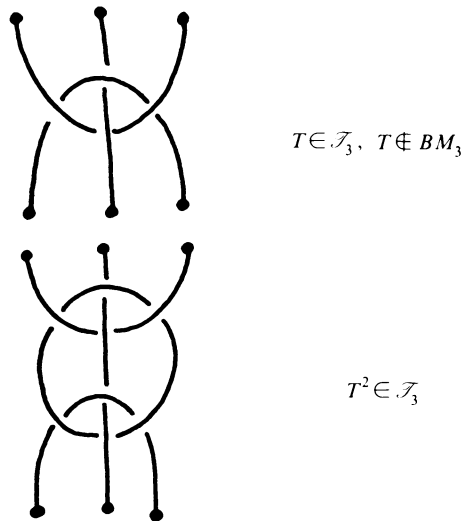


FIGURE 17

As with the braid monoid, I shall also allow the possibility that a tangle includes arbitrary many loop-components. With this caveat, *the regular isotopy classes of tangles \mathcal{T}_n becomes a monoid under standard tangle multiplication.* We have the proper inclusions $\mathcal{T}_n \supset BM_n \supset B_n$ where BM_n is the braid monoid generated by the braid group elements together with the hooks described in §3. B_n denotes the n -strand Artin braid group. That these inclusions are proper is illustrated by the example in Figure 17. This example shows that the inclusion would be proper even if we eliminated the possibility of knotted tangle strands. Figure 17 shows an element $T \in \mathcal{T}_3$, that is not in the braid monoid.

One way to see that T is not regularly isotopic to a braid monoid element is to note that in T^2 a loop appears that is linked with one of the strands. This example also underscores the reason for allowing loops in the tangles.

Just as we can define the value of the L polynomial on a braid b via the definition $L_b = L_{\bar{b}}$ where \bar{b} is the closure of b we can define

$$L: \mathcal{T}_n \rightarrow \mathcal{R} = \mathbf{Z}[a, a^{-1}, z, z^{-1}]$$

by $L_T = L_{\bar{T}}$, where \bar{T} is the standard closure of the tangle, obtained by joining top strands to bottom strands through a trivial n -strand tangle. This mapping respects monoid relations; thus (for example) $Lwh_i h_{i+1} h_i \tau = Lwh_i \tau$ where $w, \tau \in \mathcal{T}_n$ and h_i and h_{i+1} are the standard hook tangles from §3. The additive polynomial relation

$$L \begin{array}{c} \nearrow \searrow \\ \nearrow \searrow \end{array} + L \begin{array}{c} \searrow \nearrow \\ \searrow \nearrow \end{array} = z \left(L \begin{array}{c} \nearrow \nearrow \\ \searrow \searrow \end{array} + L \begin{array}{c} \searrow \searrow \\ \nearrow \nearrow \end{array} \right)$$

can be regarded in this context as the relation

$$Lw\sigma_i\tau + Lw\sigma_i^{-1}\tau = z(Lwh_i\tau + Lw\tau)$$

for elements of \mathcal{T}_n .

Formally, one would like to rewrite an equation like the one above as

$$L(w\sigma_i\tau + w\sigma_i^{-1}\tau - z(wh_i\tau + w\tau)) = 0.$$

In order to do this, we need a domain where tangle elements can be combined additively. This is provided by the definition below.

Definition 4.2. Let \mathcal{T}_n denote the tangle monoid and let $\mathcal{A}\mathcal{T}_n$ denote the free algebra over \mathcal{T}_n with coefficients in $\mathcal{R} = \mathbf{Z}[a, a^{-1}, z, z^{-1}]$, factored by the following relations:

(i) If d denotes the unknotted, unlinked loop element of \mathcal{T}_n , then d is set equal to $z^{-1}(a + a^{-1}) - 1 \in \mathcal{R}$.

$$(ii) \quad \begin{cases} \sigma_i + \sigma_i^{-1} = z(h_i + 1), \\ \sigma_i h_i = h_i \sigma_i = a h_i, \\ \sigma_i^{-1} h_i = h_i \sigma_i^{-1} = a^{-1} h_i, \\ h_i \sigma_{i \pm 1} h_i = a^{-1} h_i, \\ h_i \sigma_{i \pm 1}^{-1} h_i = a^{+1} h_i, \end{cases}$$

where σ_i, h_i are braid and hook generators in \mathcal{T}_n .

(Compare Figure 18.) Let \mathcal{AM}_n denote the restriction of \mathcal{AT}_n to tangles in the braid monoid \mathcal{BM}_n . Call \mathcal{AM}_n the *braid monoid algebra*. Because the relations in Definition 4.2 correspond to identities for the polynomial L , we have a well-defined mapping $L: \mathcal{AT}_n \rightarrow \mathbf{Z}[a, a^{-1}, z, z^{-1}]$ and its restriction

$$L: \mathcal{AM}_n \rightarrow \mathbf{Z}[a, a^{-1}, z, z^{-1}].$$

$$\begin{array}{ccc}
 \begin{array}{c} || \cdots | \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} | \cdots || \end{array} & \sigma_i h_i \Rightarrow L w \sigma_i h_i \tau = a L w h_i \tau & \\
 \\
 \begin{array}{c} || \cdots | \begin{array}{c} \cup \\ \cap \end{array} | \cdots || \end{array} & h_i & \\
 \\
 \begin{array}{c} || \cdots | \begin{array}{c} \cup \\ \cap \\ \cup \\ \cap \end{array} | \cdots || \end{array} & h_i \sigma_{i+1} h_i \Rightarrow L h_i \sigma_{i+1} h_i = a^{-1} L h_i &
 \end{array}$$

FIGURE 18

The map L is defined on individual tangle elements by evaluating L on the closure of the tangle. It is then extended additively to the entire algebra $\mathcal{A}\mathcal{T}_n(\mathcal{A}\mathcal{M}_n)$.

In order to have a purely algebraic description of this mapping, one would like a purely algebraic version (generators and relations) for the tangle algebra $\mathcal{A}\mathcal{T}_n$. This is an open problem. However, the rest of this section will give the relations for the braid monoid algebra $\mathcal{A}\mathcal{M}_n$. We shall see that it is identical to the algebra defined by Birman and Wenzel [4]. Thus our result gives a diagram/geometric context for their algebra.

In order to explicate the structure of $\mathcal{A}\mathcal{M}_n$, we first concentrate on relations for the braid monoid $\mathcal{B}\mathcal{M}_n$, and within it the diagram monoid D_n .

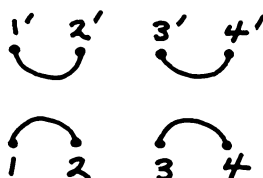
Recall that the diagram monoid is an algebra of diagrams taken up to topological equivalence in the plane (relative to an upper and lower row of n points). Each diagram is represented as a product of elementary diagrams that we have called "hooks" in §3. Each diagram appears as a collection of nonintersecting strands proceeding from points of a given row and ending either on that row, or on the other row. No two strands intersect or cross one another in the diagram monoid. In order for a given diagram of this type to be seen as an element of the diagram monoid it must be configured with paired maxima and minima that display it as a product of hooks.

The closed loop in the algebra is given to commute with all elements. Consequently, we shall assume that any given diagram is loop-free unless otherwise indicated.

There is another more simply described monoid associated to the diagram monoid. I call this the *connection monoid*, C_n . Here we again consider dia-

grams consisting of *nonintersecting strands* emanating from, and returning to the top and bottom rows, always drawn between these rows. No pairing of maxima and minima is required here. Thus an element in the connection monoid may, by pairing maxima and minima in different ways, correspond to more than one element of D_n .

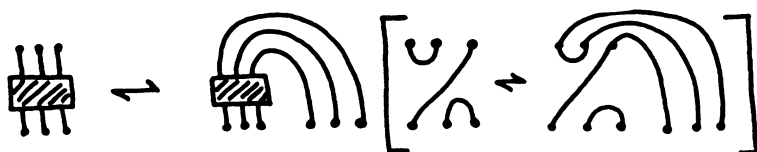
Two elements of the connection monoid C_n are said to be equivalent if there is a topological deformation of the strands relative to the upper and lower rows that carries one diagram into the other. (Just as in the diagram monoid, the strands are to remain disjoint throughout the deformation.) For the connection monoid, this implies that *two expressions are equivalent if and only if they have the same configuration of row connections*. For example, in the diagram below, the data $[1'2'] [3'4'] [12] [34]$ is sufficient to determine it as an element of C_4 . Here $[ab]$ means that the endpoints a and b are connected by a strand.



Remark. Let P_{2n} denote the set of *parenthesis structures* on a row of $2n$ points. By the term parenthesis structure I mean a partitioning of the set of points into pairs so that these pairings can be realized by nonintersecting strands in the plane above the line of products. Equivalently, if a left parenthesis $[(]$ is on the left point of a pair and a right parenthesis $])$ is placed on the right member of each pair, then the resulting nested parenthesis structure is a legal parenthesizing (in the usual sense of correct typography) involving $2n$ parentheses. For example, the elements of P_6 are shown below.



Fact. The elements of C_n are in one-to-one correspondence with the elements of P_{2n} . The proof is indicated in the diagram:



It is easy to verify that P_{2n} has cardinality the Catalan number

$$\frac{1}{n+1} \binom{2n}{n} = \frac{1}{(n+1)} \frac{(2n)!}{(n!)^2}.$$

Consequently the number of loop-free connections in C_n has the same value. To return to the diagram monoid, let \hat{D}_n denote the loop-free elements of

D_n , and \widehat{C}_n denote the loop-free elements of C_n . Define a map $e: D_n \rightarrow \widehat{C}_n$ by forgetting the paired maxima/minima structure of elements of \widehat{D}_n .

Theorem 4.3. *The mapping e defined above is a one-to-one correspondence. In fact, by using this correspondence it follows that two elements h and h' of the full diagram monoid D_n are equivalent in D_n if and only if they are related (as words in $\{h_i\}$) by a finite sequence of applications of the relations*

$$\begin{aligned} h_i^2 &= dh_i = h_id, \\ h_i h_{i\pm 1} h_i &= h_i, \\ h_i h_j &= h_j h_i, \quad \text{if } |i - j| > 1. \end{aligned}$$

Proof. Any product of the generators h_1, h_2, \dots, h_{n-1} can be put into a normal form by using the relations (*). In the normal form the product takes the form

$$h(1)h(2)\cdots h(m) = h$$

where h is a word without any consecutive appearances of a given hook (since $h_i h_i = dh_i$ and the d commutes with everybody). Furthermore, h is a product of segments of the form $h_{\alpha_1} h_{\alpha_2} h_{\alpha_3} \cdots h_{\alpha_r} = h(k)$ where the indices $\alpha_1, \dots, \alpha_r$ form a descending sequence of integers. Furthermore, the list of left-most indices of adjacent sequences is ascending as one scans from left to right. See Figure 16 for an example. (See [15] for a proof of this statement.) By following the pattern indicated in Figure 16, we associate to each element c of C_n an element h of \widehat{D}_n that is in normal form (and $e(h) = c$). Conversely, any normal form occurs in this way.

Thus we have shown that if h and h' are loop-free words in the generators of the diagram algebra with $e(h) = e(h')$, then h and h' have identical normal forms. Hence h can be obtained from h' by repeated application of the relations. Since any element can, via the relations, be written as a power of the loop times a loop-free word, this completes the proof of the theorem.

Now we turn to the structure of the full braid monoid. At this point it is worth remembering that regular isotopy is generated by Reidemeister moves of types II and III (see Figure 1). In the theory of braids the moves of type II and type III are of special importance because with braid strands being unknotted and always proceeding downward from top row to bottom row, the type I move is never needed. In other words, braid equivalence in the braid group B_n is regular isotopy relative to the top and bottom rows for that braid. A local type I move is simply not in the braid form. We should remark, however, that a global (on the 2-sphere) type I move can change a closed braid from B_n to B_{n+1} or to B_{n-1} . This is called a *Markov move*. In [3] one finds a proof of the Markov theorem stating that two knots and links represented as braids are ambient isotopic if and only if the corresponding closed braids can be transformed one to another by a combination of braid equivalence, conjugation in the braid group, and Markov moves. Thus in our terminology the ambient isotopy factors into combinations

of conjugations, special regular isotopies (the braid moves) and certain type I moves on the two-sphere (the Markov moves).

