

Description of the Topological Entanglement of DNA Catenanes and Knots by a Powerful Method Involving Strand Passage and Recombination

James H. White

*Department of Mathematics
University of California
Los Angeles, CA, U.S.A.*

K. C. Millett

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

and Nicholas R. Cozzarelli

*Department of Molecular Biology
University of California
Berkeley, CA, U.S.A.*

(Received 30 January 1987, and in revised form 3 June 1987)

We utilize a recently discovered, powerful method to classify the topological state of knots and catenanes. In this method, each such form is associated with a unique polynomial. These polynomials allow a rigorous determination of whether knotted or catenated DNA molecules that appear distinct actually are, and indicate the structure of related molecules. A tabulation is given of the polynomials for all possible stereoisomers of many of the knotted and catenated forms that are found in DNA. The polynomials for a substrate DNA molecule and the products obtained from it by either recombination or strand passage by a topoisomerase are related by a simple theorem. This theorem affords natural applications of the polynomial method to these processes. Examples are presented involving site-specific recombination by the transposon Tn3-encoded resolvase and the phage λ integrase, in which product structure is predicted as a function of crossover mechanism.

1. Introduction

All natural populations of DNA rings are to some extent interlocked as catenanes and knots (for reviews, see Kasamatsu & Vinograd, 1974; and Wasserman & Cozzarelli, 1986). The ubiquity of these forms is due to the many processes capable of generating them. The formation of a catenated intermediate bypasses a topological problem at the termination of DNA replication (Sundin & Varshavsky, 1980, 1981), and both knots and catenanes are generated by recombination and topoisomerase action (for reviews, see Nash, 1981; Wang, 1985; Wasserman & Cozzarelli, 1986). The importance of understanding the structure of these linked forms goes beyond circular DNA, because

linear chromosomes *in vivo* can be interlocked as a result of topologically confined subdomains (Stonington & Pettijohn, 1971; Benyajati & Worcel, 1976; DiNardo *et al.*, 1984; Uemura & Yanagida, 1984).

The complete stereostructure of knots and catenanes can now be determined by electron microscopy of protein-coated DNA molecules (Krasnow *et al.*, 1983; Griffith & Nash, 1985; Wasserman & Cozzarelli, 1984) and this information has provided critical insight into the mechanism of processes that generate and unravel these linked forms. The power of topological methods lies in the ability to distinguish among alternative models by predicting which of an often astronomic number of knots and catenanes should arise as products. For

example, a proposed scheme for bacteriophage λ integrative recombination correctly predicted the structure of a product knot out of 10^8 possible alternative forms (Spengler *et al.*, 1985).

In the face of such great complexity, it is imperative to have a rigorous classification scheme that determines if two molecules that appear different are actually topologically distinct; i.e. one cannot be deformed into the other without breaking backbone bonds. Manipulation of physical models for DNA, such as ribbons, can decide this unambiguously only in the simplest of cases. Classification requires associating with knots and catenanes numbers or polynomials. These are invariants because they cannot be changed by any deformation of the structure short of backbone breakage, and therefore describe a topoisomer in any of its entangled forms.

We have shown how the Schubert (1956) classification scheme can be applied to certain knots and catenanes produced by DNA replication and recombination (White & Cozzarelli, 1984). From the value of two integers called α and β , a simple theorem allows one to determine whether two forms are identical and to compute the number and associated invariants of related forms. Although the Schubert scheme classified most of the DNA knots and catenanes whose structure had been determined at that time, it has three limitations. First, the method applies only to two-bridge forms, those that can be represented in plane projection with only two DNA segments bridging any number of underpassing DNA segments. Recent work has shown the importance of forms that are not two-bridge. For example, all two-bridge forms are prime, i.e. cannot be factored (decomposed) into simpler forms, but topoisomerases can produce compound, i.e. non-prime, knots (Dean *et al.*, 1985), and recombination can produce knotted catenanes (Wasserman & Cozzarelli, 1986). Even among prime knots and catenanes, the fraction that is two-bridge is small. Second, in order to determine α and β graphically and without knowledge of other invariants, one must redraw the knot or catenane in the Schubert fashion with just two overpasses and, for all but the simplest forms, this is extremely difficult. Third, the two-bridge drawings are mathematical artifices that are very different from what DNA looks like in solution (Dean *et al.*, 1985) and have not been useful in predicting the changes in topology brought about by enzymes.