The algebraic relations in the braid group correspond directly to the type II and type III moves. Type II moves are of the form

$$\sigma_i \sigma_i^{-1} = 1 = \sigma_i^{-1} \sigma_i$$

while type III moves correspond to the braiding relations

$$\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1}.$$

The commutativity relations

$$\sigma_i \sigma_j = \sigma_j \sigma_i \quad \text{for } |i - j| > 1$$

are topological moves that do not change the configuration of crossings. They are isotopies of the underlying graphical structure of the braid projection.

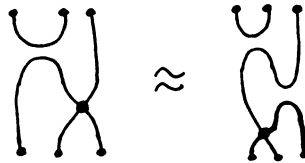
The braid group is generated by

$$\sigma_1, \sigma_1^{-1}, \sigma_2, \sigma_2^{-1}, \dots, \sigma_{n-1}, \sigma_{n-1}^{-1} \quad (B_n = n\text{-strand braid group})$$

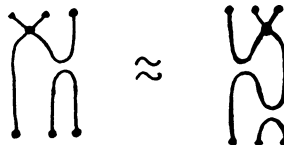
with exactly the relations described above.

In the braid monoid almost the same situation ensues. Now, however, along with the “vertical” type II move corresponding to the product of a braid generator and its inverse there is also the possibility of a “slantwise” type II move at a hook. Here one of a pair of maxima/minima separate as a single type II move occurs (see Figure 19). But, as Figure 19 shows, the reconfiguration move along with the vertical type II is sufficient to generate the slantwise type II move.

A word about the reconfiguration move ($\sigma_1 h_2 = \sigma_2^{-1} h_1 h_2$ and variants): See Figure 19. This move overlies a topological equivalence of the graph (planar graph) that is obtained by replacing the crossing at a braid generator by a 4-valent vertex. Thus, we have an underlying *graphical reconfiguration* as shown below.

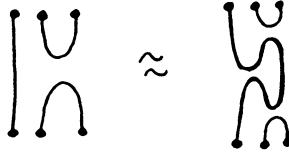


This can be symbolized as $h_i c_{i+1} = h_i h_{i+1} c_i$. There are obvious variants such as



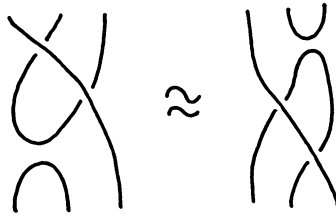
$c_i h_{i+1} = c_{i+1} h_i h_{i+1}$. Here c_i stands for the graphical vertex replacement of the braid generator.

Along with the graphical reconfiguration there is also the standard hook relation such as

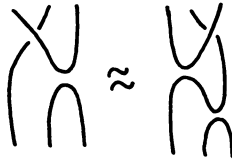


$$(h_{i+1} = h_{i+1} h_i h_{i+1}).$$

The point about the reconfiguration move and the hook relation is that these are both topological equivalences of the diagram. They are not Reidemeister moves, and they do not change the underlying graphical structure of the diagram. One further graphical move is commutativity of elements that are far apart ($c_i c_j = c_j c_i$ if $|i - j| > 1$, $h_i h_j = h_j h_i$ if $|i - j| > 1$).



slantwise type II move
($\sigma_1 \sigma_2 h_1 = h_2 \sigma_1 \sigma_2$)



reconfiguration move
($\sigma_1 h_2 = \sigma_2^{-1} h_1 h_2$)

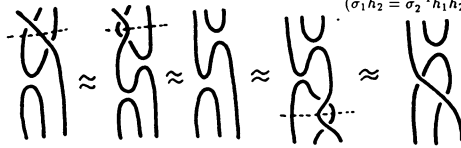


FIGURE 19. Accomplishing slantwise type II by reconfiguration and vertical ($\sigma_i \sigma_i^{-1} = 1$) type II. $\sigma_1(\sigma_2 h_1) = \sigma_1(\sigma_1^{-1} h_2 h_1) = h_2 h_1 = (h_2 h_1 \sigma_2^{-1}) \sigma_2 = h_2 \sigma_1 \sigma_2$

Finally, we come to the type III move in the braid monoid. Figure 20 shows that the type III move can be accomplished using standard braiding ($\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1}$) and reconfiguration.

Theorem 4.4. *Two diagrams in the braid monoid MB_n are regularly isotopic if and only if one can be obtained from the other by a combination of standard braiding relations, hook relations and reconfigurations.*

Proof. Any regular isotopy of diagrams can (by definition) be factored into a sequence of type II moves and type III moves. Up to a topological deformation

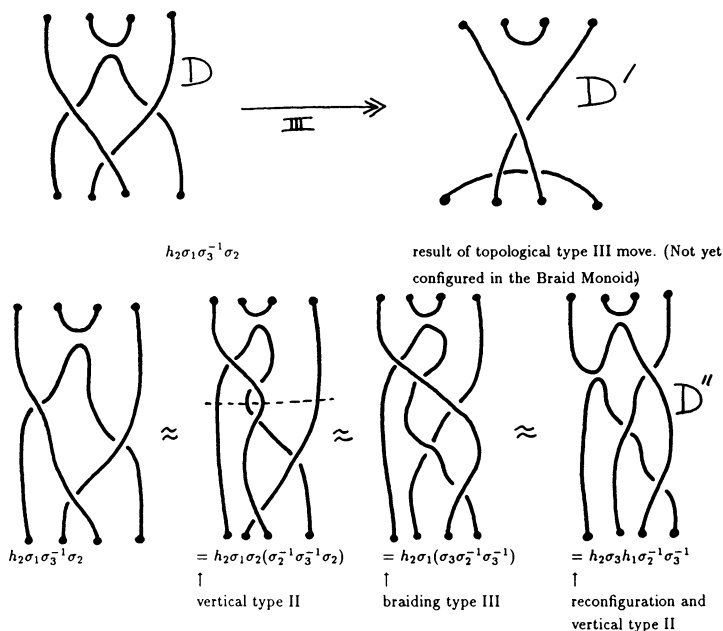


FIGURE 20. D'' is a braid monoid element regularly isotopic to D' . Accomplishing a type III move via braiding and reconfiguration.

of the diagram, any type II move is either “vertical” and hence of the form $\sigma_i\sigma_i^{-1} = 1$ or “slantwise” as shown in Figure 19. Figure 19 shows that the slantwise type II can be accomplished via braiding and reconfiguration. Similarly, Figure 20 shows that a nonbraiding type III move can be accomplished by braiding and reconfiguration. Thus if d and d' are diagrams in BM_n with $d \approx d'$, then each type II or type III move in this regular isotopy can be replaced by a move or sequence of moves, resulting in an algebraic equivalence (i.e., by braiding, hooks and reconfigurations) $d \equiv d''$ where d'' is also in BM_n and d'' and d' are topologically equivalent diagrams.

The next stage in the process described by this proof is to show that d'' and d' are equivalent via commuting relations, hook relations and reconfigurations. In order to do this I indicate two more stages in the procedure: Note that d' and d'' have the same number of crossings, and that in d'' these crossings occur as appearances of σ_i or σ_i^{-1} for various choices of i in the word for d'' . I now apply two lemmas, stated and proved below. By Lemma A, a series of reconfiguration moves will transform d'' into d''' where d''' has, up to commutation, the same pattern of σ_i 's as d' . They, by Lemma B, d''' is transformed to d' by commutations and hook relations.

Lemma A. Let d_1 and d_2 be two elements of the braid monoid BM_n , and suppose that d_1 and d_2 are graphically equivalent diagrams. Let $\Sigma(d)$ denote the word in $\{\sigma_1, \dots, \sigma_{n-1}, \sigma_1^{-1}, \dots, \sigma_{n-1}^{-1}\}$ obtained by setting each h_i in d

equal to 1. Then there exists a series of reconfiguration moves taking d_1 to d'_1 so that $\Sigma(d'_1)$ and $\Sigma(d_2)$ differ only by commutation relations ($\sigma_i \sigma_j = \sigma_j \sigma_i$ for $|i - j| > 1$).

Proof. See Figure 21 for an example of this lemma. The proof goes by induction on the length of the word $\Sigma(d)$. Choose the right-most occurrence of σ_i or σ_i^{-1} in $\Sigma(d)$. Either this index must be changed or it will be left alone in the transition to d'_1 . In order to decide this, compare the tangle-graphs underlying d_1 and d_2 . Since we are given that the graphs are equivalent, it is possible to label the crossings in one graph that correspond to crossings in the other graph. Thus σ_i in d_1 corresponds via the graph-isomorphism to σ_j in d_2 (their exponents may differ if $i \neq j$). If $i \neq j$ then we desire a sequence of reconfiguration moves that will shift this index. That these moves exist follows from the equivalence of the graphs: Note the two strands that cross at σ_i . Use reconfigurations with adjacent maxima and minima to raise or lower the index i . These can be accomplished because, via the planar isotopy between d_1 and d_2 , there is an isotopy moving one strand along the other that carries the crossing to its desired position. This completes the proof of Lemma A.

Lemma B. Let d_1 and d_2 be two elements of the braid monoid BM_n , and suppose that d_1 and d_2 are graphically equivalent diagrams. Let $\Sigma(d_1)$ and $\Sigma(d_2)$ be as in Lemma A. Assume that $\Sigma(d_1)$ and $\Sigma(d_2)$ are identical words in the braid generators. Then there exists a sequence of hook relations carrying d_1 to d_2 .

Proof. Apply Theorem 4.3 to each full word in the h_i 's that is a subword of d_1 . The equivalence of the graphs establishes a correspondence between the set of h_i -subwords of d_1 and the h_i -subwords of d_2 . The hypotheses of 4.3 apply, and show that each pair of words are equivalent in the diagram monoid. This completes the proof of Lemma B.

The proof of Theorem 4.4 is now complete.

Remark. Another good example for Theorem 4.4 is the presence of the regular isotopy version of the Whitney trick in the braid monoid. This is shown in Figure 22.

Remark. The connection monoid C_n naturally embeds in the Brauer monoid BR_n of all possible connections between two rows of n points. This monoid is denoted BR after R. Brauer [7] who studied it, and an algebra BR_n as coefficients in the context of representations of the orthogonal group. We have an obvious map from the braid monoid to the Brauer monoid that is obtained by taking only the connection structure of a given element of the braid monoid. The Brauer monoid has a free commuting loop element, and it follows from our discussion that the braid monoid modulo an appropriate equivalence relation will give the Brauer monoid. The equivalence relation includes equivalence of braid generators and their inverses, plus curl-eliminating relations such as

$$ch = hc = h, \quad hc'h = h,$$

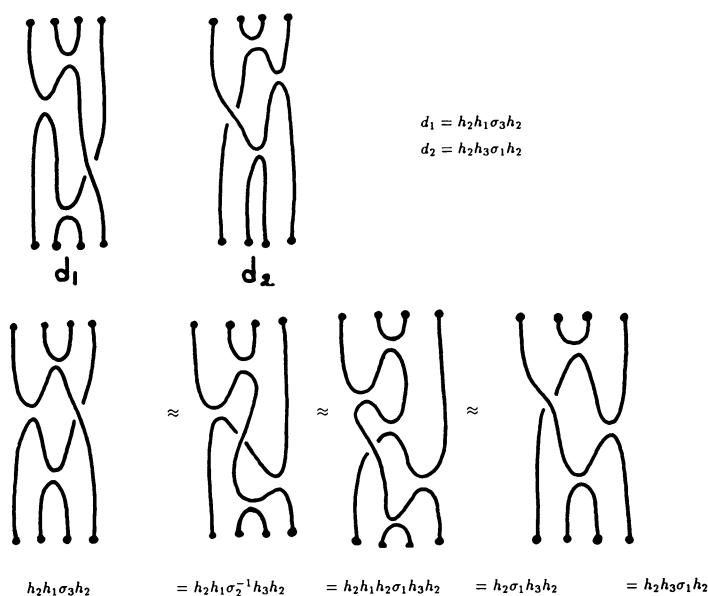


FIGURE 21

where c is a braid generator in the same column as h , and c' is shifted by a column from h . A corresponding result shows that the Brauer algebra is isomorphic to a specialization of the braid monoid algebra AM_n . Curiously, the number of multiplicative generators of AM_n and of the Brauer monoid is the same. This follows from the exchange identity

$$c + c^{-1} = z(1 + h).$$

This does not set braid generators and their inverses equal to one another, but it does reexpress inverses in terms of the rest of the algebra.