All three problems are solved simultaneously by a recent very important advance in knot and catenane theory involving polynomial invariants (Jones, 1985; Freyd *et al.*, 1985; Lickorish & Millett, 1987; Hoste, 1986). The polynomial methods classify not just the two-bridge knots and catenanes but all knots and catenanes. The polynomials can be calculated by a simple iterative procedure starting with standard drawings of knots and catenanes such as found in the Rolfsen (1976) Tables. It is of particular importance to the biologist that a central theorem relates the changes

in the polynomials made by strand passage and strand exchange. These two operations are the hallmarks of topoisomerases and recombination, the two principal means known for changing DNA structure in biological systems (Wasserman & Cozzarelli, 1986). Therefore, the polynomials allow one to formulate and test the mechanism of the processes with mathematical precision.

The first of the new generation of classification methods using polynomial invariants was discovered by Jones (1985). The second was discovered independently and simultaneously by Lickorish & Millett, Freyd & Yetter, Ocneanu, and Hoste (Freyd *et al.*, 1985), all of whom were attempting to clarify and extend the pioneering results of Jones. Unlike the Jones polynomial, which has one variable, the later methods use a two-variable polynomial. These invariants were foreshadowed by the classical Alexander polynomial (Alexander, 1928) and its subsequent elaboration by Conway (1969). The Alexander polynomial, however, does not distinguish mirror images, and enzymes generally have chiral products.

The polynomial methods are so new and rapidly evolving that even in the mathematical literature they are published only in outline form. Thus, the treatment of the methods in this paper are self-contained and require no knowledge of the topological literature. The second of the seven sections of this paper presents the two-variable polynomial theory of Lickorish & Millett (1987) and the main theorems useful in classification. We introduce the method using the two-variable polynomial rather than the simpler Jones polynomial for two reasons. First, *via* a simple substitution, the Jones polynomial can be calculated from the two-variable polynomial but not *vice versa*. Second, the two-variable polynomial is more powerful, in that there are fewer examples of distinct structures with the same polynomial. Nonetheless, the Jones polynomial classifies uniquely nearly all cases of current interest to biologists and has valuable properties lacking in the two-variable polynomials. In section 3, therefore, we discuss the features of the Jones polynomial. In section 4, we give a number of examples of how one computes both polynomials and discuss the ways in which these polynomials reflect fundamental topological properties of the forms with which they are associated. In section 5, we apply the new polynomials to site-specific recombination by the transposon Tn3 resolvase and phage λ integration systems and show how the products of these reactions can be rigorously predicted. Finally, in section 6 we provide for easy reference a Table of polynomials for many common knots and catenanes.

2. The Lickorish–Millett Three-state Polynomial and the Main Theorem

In this section we outline the main theorem on polynomial invariants associated with knots and catenanes that is given in the work of Lickorish &

F
arr
tan
shc
(<
Clo
(+
—
Mi
a c
a
pa
kn
co
cu
di
A
cr
pr
tw
ac

or
pe
cc
—
—
li
ci
(
n
S
n
a
c

s
i:
a
F
r
c
c
y

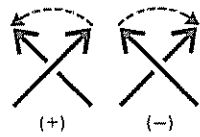


Figure 1. Sign convention for nodes. The crossed arrows represent segments of oriented curves (or unit tangent vectors of the curves), and the broken arrows show the direction the segment on top must be rotated ($<180^\circ$) to be congruent with the underlying segment. Clockwise and counterclockwise motion define (-) and (+) nodes, respectively.

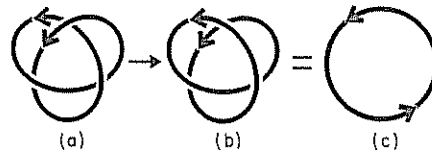


Figure 3. Segment passage. Segment passage is illustrated with a (-) trefoil. At the node at the top where the orientation arrows cross, the overlying segment is passed through the underlying one. This passage inverts the sign of the node and thereby unties the knot because all 3 nodes of a trefoil must have the same sign.

Millett (1987) A knot is defined, in mathematics, as a closed curve in three-dimensional space; therefore, a circle is a knot. To reduce confusion from this paradoxical usage we will call the circle the trivial knot or the unknot (unknotted knot). A second concept we need to introduce is that of the oriented curve. This is simply a curve along which a specific direction in which to travel has been chosen. A third needed concept is that of the node, or the crossing of segments of oriented curves in plane projection. Two oriented segments can cross in only two ways; the resulting nodes are given a sign according to the convention shown in Figure 1.