The upshot of these considerations is that for computing L_K using the algebra AM_n , it suffices to know explicitly a set of polynomial values in one-to-one correspondence with the elements of the Brauer monoid. For BR_n there are $(2n-1)(2n-3)\cdots(3)(1)$ such elements. Thus for three-strand braids it suffices to know the value of the L -polynomial for 15 particular braids, in order to get any other by pure algebra! Clearly a computer program is called for here.

V. REGULAR ISOTOPY

We have been using the concept of regular isotopy as a formal trick: restricting knot moves to the Reidemeister types II and III. In this section I will discuss geometry behind this equivalence relation, and its relevance for the knot polynomials.

Recall the notion of *regular homotopy* [47]. Two immersions of the circle into the plane are said to be regularly homotopic if there is a time-parameter (t) family of immersions of the circle that restricts to one map at $t = 0$ and

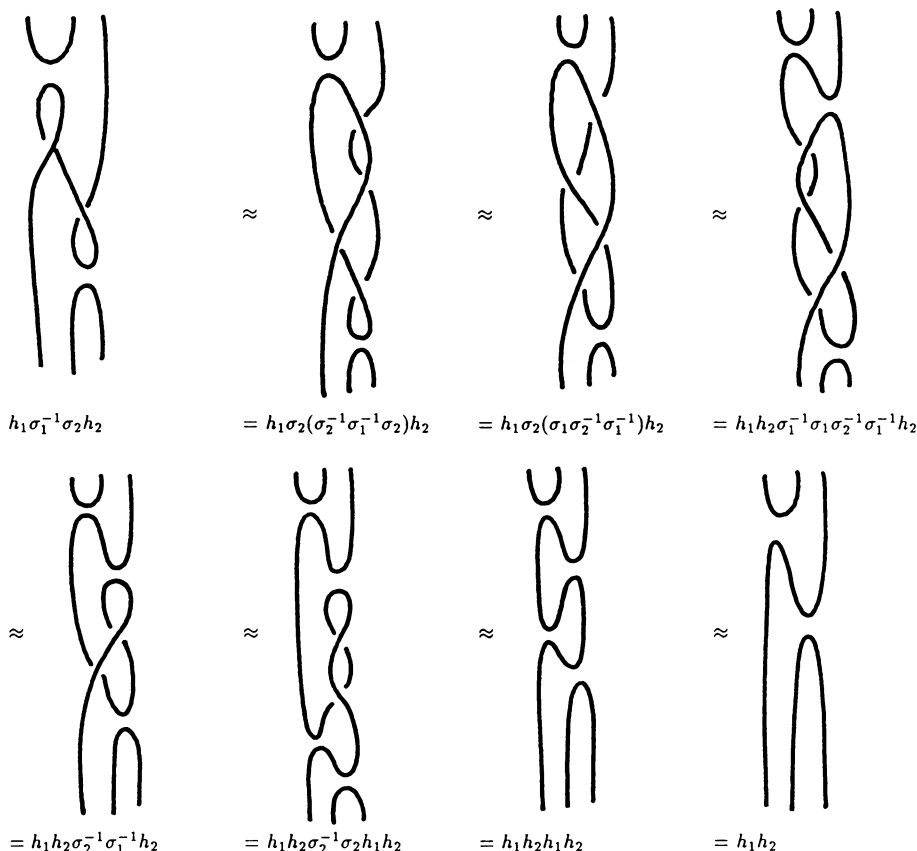


FIGURE 22. Whitney trick in the braid monoid

to the other at $t = 1$ (for t varying in the interval from 0 to 1). It is required that the family be differentiable in the variable t .

This means that the shadows of the Reidemeister II and III moves can be seen as regular homotopies, but that the shadow of the type I move cannot be a regular homotopy, since the contraction of the loop would violate differentiability. For planar curves, one can discretize the notion of regular homotopy by taking the shadows of the Reidemeister type II and type III moves as generators for the equivalence relation.

Therefore any regular isotopy of link diagrams projects to a regular homotopy of the underlying plane curves. This is the genesis of the choice of terminology. It also leads to another invariant of regular isotopy: the *Whitney degree of the underlying plane curve(s)*.

We define the Whitney degree, $d(U)$, for any oriented curve or curves U in the plane with ordinary double point intersections (I call such a collection of curves a *link-shadow*, or a *universe*) by splitting each crossing in an oriented fashion, producing a collection C (the Seifert circuits) of disjoint oriented curves in the plane. Let d assign plus 1 or minus 1 to each curve according to its

counter-clockwise or clockwise orientation, respectively. Then let $d(U) = d(C)$ denote the sum of these plus or minus ones. The resulting integer, $d(U)$, is an invariant of regular homotopy for U . See [18] for a combinatorial derivation of this fact.

The integer $d(U)$ is identical in the case of differentiable plane curves, with the degree defined by Whitney and Graustein. They define degree to be the total turning of a unit tangent vector to the curve as the curve is traversed once (and the sum of these total turns for each component of a multiple curve). The fundamental result is that *two oriented curves are regularly homotopic if and only if they have the same degree*.

Thus we may define for a given oriented link diagram L , the invariant $d(L) = d(U)$ where U is the planar shadow of L . Then $d(L)$ is an invariant of regular isotopy. It will be referred to as *the Whitney degree of the oriented diagram L* .

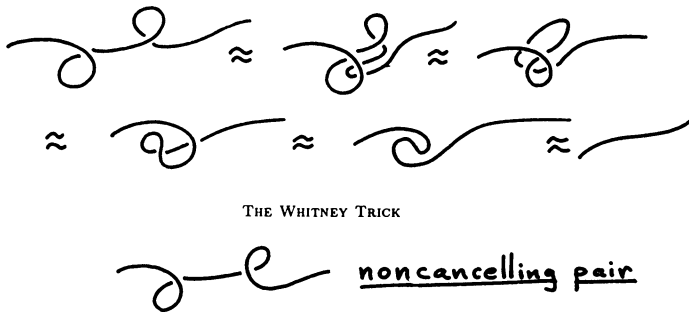


FIGURE 23

Example. View Figure 23 for a regular isotopy version of the Whitney trick. This figure also shows a situation of two curls on a string that cannot cancel up to regular isotopy, due to the nonvanishing of the writhe. It is sometimes useful to consider knots on a string (isotopy relative to the endpoints) rather than with closed loops. Then we assign Whitney degree zero to the unknotted string, and calculate it for other diagrams by the Seifert circuit description given above. The decomposition into Seifert circuits yields a collection of closed curves and one unknotted string. The following lemma is a direct generalization of the Whitney-Graustein theorem.

Lemma 5.1. *Let S_1 and S_2 be unknotted strings. Then S_1 is regularly isotopic to S_2 (relative to end-points) if and only if S_1 and S_2 have the same Whitney degree and the same writhe.*

For a proof of this lemma see [19 or 46].

Example. The figure eight knot as shown in Figure 24 has Whitney degree -1 . A standard way to turn the figure eight into its mirror image reverses the orientation on the knot. Since this orientation reversal changes the Whitney degree

to $+1$, we conclude that the figure eight knot is not regularly isotopic to its orientation-reversed mirror image. The figure eight knot is regularly isotopic to its orientation-preserved mirror image, as shown in Figure 24. In fact,

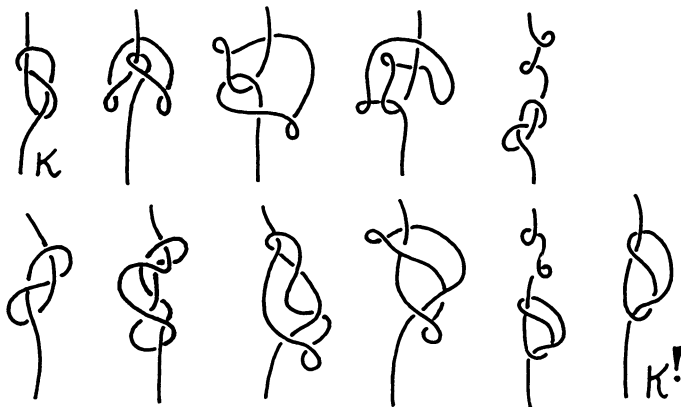


FIGURE 24. Regular isotopy of figure eight knot to its mirror image

Theorem 5.2. *If an oriented knot K of writhe zero ($w(K) = 0$) is oriented equivalent (ambient isotopic) to its mirror image, then this equivalence can be accomplished through regular isotopy.*

Proof. K is ambient isotopic to its mirror image $K!$ if and only if

$$K + S_1 \approx K! + S_2$$

where $+$ denotes connected sum along the string, and S_1 and S_2 are unknotted strings that catalogue the type I moves needed or performed during the isotopy. Since writhe and Whitney degree are additive under connected sum, we have

$$\begin{aligned} w(K) + w(S_1) &= w(K!) + w(S_2), \text{ hence} \\ w(S_1) &= w(S_2) \quad (K \text{ has zero writhe}). \end{aligned}$$

Similarly, since $d(K) = d(K!)$, we have $d(S_1) = d(S_2)$. Therefore, by Lemma 5.1, S_1 and S_2 are regularly isotopic. Thus, letting S denote S_1 , $K + S \approx K! + S$. Since S is a string of curls, we may choose $S^\#$, a string of opposite curls in the sense of the Whitney trick so that

$$S + S^\# \approx \sim.$$

Then, adding $S^\#$ to both sides,


$$K + S + S^\# \approx K! + S + S^\#,$$

hence $K \approx K!$. This completes the proof.

Framed links and twisted bands. Another context for regular homotopy arises by interpreting our links as *framed links*. That is, we assume that each link

component is endowed with a unit normal vector field. This is equivalent to replacing the embedding of each component by an embedded band whose core is the given component, and such that bands for different components do not intersect. As Figure 25 illustrates, the Reidemeister moves of type II and type III on the cores of bands extend to the bands themselves, while the type I move does not extend. (Rather it corresponds to a full twist on the band.) Consequently, *regular isotopy corresponds to ambient isotopy of framed links or to ambient isotopy of embedded bands*. Note, however, that regular isotopy is actually more restricted than the corresponding band-isotopy. Thus



as bands, relative to the endpoints, while there is no cancelling regular isotopy for  relative to its endpoints. However, by the rules for the L -polynomial, $L \text{ (crossing) } = aa^{-1} L \text{ (parallel) } = L \text{ (parallel) }$ (since the Whitney degree does not play a role in this evaluation). As a result we can interpret the L -polynomial as an ambient isotopy invariant of links of bands embedded in three dimensional space. The axioms are given below.

Axioms for L as a Band-Invariant

$$1. \quad L \text{ (crossing) } + L \text{ (crossing) } = z(L \text{ (parallel) } + L \text{ (parallel) }),$$

$$2. \quad L \text{ (twist) } = L \text{ (parallel) } = aL \text{ (parallel) }, \quad L \text{ (full twist) } = 1,$$

$$L \text{ (full twist) } = L \text{ (parallel) } = a^{-1}L \text{ (parallel) }.$$

Note that in this interpretation, the second variable (a) measures the twisting of the bands. This measurement is distributed throughout the skein calculation.

The same remarks apply to the homfly polynomial. We can start by defining a polynomial invariant of regular isotopy, call it H_K , satisfying the axioms:

Axioms for H

$$1. \quad H \text{ (crossing) } - H \text{ (crossing) } = zH \text{ (parallel) },$$

$$2. \quad H \text{ (twist) } = aH \text{ (parallel) }, \quad H \text{ (full twist) } = 1, \quad H \text{ (full twist) } = a^{-1}H \text{ (parallel) }.$$

Then define $P_K = a^{-w(K)} H_K$ to obtain an invariant of ambient isotopy. The polynomial P_K is easily seen to be equivalent to the homfly polynomial. In fact, P_K satisfies the identity

$$a^{-1}P \begin{array}{c} \nearrow \\ \searrow \end{array} - aP \begin{array}{c} \nearrow \\ \searrow \end{array} = zP \begin{array}{c} \longrightarrow \end{array}.$$

This point of view also shows how the homfly polynomial is related to the Hecke algebra obtained by adjoining the relations $\sigma_i - \sigma_i^{-1} = z$ to the free additive algebra over the braid group. (See [13, 41].) The braiding relation $\sigma_i - \sigma_i^{-1} = z$ is the direct analog of the polynomial relation $H \begin{array}{c} \nearrow \\ \searrow \end{array} - H \begin{array}{c} \searrow \\ \nearrow \end{array} = zH \begin{array}{c} \longrightarrow \end{array}.$

To summarize, we have indicated two motivations for regular isotopy, one as a lifting of regular homotopy of plane curves, the other as a version of ambient isotopy of framed links. Because of the multiplicity of interpretations, it has been convenient to isolate regular isotopy as a formal equivalence of link diagrams.

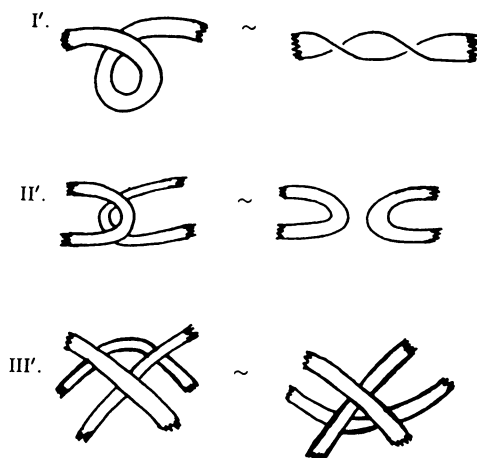


FIGURE 25. Ambient isotopy of twisted bands

VI. WELL-DEFINEDNESS AND INVARIANCE

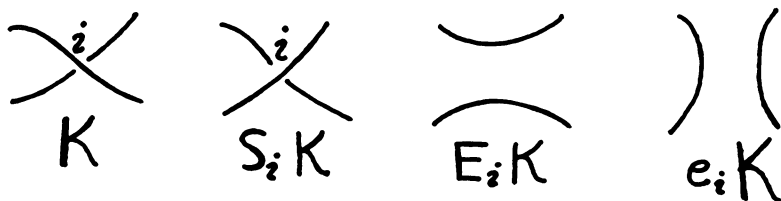
This section defines the polynomial L inductively, and proves that it is an invariant of regular isotopy.

For these purposes it is useful to establish notation for the switching and elimination of crossings. Thus I shall write $S_i K$ for the link obtained by switching K at the i th crossing, and $E_i K$, $e_i K$ respectively for the two modes of splicing the i th crossing. See Figure 26.

We wish to give a definition of L_K so that the identity

$$L_K + L_{S_i K} = z(L_{E_i K} + L_{e_i K})$$

is a consequence of this definition. The motivation for the definition we will adopt is given by the following remarks. The definition (Definition 6.4) will follow these remarks.

FIGURE 26. $L_K + L_{S_i K} = z(L_{E_i K} + L_{e_i K})$

Motivating the inductive definition. For a set of crossings labelled $0, 1, 2, 3, \dots, n$ in the diagram K consider the result, $\hat{K} = S_n S_{n-1} \cdots S_1 S_0 K$, of switching all these crossings. We then have the following list of equations (assuming the exchange identity for L):

$$\begin{aligned} L_K + L_{S_0 K} &= z(L_{E_0 K} + L_{e_0 K}) \\ L_{S_0 K} + L_{S_1 S_0 K} &= z(L_{E_1 S_0 K} + L_{e_1 S_0 K}) \\ &\vdots \\ L_{S_{n-1} \cdots S_0 K} + L_{\hat{K}} &= z(L_{E_n S_{n-1} \cdots S_0 K} + L_{e_n S_{n-1} \cdots S_0 K}) \end{aligned}$$

Definition 6.2. Define abbreviated switching and elimination operators A_i and B_i by the following formulas:

$$A_i K = E_i S_{i-1} \cdots S_0 K, \quad B_i K = e_i S_{i-1} \cdots S_0 K.$$

Successive addition and subtraction of the above equations, then shows that

$$(*) \quad L_K = (-1)^{n+1} L_{\hat{K}} + z \sum_{i=0}^n (-1)^i [L_{A_i K} + L_{B_i K}].$$

Thus, formula $(*)$ shows how to compute L_K in terms of $L_{\hat{K}}$ and the results of L applied to smaller (fewer crossings) knots and links.

Since we can choose a switching sequence so that \hat{K} is unknotted when K is a knot, or \hat{K} is a split link when K is a link, this formula gives an inductive procedure for calculating the polynomial L . Note that in calculating with links we also have the rule

$$(**) \quad L_{K_1 \sqcup K_2} = d L_{K_1} L_{K_2}$$

whenever the diagram of K_1 *overlies* the diagram of K_2 . (I say that K_1 *overlies* K_2 if every crossing of K_1 and K_2 has the overcrossing segment in K_1 .) Here $d = (a + a^{-1})(z^{-1}) - 1$ as in §2. Note that this includes the case where the diagrams are disjoint in the plane.

The key to creating an inductive definition is the use of the *standard unknot* [2] associated with a given knot-diagram with directed base-point. Let K be any knot diagram and U its planar shadow. Let p be an interior point of an arc of U , together with a direction of travel along the projection U . Call p a *directed basepoint* for U .

The *standard unknot* is constructed as follows: Given U and a directed basepoint p , draw a knot diagram $\hat{K} = \hat{K}(U, p)$ by tracing along U in the direction from p and *over-crossing on the first pass through each crossing*. This produces an unknotted diagram as shown in Figure 27. I call $\hat{K} = \hat{K}(U, p)$ the standard unknot associated with U and the directed basepoint p .

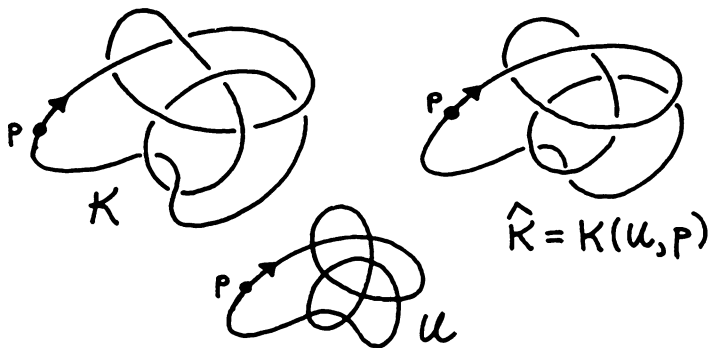


FIGURE 27

The standard unknot $\hat{K} = \hat{K}(U, p)$ can be used to produce a *specific unknotting sequence for the knot K* : Traverse K from the basepoint p , and mark each crossing that differs from the corresponding crossing in \hat{K} . Label the marked crossings in descending order from the basepoint, say by the integers $n, n-1, \dots, 1, 0$. Then \hat{K} is obtained from K by switching these crossings, and we have the formula: $\hat{K} = S_n S_{n-1} \cdots S_1 S_0 K$.

This switching sequence is determined by the choice of directed basepoint on K .

The value of L on standard unknots is given by the formula

$$(***) \quad L_{\hat{K}(U, p)} = a^{w(\hat{K}(U, p))}$$

where w denotes the writhe as defined in §2.

In order to make use of formula $(**)$ ($L_{K_1 \sqcup K_2} = d L_{K_1} L_{K_2}$) it will also be necessary to separate components by a switching sequence. Formula $(*)$ then applies with \hat{K} a split link rather than an unknot.

We are now in possession of a method of recursive calculation using $(*)$, $(**)$ and $(***)$ so that the calculations will eventually depend only on the values for L at standard unknots $\hat{K} = \hat{K}(U, p)$ as described above.

In order to formalize these procedures to give an inductive definition, it is useful to adopt a notation for the right-hand side of formula $(*)$.

Definition 6.3. Let K be any link and let $\lambda = (\lambda_n, \lambda_{n-1}, \dots, \lambda_1, \lambda_0)$ be any ordered sequence of labels for crossings of K . Let A_i^λ and B_i^λ denote the operators given by the formulas

$$A_i^\lambda = E_i S_{\lambda_i} S_{\lambda_{i-1}} \cdots S_{\lambda_1} S_{\lambda_0}, \quad B_i^\lambda = e_i S_{\lambda_i} S_{\lambda_{i-1}} \cdots S_{\lambda_1} S_{\lambda_0}.$$

Let $\widehat{K}(\lambda) = S_{\lambda_n} S_{\lambda_{n-1}} \cdots S_{\lambda_0} K$. Let

$$\sum_K(\lambda) = \sum_{i=0}^n (-1)^i (L_{A_i^* K} + L_{B_i^* K})$$

and

$$\Omega_K(\lambda) = (-1)^{|\lambda|+1} L_{\widehat{K}(\lambda)} + z \sum_K(\lambda),$$

where $|\lambda| = n$ for λ as shown above. Note that we intend that $\Omega_K(\lambda)$ be equal to L_K . In the inductive definitions, a modification of $\Omega_K(\lambda)$ will be used for logical purposes.

The inductive definition of L . Suppose that $K = K_1 \cup K_2 \cup \cdots \cup K_n$ is a link of n components. Let $K - K_i$ denote the link obtained by removing the i th component from K . Let K_i (alone) stand for the knot diagram resulting from K by deleting all the components $K_1, K_2, \dots, K_{i-1}, K_{i+1}, \dots, K_n$.

Definition 6.4. The following is an inductive definition of the L -polynomial. Each stage in the definition depends upon the choice of basepoint on the diagram. (It will be shown that the polynomial is independent of these choices.)

(a) If $\widehat{K} = K(U, p)$ is a standard unknot (see the discussion above), then $L_{\widehat{K}} = a^{w(\widehat{K})}$ where $w(\widehat{K})$ is the writhe of \widehat{K} . (Note that $w(\widehat{K})$ is independent of the orientation assigned to this diagram by the directed basepoint p .)

(b) If K_1 is a knot diagram, and K_1 overlies a link diagram K_2 , then $L_{K_1 \sqcup K_2} = d L_{K_1} L_{K_2}$ where $d = (a + a^{-1})(z^{-1}) - 1$. (See the previous discussion for the definition of the term overlies.)

(c) If $K = K_1 \cup K_2 \cup \cdots \cup K_n$ is a link diagram with components K_1, K_2, \dots, K_n then

(i) If any component overlies all the others, apply part (b).

(ii) If no component K_i overlies all the others, let p_1, p_2, \dots, p_n be directed basepoints on K_1, \dots, K_n . Let $\bar{p}_1, \dots, \bar{p}_n$ be the same basepoints endowed with the opposite direction. Let $\lambda(p_i)$ be the sequence of undercrossings (order determined by directions of p_i) of K_i with $K - K_i$ so that $\widehat{K}(\lambda(p_i)) = K_i \sqcup (K - K_i)$ with K_i over-crossing the rest of these components. Since p_i determines $\lambda(p_i)$, $\sum_K(\lambda(p_i))$ depends only on the choice of directed basepoint p_i . Now define L_K in this case by the formula

$$L_K = \frac{1}{2n} \left[\sum_{i=1}^{|\lambda(p_i)|} (-1)^{|\lambda(p_i)|+1} d L_{K_i} L_{(K-K_i)} + z \sum_K(\lambda(p_i)) \right. \\ \left. + \sum_{i=1}^{|\lambda(\bar{p}_i)|} (-1)^{|\lambda(\bar{p}_i)|+1} d L_{K_i} L_{(K-K_i)} + z \sum_K(\lambda(\bar{p}_i)) \right].$$

(d) If K is a knot diagram, let p be a directed basepoint for K , \bar{p} the same basepoint with reversed direction. Let $\lambda(p)$ be the switching sequence

determined by p , and $\lambda(\bar{p})$ the switching sequence determined by \bar{p} . Define L_K by the formula

$$L_K = \frac{1}{2} \left((-1)^{|\lambda(p)|+1} L_{\widehat{K}(\lambda(p))} + z \sum_K (\lambda(p)) + (-1)^{|\lambda(\bar{p})|+1} L_{\widehat{K}(\lambda(\bar{p}))} + z \sum_K (\lambda(\bar{p})) \right).$$

This completes the inductive definition of L_K . Since, for each basepoint choice, we have included summations for both of the associated orientations, it suffices to show inductively that these definitions do not depend upon the choice of basepoint.

All induction arguments will be based on the number of crossings in the link diagrams. Thus I shall assume in each case that L_K has been verified to have a given property for all diagrams with less than n vertices, and then show that Definition 6.4 results in this property for links of n vertices. In all cases it is easy to check the bottom of the induction on knot and link diagrams with few vertices.

Definition 6.5 (the inductive hypothesis). The inductive hypothesis for the polynomial L_K defined in 6.4 is as follows:

(a) L_K is well defined (independent of basepoint choice) on diagrams with less than N crossings.

(b) L_K satisfies the axioms of evaluation:

$$\begin{cases} L_K + L_{S,K} &= z(L_{e,K} + L_{E,K}), \\ L_{\text{crossing}} &= aL, \\ L_{\text{loop}} &= a^{-1}L, \end{cases}$$

whenever K has $< N$ crossings (and whenever crossing has $< N$ crossings).

(c) L_K is invariant under Reidemeister moves of type II and type III that do not increase the number of crossings in the diagram. That is, if K has $< N$ crossings and K' is obtained from K by a Reidemeister (II, III) move that does not increase the number of crossings, then $L_K = L_{K'}$.

The proof of well-definedness will be a multiple induction on these properties.

The issue that must be addressed in order to prove that L_K is well defined on diagrams is its invariance with respect to change of basepoint in 6.4(c) and 6.4(d). The next lemma is relevant to 6.4(c).

Lemma 6.6. Let $\lambda = (\lambda_n, \lambda_{n-1}, \dots, \lambda_1, \lambda_0)$ be any choice of labels for a subset of distinct crossings in a link diagram K . Let $\mu = (\lambda_0, \lambda_n, \lambda_{n-1}, \dots, \lambda_2, \lambda_1)$. Then $\sum_K(\lambda) = \sum_K(\mu)$ where $\sum_K(\lambda)$ is defined as in Definition 6.3. In other words, $\sum_K(\lambda)$ is invariant under cyclic permutation of λ .

Proof. The proof is by induction on the number of crossings in the link diagram K . Thus we assume the lemma and the induction hypothesis 6.5 for all link diagrams with fewer crossings than K .

Notation. It is sufficient to take

$$\lambda = (n, n-1, \dots, 1, 0) \quad \text{and} \quad \mu = (0, n, n-1, \dots, 2, 1).$$

Then

$$\sum_K (\lambda) = \sum_{i=0}^n (-1)^i (L_{A_i^\lambda K} + L_{B_i^\lambda K})$$

where

$$A_i^\lambda = E_i S_{i-1} \cdots S_1 S_0, \quad B_i^\lambda = e_i S_{i-1} \cdots S_1 S_0,$$

and

$$\sum_K (\mu) = \sum_{i=0}^n (-1)^i (L_{A_i^\mu K} + L_{B_i^\mu K})$$

where

$$\begin{aligned} A_0^\mu &= E_1 \\ A_1^\mu &= E_2 S_1 \\ A_2^\mu &= E_3 S_2 S_1 \\ &\vdots \\ A_{n-1}^\mu &= E_{n-1} S_{n-1} \cdots S_2 S_1 \\ A_n^\mu &= E_0 S_n S_{n-1} \cdots S_2 S_1 \end{aligned}$$

and B_i^μ is obtained from A_i^μ by replacing E by e .

Thus

$$\begin{aligned} \sum_K (\lambda) - \sum_K (\mu) &= (L_{E_0 K} + L_{e_0 K}) - (L_{E_1 S_0 K} + L_{e_1 S_0 K}) + \cdots \\ &\quad + (-1)^n (L_{E_n S_{n-1} \cdots S_0 K} + L_{e_n S_{n-1} \cdots S_0 K}) \\ &\quad - [(L_{E_1 K} + L_{e_1 K}) - (L_{E_2 S_1 K} + L_{e_2 S_1 K}) + \cdots \\ &\quad + (-1)^{n-1} (L_{E_n S_{n-1} \cdots S_1 K} + L_{e_n S_{n-1} \cdots S_1 K}) \\ &\quad + (-1)^n (L_{E_0 S_n \cdots S_2 S_1 K} + L_{e_0 S_n \cdots S_2 S_1 K})] \\ &= (L_{E_0 K} + L_{e_0 K}) + (-1)^{n+1} (L_{E_0 S_n \cdots S_1 K} + L_{e_0 S_n \cdots S_1 K}) \\ &\quad - [L_{E_1 S_0 K} + L_{E_1 K} + L_{e_1 S_0 K} + L_{e_1 K}] \\ &\quad + [L_{E_2 S_1 S_0 K} + L_{E_2 S_1 K} + L_{e_2 S_1 S_0 K} + L_{e_2 S_1 K}] + \cdots \\ &\quad + (-1)^n [L_{E_n S_{n-1} \cdots S_0 K} + L_{E_n S_{n-1} \cdots S_1 K} + L_{e_n S_{n-1} \cdots S_0 K} + L_{e_n S_{n-1} \cdots S_1 K}]. \end{aligned}$$

Now use the exchange identity repeatedly to conclude (via induction) that

$$\begin{aligned}
 \sum_K (\lambda) - \sum_K (\mu) &= (L_{E_0 K} + L_{e_0 K}) + (-1)^{n+1} (L_{S_n \dots S_1 E_0 K} + L_{S_n \dots S_1 e_0 K}) \\
 &\quad - z[L_{E_1 e_0 K} + L_{E_1 E_0 K} + L_{e_1 e_0 K} + L_{e_1 E_0 K}] \\
 &\quad + z[L_{E_2 S_1 e_0 K} + L_{E_2 S_1 E_0 K} + L_{e_2 S_1 e_0 K} + L_{e_2 S_1 E_0 K}] + \dots \\
 &\quad + (-1)^n z[L_{E_n S_{n-1} \dots S_1 e_0 K} + L_{E_n S_{n-1} \dots S_1 E_0 K} \\
 &\quad \quad \quad + L_{e_n S_{n-1} \dots S_1 e_0 K} + L_{e_n S_{n-1} \dots S_1 E_0 K}] \\
 &= (L_{E_0 K} + (-1)^{n+1} L_{S_n S_{n-1} \dots S_1 E_0 K}) \\
 &\quad - z\{[L_{E_1 E_0 K} + L_{e_1 E_0 K}] - [L_{E_2 S_1 E_0 K} + L_{e_2 S_1 E_0 K}] \\
 &\quad \quad \quad \pm \dots - (-1)^n [L_{E_n S_{n-1} \dots S_1 E_0 K} + L_{e_n S_{n-1} \dots S_1 E_0 K}]\} \\
 &\quad + (L_{e_0 K} + (-1)^{n+1} L_{S_n S_{n-1} \dots S_1 e_0 K}) \\
 &\quad - z\{[L_{E_1 e_0 K} + L_{e_1 e_0 K}] - [L_{E_2 S_1 e_0 K} + L_{e_2 S_1 e_0 K}] \\
 &\quad \quad \quad \pm \dots - (-1)^n [L_{E_n S_{n-1} \dots S_1 e_0 K} + L_{e_n S_{n-1} \dots S_1 e_0 K}]\} \\
 &= 0 + 0 = 0
 \end{aligned}$$

by induction using 6.5(b) (repeatedly as in formula (*)). Note that we have used the independence of order of application of the operators S_i , E_i , e_i for a specific link diagram. This order-independence follows from the fact that these operators make only local changes in the diagram. This completes the proof of the lemma.

Remark. It follows at once from Lemma 6.6 that the formula for L_K in 6.4(c) is independent of the basepoint choice. In this case, the basepoint determines the sequence of crossing changes up to cyclic order, and we have included both directions for any given basepoint. Thus it remains to examine basepoint dependence in case 6.4(d).

Lemma 6.7. *Consider the two ways of splicing a standard unknot at the first crossing past (or prior) to a directed basepoint. In one case this splice produces an unknot. In the other case it produces an unlink composed of two standard unknots with one overlying the other.*

Proof. This lemma follows directly from the definition of the standard unknot. Consider the first crossing (i) past the basepoint. Starting at the basepoint and travelling in the direction that it indicates, one travels *over* the crossing i (this is the definition of the standard unknot: first encounters after the basepoint are overcrossings). The diagram that is then drawn (up until the return to i as an under-crossing) lies over the rest of the unknot diagram. One splice at i separates this first part from the rest, producing an unlink composed of two overlaid unknots. The other splice produces an unknot consisting in the connected sum of the two unknots discussed in the last sentence. See Figure 28.

This unknot diagram is not standard. See the comments following this lemma. This completes the proof of the lemma.

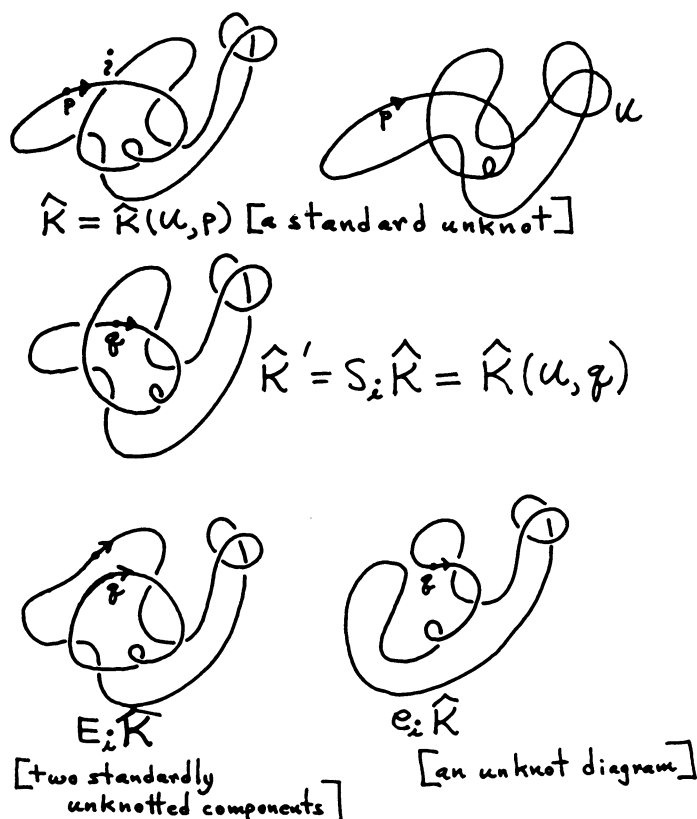


FIGURE 28

Comment. As Figure 28 illustrates, there are two splices $e_i \hat{K}$ and $E_i \hat{K}$. Here i is the first crossing encountered upon proceeding from the basepoint in a standard unknot \hat{K} . In this figure $E_i \hat{K}$ is a link (trivial link) of two standard unknots, while $e_i \hat{K}$ is an unknot diagram. If we write $E_i \hat{K} = K_1 \cup K_2$ where K_1 and K_2 are the two standard unknots in this link, then $e_i \hat{K} = K_1 \# K_2^*$ where this connected sum is taken at the crossing. As is easy to see, the standard unknot corresponding to $e_i \hat{K}$ is $K_1 \# K_2^*$ where K_2^* is obtained from K_2 by reversing all of its crossings.

A fact about standard unknot diagrams generalizes to the diagram $e_i \hat{K}$. This fact is that *standard unknot diagrams are either composed entirely of curls (∂^-) or they admit simplifying Reidemeister moves of type II and type III*. That is, a standard unknot diagram can be transformed by *noncomplexifying* moves of type II and type III to a diagram consisting only of curls. The proof of this remark involves noting that it is possible to move an arc that overpasses a

diagram across that diagram by regular isotopy. (See §2 (Figure 2).) These remarks apply to $e_i \hat{K}$, and consequently, we know that $L_{e_i \hat{K}} = a^{w(e_i \hat{K})}$ by applying 6.5(c) and the regular isotopy invariance of the writhe.

Lemma 6.8. *Let $\hat{K} = \hat{K}(U, p)$ be a standard unknot. Let i denote the first crossing on K that is encountered when travelling from the basepoint. Let q be the new directed basepoint obtained by sliding p just past the crossing i , retaining the direction from q . Then*

$$(i) \ S_i \hat{K}(U, p) = \hat{K}(U, q),$$

(ii) $L_{\hat{K}(U, p)} + L_{\hat{K}(U, q)} = z(L_{e_i \hat{K}(U, p)} + L_{E_i \hat{K}(U, p)})$ where these evaluations are based inductively on Definition 6.4, and the induction hypothesis 6.5.

Proof. The situation for this lemma is illustrated in Figure 28. We are given a standard unknot $\hat{K} = \hat{K}(U, p)$ and $\hat{K}' = \hat{K}(U, q)$ where q is the new basepoint described in the statement of the lemma. It follows at once from the definition of the standard unknot that $\hat{K}' = S_i \hat{K}$ (since the first pass through i will now be done by the other arc at this crossing). Represent the writhe of \hat{K} by $w(\hat{K}) = w + 1$ and $w(\hat{K}') = w - 1$. (They will differ by two, possibly with reversed signs. The argument to follow will go through exactly as given, with this sign-reversal.) Then $w(e_i \hat{K}) = w$ where $e_i \hat{K}$ is the connected splice. Let w_1 and w_2 be the writhes of the two components of $E_i \hat{K}$. Then $w_1 + w_2 = w$ since, by the Jordan curve theorem, the sum of the crossing sign contributions between these two components adds up to zero (one curve lies entirely over the other). Thus, using the *comment* prior to this lemma, we have

$$\begin{aligned} L_{\hat{K}} &= a^{w+1}, \\ L_{S_i \hat{K}} &= a^{w-1}, \\ L_{e_i \hat{K}} &= a^w, \\ L_{E_i \hat{K}} &= d a^{w_1} a^{w_2} = d a^{w_1+w_2} = d a^w. \end{aligned}$$

Since, by definition of d , $a + a^{-1} = z(1 + d)$, we have $a^{w+1} + a^{w-1} = z(a^w + d a^w)$, it follows that $L_{\hat{K}} + L_{S_i \hat{K}} = z(L_{e_i \hat{K}} + L_{E_i \hat{K}})$, proving the lemma. Note that the evaluations on the left-hand side are evaluations of standard unknots—part of the definition of L_K . The evaluations on the right-hand side follow by induction with these links having smaller number of crossings.

Lemma 6.9. *Let K be a knot diagram. Then the value of L_K as described in Definition 6.4 (d) is independent of the choice of basepoint. More particularly, let p be a directed basepoint for K . Let $\lambda(p)$ be the switching sequence determined by p . Then*

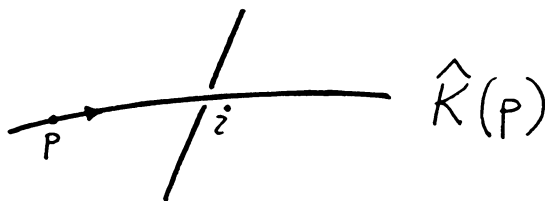
$$\Omega_K(p) = (-1)^{|\lambda(p)|+1} L_{\hat{K}(p)} + z \sum_K(p)$$

is independent of the choice of basepoint. (Here $\hat{K}(p) = \hat{K}(\lambda(p))$ is the standard

unknot associated with p , and $\sum_K(p) = \sum_K(\lambda(p))$ is the summation defined in Definition 6.3.)

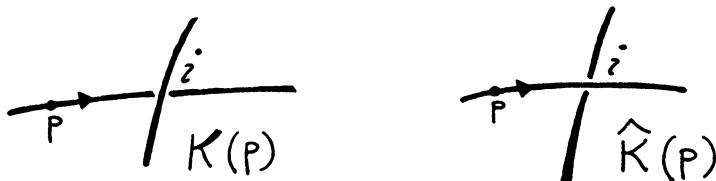
Proof. We may assume that the switching sequence determined by p is labelled $\lambda = (n, n-1, \dots, 1, 0)$ so that $\Omega_K(p) = (-1)^{n+1} L_{\hat{K}(p)} + z \sum_K(p)$. Since $L_K(p) = \frac{1}{2}(\Omega_K(p) + \Omega_K(\bar{p}))$ it suffices to prove that $\Omega_K(p)$ is independent of basepoint. The proof is by induction on the number of crossings in the link diagram, and depends upon the inductive definition and induction hypothesis.

Consider the first crossing i past the basepoint in K and in $\hat{K}(p)$. In $\hat{K}(p)$ this is an over-crossing by definition.



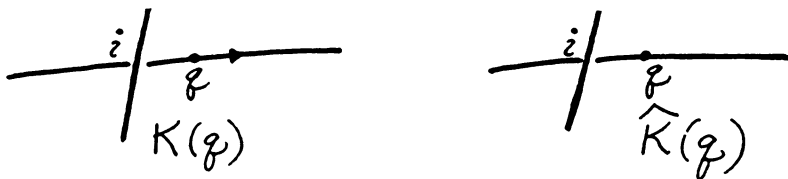
In $K(p)$ the crossing may be either an undercrossing or an undercrossing.

Case 1. The crossing i is an undercrossing in $K(p)$:



In this case i is part of the switching sequence for \hat{K} relative to K . In fact, by our convention, i receives the label n (highest index) for the switching sequence $\hat{K}(p) = S_n S_{n-1} \cdots S_1 S_0 K$.

Now consider the result, in this case, of sliding the basepoint p to position q just past the crossing. Then we have



that the crossing i is not switched in $\hat{K}(q)$. Thus the switching sequence for $K(q)$ is $(n-1, n-2, \dots, 1, 0)$ and $\hat{K}(q) = S_{n-1} S_{n-2} \cdots S_1 S_0 K$.

We wish to show that $\Omega_K(q) = \Omega_K(p)$. We have

$$\Omega_K(p) = (-1)^{n-1} L_{\hat{K}(p)} + z \sum_K(p),$$

$$\Omega_K(q) = (-1)^{(n-1)+1} L_{\hat{K}(q)} + z \sum_K(q),$$

where

$$\sum_K(p) = \sum_{i=0}^n (L_{A_i K} + L_{B_i K})(-1)^i$$

and

$$\sum_K(q) = \sum_{i=0}^{n-1} (L_{A_i K} + L_{B_i K})(-1)^i.$$

The switching sequence for $K(q)$ is a subset of the switching sequence for $K(p)$. Thus

$$\Omega_K(p) - \Omega_K(q) = (-1)^{n+1} [L_{\widehat{K}(p)} + L_{\widehat{K}(q)}] + z(L_{A_n K} + L_{B_n K})(-1)^n.$$

However, we have

$$\begin{aligned} A_n K &= E_n S_{n-1} \cdots S_1 S_0 K = E_n \widehat{K}(p), \\ B_n K &= e_n S_{n-1} \cdots S_1 S_0 K = e_n \widehat{K}(p), \\ \widehat{K}(q) &= S_n \{\widehat{K}(p)\} \end{aligned}$$

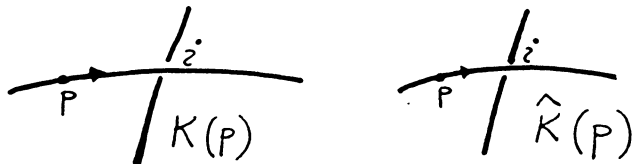
and, by Lemma 6.8,

$$L_{\widehat{K}(p)} + L_{\widehat{K}(q)} = z(L_{e_n \widehat{K}(p)} + L_{E_n \widehat{K}(q)}).$$

Therefore $\Omega_K(p) = \Omega_K(q)$.

We have verified that in Case 1 the value of $\Omega_K(p)$ is unchanged by sliding the basepoint past the first crossing along its direction.

Case 2. The crossing i is an overcrossing in $K(p)$:



In this case, i is not part of the switching sequence for $K(p)$. When the basepoint is slid across i we have



that i is part of the switching sequence for $K(q)$. Let i stand for the place in this switching sequence so that

$$\widehat{K}(q) = S_n S_{n-1} \cdots S_{i+1} S_i S_{i-1} \cdots S_1 S_0 K$$

(i.e., $\lambda(q) = (n, n-1, \dots, i+1, i, i-1, \dots, 1, 0)$). Then $\lambda(p) = (n, n-1, \dots, i+1, i-1, \dots, 1, 0)$ and $\widehat{K}(p) = S_n S_{n-1} \cdots S_{i+1} S_{i-1} \cdots S_1 S_0 K$. Note

that $S_i \hat{K}(q) = \hat{K}(p)$, and that Lemma 6.8 applies to the pair $\hat{K}(p)$, $\hat{K}(q)$ with respect to i , this being the case of the first crossing in $\hat{K}(q)$ just prior to the basepoint—the same arguments apply. However, we can also simplify the calculations for $\sum_K(p)$ and $\sum_K(q)$ by using Lemma 6.6. Lemma 6.6 says that these sums are invariant under cyclic permutation. Therefore, we can replace $\lambda(p)$ and $\lambda(q)$ by

$$\begin{aligned}\lambda'(p) &= (i-1, \dots, 1, 0, n, n-1, \dots, i+1), \\ \lambda'(q) &= (i, i-1, \dots, 1, 0, n, n-1, \dots, i+1).\end{aligned}$$

At this point the structure of the argument is identical to that of the first case, and the same form of calculation shows that $\Omega_K(p) = \Omega_K(q)$.

Thus we have verified that $\Omega_K(p)$ is independent of the placement of the oriented basepoint along the knot diagram. This completes the proof of Lemma 6.9.

Lemma 6.10. *Let i be any crossing in a link diagram K . Then L_K satisfies the identities*

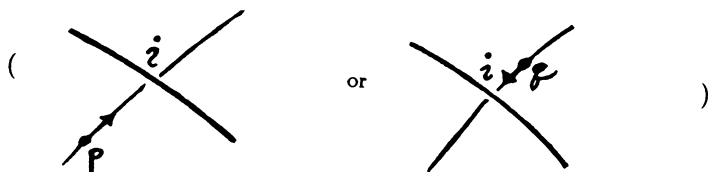
$$(a) \quad L_K + L_{S_i K} = z(L_{e_i K} + L_{E_i K}),$$

$$(b) \quad L_{\overline{\partial}^-} = aL, \quad L_{\overline{\partial}^+} = a^{-1}L.$$

Proof. The proof is by induction on the number of crossings in the diagram. We use the already-proved fact (Lemma 6.6 and Lemma 6.9) that L_K as defined in 6.4 is independent of the choice of basepoint. Note that at no point in the argument up to this point have we used the conclusion of Lemma 6.10, except in the special case form of Lemma 6.8.

We have shown in 6.9 that in the case of a single component diagram the individual summations $\Omega_K(p)$ and $\Omega_K(\overline{p})$ are each independent of the placement of the basepoint.

1. Suppose that K has one component and a given crossing i . Let SK be the knot obtained by switching this crossing, and EK , eK the two diagrams obtained by splicing this crossing. By choosing the basepoint correctly



we can make i the “ n th” crossing in the switching sequence for K . Comparison with 6.4(d) shows that the identity $L_K + L_{SK} = z(L_{EK} + L_{eK})$ follows directly from the difference of the expansions of $\Omega_K(p)$ and $\Omega_{SK}(p)$.

2. Suppose that K has more than one component, and that at the crossing i we have a crossing of a component with itself. Then this crossing will not be involved in the lift sequence for any component. Hence the identity follows by induction, using 6.4(c).

3. Suppose that K has more than one component, and that the two strands

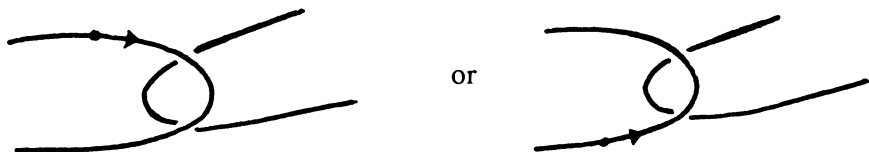
in the crossing are from different components. Recall that L_K is by 6.4(c) the average over evaluations from lift sequences for liftable components. (A component is liftable if it does not overlie the entire diagram.) Consider one of these summands. If the component being lifted does not contain either of the two strands at the crossing, then the exchange formula follows by induction. If one of the strands is involved in the lift sequence, then appropriate choice of basepoint exhibits the identity as before. Since the identity is established for each summand (for a given lift sequence), it is established for their average.

This completes the proof of part (a). To verify part (b) simply note that the curl will, in the recursive Definition 6.4, eventually be part of a knot-evaluation, and that by choosing the position of the basepoint this curl will be unchanged on the corresponding standard unknot. Hence all terms in the right-hand side of 6.4(d) contain identical copies of this curl. (b) then follows by induction.

This completes the proof of the lemma.

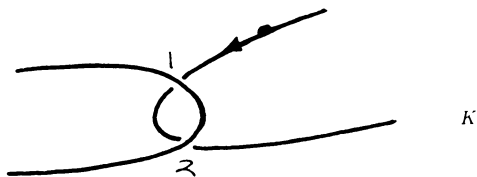
Lemma 6.11. *Let K be any link diagram and K' another link diagram regularly isotopic to K . Then the polynomials L_K and $L_{K'}$ are equal. L_K is an invariant of regular isotopy.*

Proof. If K is a knot, the invariance under moves of type II and type III (these generate regular isotopy) can be seen, inductively, by appropriate choice of basepoint: Thus in the case of a type II move, choose the basepoint as shown below:



Then the two crossings involved in the move are not switched by the switching sequence for K . Since we inductively assume (6.5(c)) that each term in the expansion of 6.4(d) is invariant under type II (simplifying) moves, it follows that L_K is also invariant under the type II move. Here it is assumed that the proposition (invariance under simplifying type II moves) has been verified for all knots/links with fewer crossings than K .

If K has more than one component, then we must consider those cases of type II moves where one of the strands is involved in the lift sequence. The worst possible case then corresponds to a basepoint choice of the form

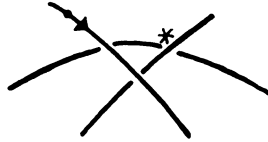


where the basepoint is necessarily on the undercrossing strand in order to accomplish the lift sequence. However, Figure 29 coupled with Lemma 6.10 shows

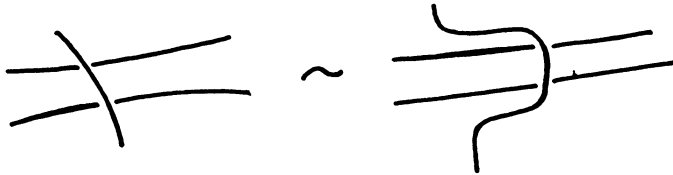
that $L_K = L_{S_1 S_2 K}$. In the link $S_1 S_2 K$ the type II move will not be involved in a lift sequence and invariance follows by induction as before.

Having verified that L_K is invariant under simplifying type II moves for all knots and links, we conclude that L_K is type II invariant for both directions of the type II move.

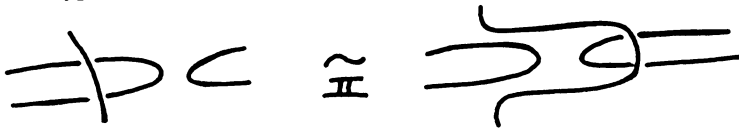
Now consider the type III move. Suppose that K is a knot diagram. By choosing the basepoint, we arrange that two out of three crossings are uninvolved in the switching sequence:



The remaining starred crossing may be switched or spliced in any of the terms of the expansion formula 6.4(d). If it is switched then the type III move is applicable inductively. If it is spliced, then invariance follows either by direct diagram equivalence

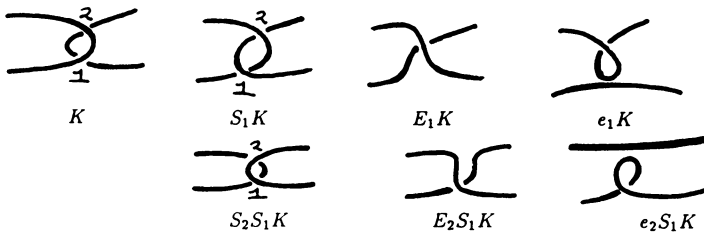


or by using the type II invariance



For the type III move in the multi-component case, the ideas are the same. Here two consecutive crossings may be involved in a lift sequence. If these are labelled 1, 2 as in Figure 29 then $L_K = L_{S_2 S_1 K}$ again applies to simplify the situation, and allows us to conclude invariance under the type III move by mathematical induction.

This completes the proof.



$$\begin{aligned}
 L_K + L_{S_1 K} &= z(L_{E_1 K} + L_{e_1 K}), \\
 L_{S_1 K} + L_{S_2 S_1 K} &= z(L_{E_2 S_1 K} + L_{e_2 S_1 K}), \\
 L_{E_1 K} &= L_{E_2 S_1 K}, L_{e_1 K} = aL_{\sim} = L_{e_2 S_1 K}, \\
 &\Rightarrow L_K = L_{S_2 S_1 K}.
 \end{aligned}$$

FIGURE 29

VII. The Dubrovnik polynomial

The Dubrovnik polynomial $D_K(a, z)$ is the name I give to another version of L_K . (It was discovered in the old city of Dubrovnik, Yugoslavia and told to W. B. R. Lickorish by postcard. That postcard received an immediate reply to the effect that D_K and L_K are interconvertible!) The Dubrovnik polynomial is also an invariant of regular isotopy, and it satisfies the axioms

1. $D_{\nearrow \searrow} - D_{\searrow \nearrow} = z(D_{\leftrightsquigarrow} - D_{\rightleftharpoons})$,
2. $D_{\bigcirc} = aD$, $D_{\bigcirc} = 1$, $D_{-\bigcirc} = a^{-1}D$.

Note that the minus sign makes sense relatively speaking. Thus $\nearrow : \leftrightsquigarrow :: \searrow : \rightleftharpoons$ in the sense that the second pair is obtained from the first pair by 90° rotation, and we can, by convention, associate \nearrow and \leftrightsquigarrow .

Proposition (Lickorish). $D_K(a, z) = -i^{-w(K)}(-1)^{c(K)} L_K(ia, -iz)$ where i denotes the square root of negative one ($i = \sqrt{-1}$), and $c(K)$ is the number of components of K .

I shall omit the proof of this proposition, but give one example to illustrate it.

Example. We have

$$\begin{aligned} D_{\bigcirc} - D_{-\bigcirc} &= z(D_{\bigcirc} - D_{-\bigcirc}) \\ &\Rightarrow a - a^{-1} = z(D_{\bigcirc} - 1). \end{aligned}$$

Thus $D_{\bigcirc} = z^{-1}(a - a^{-1}) + 1$. Now

$$\begin{aligned} c(\bigcirc) &= 2, \\ w(\bigcirc) &= 0. \end{aligned}$$

Thus if $\Lambda = \bigcirc$, then

$$\begin{aligned} &-i^{-w(\Lambda)}(-1)^{c(\Lambda)} L_{\Lambda}(ia, -iz) \\ &= -(1)(-1)^2 L_{\Lambda}(ia, -iz) \\ &= -L_{\Lambda}(ia, -iz) \quad (L_{\Lambda} = z^{-1}(a + a^{-1}) - 1) \\ &= -[(-iz)^{-1}(ia - ia^{-1}) - 1] \\ &= -[iz^{-1}(a - a^{-1})i - 1] \\ &= z^{-1}(a - a^{-1}) + 1 \\ &= D_{\Lambda}(a, z). \end{aligned}$$

There are some formal advantages in using the Dubrovnik polynomial. For example, we can construct an algebra analogous to the Birman-Wenzel algebra described in §4. Then the basic relation would take the form $c_i - c_i^{-1} = z(1 - h_i)$. Then at $z = 0$, the algebra collapses directly to the group algebra of the symmetric group.

I would like to give an account here of how I came upon the Dubrovnik polynomial. This will show another facet of this reformulation: Knowing that the bracket could be constructed as a state-model via $\langle \bowtie \rangle = A \langle \asymp \rangle + A^{-1} \langle \rangle \langle \rangle$ I decided to try (in October 1985) to construct a more general model in the form

$$[\bowtie] = A [\asymp] + B [\rangle \langle] + C [\times]$$

where the third local configuration is a simple crossing (not under or over). In other words it is a 4-valent graphical vertex. Now an inductive expansion of this type is easily seen to be well defined *if we know the values of $[G]$ for 4-valent planar graphs*. There are many possibilities for graphical evaluation, and I tried what seemed the simplest, namely $[G] = d^{|G|-1}$ where $|G|$ is the number of “knot-theoretic circuits” in G . That is $|G|$ is the number of link components if G is regarded as a link projection. Thus

$$[\bigcirc \bigcirc] = [\bigcirc \bigcirc] = [\bigcirc \bigcirc] = d^{2-1} = d^1 = d$$

since these graphs all have two link components.

With this definition, we have a generalization of the bracket [20], and can investigate its invariance under Reidemeister moves. The case of the bracket is $C = 0$, $B = A^{-1}$, $d = -A^2 - A^{-2}$. I found that the only other solution that is a regular isotopy invariant (with this form of graphical evaluation) has the form

$$\begin{cases} [\bowtie] = (\frac{\alpha - \alpha^{-1}}{2}) [\asymp] + (\frac{\alpha^{-1} - \alpha}{2}) [\rangle \langle] + (\frac{\alpha + \alpha^{-1}}{2}) [\times], \\ [\wr] = \alpha [\neg] , [\neg \neg] = \alpha^{-1} [\neg] , \\ [\bigcirc] = 1, \\ [G] = 2^{|G|-1}, G \text{ a 4-valent planar graph.} \end{cases}$$

Note that it follows at once from these identities that this regular isotopy invariant satisfies the identities

$$\begin{aligned} [\bowtie] - [\times] &= (\alpha - \alpha^{-1}) ([\asymp] - [\rangle \langle]), \\ [\wr] &= \alpha [\neg] , [\neg \neg] = \alpha^{-1} [\neg] , \\ [\bigcirc] &= 1. \end{aligned}$$

Thus $[K]$ is a special case of the Dubrovnik polynomial.

In fact, we have

Proposition 7.1. $[K]$ with $A = ((\alpha - \alpha^{-1})/2)$, $B = ((\alpha^{-1} - \alpha)/2)$, $C = ((\alpha + \alpha^{-1})/2)$, and $|G| =$ the number of link components in a 4-valent planar graph G , $[G] = 2^{|G|-1}$, is given by

$$(*) \quad [K] = \frac{1}{2} \sum_{\mathcal{O}} \alpha^{w(K^{\mathcal{O}})}$$

where \mathcal{O} runs over all orientations $K^{\mathcal{O}}$ of the unoriented link K . Here $w(K^{\mathcal{O}})$ is the writhe of the oriented link $K^{\mathcal{O}}$.

Proof (sketch). Extend the writhe to oriented diagrams with knot-theoretic (\nearrow, \searrow) and graph-theoretic (\bowtie) vertices by summing over the signs of the knot theoretic vertices. ($w(\nearrow) = +1$.) Define $[K]$ by formula (*) in Proposition 7.1. Let $[\bowtie]_+$ denote that part of the summation where this particular crossing has positive sign. Let $[\bowtie]_-$ denote that part of the summation where the indicated crossing has negative sign. Thus

$$[\bowtie] = [\bowtie]_+ + [\bowtie]_-$$

by definition. Then verify that

$$[\nearrow]_+ = \left(\frac{\alpha}{2}\right)([\bowtie] + [\smile] - [\) (\])$$

and

$$[\nearrow]_- = \left(\frac{\alpha^{-1}}{2}\right)([\bowtie] - [\smile] + [\) (\]).$$

(These last two formulas are verified by checking the various cases of local orientations. Thus $[\bowtie]$ can take the local forms $\bowtie, \bowtie, \bowtie, \bowtie$ while $[\smile]$ can take the forms $\Rightarrow, \Leftarrow, \Rightarrow, \Leftarrow$. Note that forms of type \Rightarrow, \Leftarrow are in 1-1 correspondence with forms of type $\downarrow\uparrow, \uparrow\downarrow$ by local resplicing. Thus $[\smile] - [\) (\]$ counts \Rightarrow, \Leftarrow and subtracts $\uparrow\uparrow, \downarrow\downarrow$. Hence $[\bowtie] + [\smile] - [\) (\]$ counts \bowtie and \bowtie twice since \bowtie and \bowtie are also in one-to-one correspondence. Thus $\frac{1}{2}([\bowtie] + [\smile] - [\) (\])$ counts all cases of positivity. It must be multiplied by α to take the crossing into account.) We then obtain the defining formula for

$$[\bowtie] = \left(\frac{\alpha + \alpha^{-1}}{2}\right)[\bowtie] + \left(\frac{\alpha - \alpha^{-1}}{2}\right)[\smile] + \left(\frac{\alpha^{-1} - \alpha}{2}\right)[\) (\]$$

by adding these two formulas for the \pm writhe contributions. This completes the proof.

Remark. If K is a link of two components then $[K] = \alpha^a + \alpha^b$ where $|(a - b)|/4 = |\text{lk}(K)|$. Thus, this special case state model for the Dubrovnik polynomial calculates the absolute value of the linking number of two curves from an unoriented diagram for the link.

This is a remarkable fact. In principle the absolute linking number does not depend upon orientation. That it is possible to give an algorithm to calculate it that does not use orientation is amazing. (A similar version of this model and the corresponding conclusion about linking numbers was independently obtained by Andrew Lipson [32].)

There is much more to say about state models. I have shown this one because it is simple, gives a different view of linking numbers, and is a special case of the Dubrovnik polynomial.

The *form* of expansion

$$[\bowtie] = A[\smile] + B[\) (\] + [\bowtie],$$

and

$$[\text{X}] = B[\text{Y}] + A[\text{Z}] + [\text{W}]$$

will always satisfy the Dubrovnik identity $[\text{X}] - [\text{Y}] = z([\text{Z}] - [\text{W}])$ with $z = A - B$. Pierre Vogel and the author [25] observe that this triple of equations can be regarded as the form of *definition* for $[K]$ on graph-diagrams. This definition depends upon the theory of the polynomial L_K . In [25] we give a graphical calculus for evaluating $[G]$ on planar graphs in this generality and show that this extension of the polynomials to three variables gives rigid-vertex isotopy invariants for graph embeddings in three space.

Remark (added in proof). State models using solutions to the Yang-Baxter Equation (see [0], [15.1], [26.1], [45.1]) can be used to establish the existence of the polynomial invariants discussed in this paper. In [48] Witten has initiated three dimensional interpretations and generalizations of these invariants via quantum field theory.

REFERENCES

0. Y. Akutsu and M. Wadati, *Exactly solvable models and new link polynomials. I. N-state vertex models*, J. Phys. Soc. Japan **56** (1987), 3039–3051.
1. J. W. Alexander, *Topological invariants of knots and links*, Trans. Amer. Math. Soc. **20** (1923), 275–306.
2. R. Ball and M. L. Mehta, *Sequence of invariants for knots and links*, J. Physique **42** (1981), 1193–1199.
3. J. S. Birman, *Braids, links and mapping class groups*, Ann. of Math. Stud., no. 82, Princeton Univ. Press, Princeton, N. J., 1976.
4. J. S. Birman and H. Wenzel, *Braids, link polynomials and a new algebra*, preprint, 1986.
5. G. Burde and H. Zieschang, *Knots*, De Gruyter Studies in Math. no. 5, De Gruyter, 1985.
6. R. D. Brandt, W. B. R. Lickorish and K. C. Millett, *A polynomial invariant for unoriented knots and links*, Invent. Math. **74** (1986), 563–573.
7. R. Brauer, *On algebras which are connected with the semisimple continuous groups*, Ann. of Math. **38** (1937).
8. J. H. Conway, *An enumeration of knots and links and some of their algebraic properties*, Computational Problems in Abstract Algebra, Pergamon Press, New York, 1970, pp. 329–358.
9. R. H. Crowell and R. H. Fox, *Introduction to knot theory*, Blaisdell, 1963.
10. P. Freyd, D. Yetter, J. Hoste, W. B. R. Lickorish, K. C. Millett and A. Ocneanu [HOMFLY], *A new polynomial invariant of knots and links*, Bull. Amer. Math. Soc. **12** (1985), 239–246.
11. W. Graue, *Die semilinearen abbildungen*, S. B. Heidelberg Adad. Wiss. Math.-Nat. Kl., 1950, pp. 205–272.
12. C. F. Ho, *A new polynomial invariant for knots and links—preliminary report*, Abstracts Amer. Math. Soc. **6** (1985), 300.
13. V. F. R. Jones, *A new knot polynomial and von Neumann algebras*, Notices Amer. Math. Soc. (1985).
14. —, *A polynomial invariant for links via von Neumann algebras*, Bull. Amer. Math. Soc. **12** (1985), 103–112.

15. —, *Hecke algebra representations of braid groups and link polynomials*, Ann. of Math. **126** (1987), 335–388.
- 15.1. V. F. R. Jones, *On knot invariants related to some statistical mechanical models*, Pacific J. Math. **137** (1989).
16. T. Kanenobu and M. Sakuma, *A note on the Kauffman polynomial*, preprint, 1986.
17. L. H. Kauffman, *The Conway polynomial*, Topology **20** (1980), 101–108.
18. —, *Formal knot theory*, Princeton Univ. Press Math. Notes, no. 30, Princeton Univ. Press, 1983.
19. —, *On knots*, Ann. of Math. Stud., no. 115, Princeton Univ. Press, Princeton, N. J., 1987.
20. —, *State models and the Jones polynomial*, Topology **26** (1987), 395–407.
21. —, *New invariants in the theory of knots* (lectures given in Rome, June 1986), (to appear).
22. —, *New invariants in the theory of knots*, Amer. Math. Monthly Vol. 95 (1988), 195–242.
23. —, *Statistical mechanics and the Jones polynomial*, Proc. Conf. on Artin's Braid Group, Santa Cruz, Calif., 1986, Contemp. Math., vol. 78, Amer. Math. Soc., Providence, R.I., 1988, pp. 263–297.
24. —, *An invariant of regular isotopy*, Announcement, 1985.
25. —, *Knots and physics* (in preparation).
26. L. H. Kauffman and P. Vogel, *Link polynomials and a graphical calculus*, Announcement, 1987.
- 26.1. L. H. Kauffman, *Knots, abstract tensors and the Yang-Baxter equation*, Knots, Topology and Quantum Field Theories - Proceedings of the Johns Hopkins Workshop on Current Problems in Particle Theory 13, Florence (1989). Edited by L. Lusanna, World Scientific, 1989, pp. 179–334.
27. M. Kidwell, *On the degree of the Brandt-Lickorish-Millett-Ho polynomial of a link*, Proc. Amer. Math. Soc. **100** (1987), 755–761.
28. T. P. Kirkman, *The enumeration, description and construction of knots with fewer than 10 crossings*, Trans. Roy. Soc. Edinburgh **32** (1865), 281–309.
29. W. B. R. Lickorish and K. C. Millett, *A polynomial invariant for oriented links*, Topology **26** (1987), 107–141.
31. W. B. R. Lickorish, *A relationship between link polynomials*, Math. Proc. Cambridge Philos. Soc. **100** (1986), 109–112.
32. A. S. Lipson, *Smith's prize essay*, Univ. of Cambridge, 1987.
33. C. N. Little, *Non-alternate $+-$ knots*, Trans. Roy. Soc. Edinburgh **35** (1889), 663–664.
34. H. R. Morton, *Seifert circles and knot polynomials*, Math. Proc. Cambridge Philos. Soc. **99** (1986), 107–109.
35. H. R. Morton and H. B. Short, *Calculating the 2-variable polynomial for knots presented as closed braids*, preprint, 1986.
36. H. Murakami, *A formula for the two-variable Jones polynomial*, Topology **26** 409–412.
37. J. Murakami, *Kauffman polynomial of links and representation theory*, preprint, 1986.
38. K. Murasugi, *Jones polynomials and classical conjectures in knot theory*, Topology **26** (1987), 187–194.
39. —, *Jones polynomials and classical conjectures in knot theory*. II, preprint, 1986.
40. J. Przytycki, *Conway formulas for Jones-Conway and Kauffman polynomials*, preprint, 1986.
41. K. Reidemeister, *Knotentheorie*, Chelsea, New York, 1948.
42. D. Rolfsen, *Knots and links*, Publish or Perish, Berkeley, Calif., 1976.
43. P. G. Tait, *On knots*. I, II, III, Scientific Papers. Vol. I, Cambridge Univ. Press, London, 1898, pp. 273–347.
44. M. Thistlethwaite, *A spanning tree expansion of the Jones polynomial*, Topology **26** 297–309.

- 45. —, *Kaufmann's polynomial and alternating links*, *Topology* **27** (1988), 311–318. 1986.
- 45.1. V. G. Turaev, *The Yang-Baxter equations and invariants of links*, *Invent. Math.* **92** (1988), 527–553.
- 46. B. Trace, *On the Reidemeister moves of a classical knot*, *Proc. Amer. Math. Soc.* **89** (1983).
- 47. H. Whitney, *On regular closed curves in the plane*, *Comp. Math.* **4** (1937), 276–284.
- 48. E. Witten, *Quantum field theory and the Jones polynomial*, *Comm. Math. Phys.* **121** (1989), 351–399.
- 49. D. Yetter, *Markov algebras*, *Proc. Conf. on Artin's Braid Group*, Santa Cruz, Calif., 1986, *Contemp. Math.*, vol. 78, Amer. Math. Soc., Providence, R.I., 1988, pp. 705–730.

DEPARTMENT OF MATHEMATICS, STATISTICS, AND COMPUTER SCIENCE, UNIVERSITY OF ILLINOIS
AT CHICAGO, CHICAGO, ILLINOIS 60680