Lickorish & Millett have shown that with every oriented knot or catenane one can associate a polynomial in two variables, l and m , with integer coefficients. For example, the polynomials of three common structures shown in Figure 2 are $-2l^2 - l^4 + l^2 m^2$ for the three-noded knot (trefoil), $-l^{-2} - 1 - l^2 + m^2$ for the four-noded knot, and $lm^{-1} + l^3 m^{-1} - lm$ for the singly interlocked catenane. We will call the trefoil in Figure 2(a) the (-) trefoil because it has three (-) nodes and its mirror image with three (+) nodes, the (+) trefoil. Similarly, the catenane in Figure 2(c) with (-) nodes will be called the (-) singly linked catenane and its mirror image the (+) singly linked catenane.

The main theorem involving the association of such polynomials with knots and catenanes falls into two parts. The first part asserts that, associated with each oriented form there is a unique polynomial such that if two objects have polynomials that are different, in so much as a single coefficient, the two objects are topologically distinct. In the examples given, the occurrence of m^{-1} distinguishes the catenane from the knots and

the powers of the variable l in the two knots are different. Hence, all three examples are topologically distinct. The converse of the main theorem, that no two distinct forms have the same polynomial, is true for all prime knots containing up to nine nodes, and there are only a few exceptions for prime knots with 10 to 12 nodes. Little information is available for more complex knots.

The second part of the main theorem asserts that there is a straightforward algorithm for computing the polynomials that depends on two simple rules, and that any way of applying these rules to topologically equivalent forms will give precisely the same polynomial. The computational algorithm (Alexander, 1928) employs the method of changing a planar projection of an oriented knot or catenane into that of another knot or catenane whose polynomial is already known by a method that relates the polynomials. This is done by either segment passage or segment exchange. In passage, a segment is passed through a crossing segment, thereby changing the sign of the node. An example of a passage is shown in Figure 3, in which a trefoil is converted into a circle.

In segment exchange, a (+) or (-) node is converted to a (0) node, or *vice versa*. The (0) node is actually the deletion of a node. The generation of the (0) node is illustrated by Figure 4. A (+) node is approached along each of the crossing segments in the direction given by their orientation. When the node is reached, it is not crossed but one switches segments and proceeds in the direction specified by their orientation. If a (+) or (-) node interlocking the rings of a catenane is changed to a (0) node, the resulting object is a knot. This is due

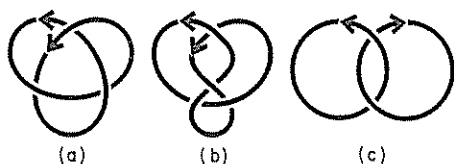


Figure 2. Three simple knots and catenanes. A plane projection is shown of (a) a (-) trefoil knot, (b) the 4-noded knot, and (c) the (-) singly interlocked catenane. The arrows indicate the orientation of the curves.

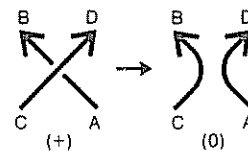


Figure 4. The zero node. The (0) node is a mathematical limiting case in which segments do not cross. The (0) node can be obtained from the (+) node shown by the following procedure. Start at C on one of the incoming branches and proceed to B on the other outgoing branch but avoid the crossing. Similarly, start at A on the other incoming branch and proceed to D.



Figure 5. Comparison of (+), (-), and (0) nodes. The arrows indicate the orientation of curves as for Fig 3

to the fact that the two separate rings have been joined to become one ring. If, on the other hand, the (+) or (-) node of a knot is changed to a (0) node, the resulting object is a catenane. The reverse is true also; namely, changing a (0) node formed by the two rings of a catenane into a (+) or (-) node changes the object to a knot; and changing a (0) node of a knot to a (+) or (-) node changes the object into a catenane.

Finally, we let K_+ , K_- and K_0 be planar representations of knots or catenanes that are exactly the same except in the vicinity of a (+), (-) or (0) node, respectively, where they have the configurations shown in Figure 5. We shall call this collection a state set, and it is illustrated below by the first three forms in Figure 7. We let $PK_+(l, m)$, $PK_-(l, m)$ and $PK_0(l, m)$ be the polynomials in l and m associated with the corresponding three states. The second part of the main theorem asserts that:

$$lPK_+(l, m) + l^{-1}PK_-(l, m) + mPK_0(l, m) = 0. \quad (1)$$

To calculate a polynomial, one needs to choose a normalization for the polynomial associated with the unknot, which we will designate U . Following the tradition of the Alexander polynomial, the polynomial associated with the unknot is defined to be equal to 1, the simplest non-zero polynomial. Sample calculations of polynomials using equation (1) are given in section 4, below

3. The Jones Polynomial and the Fourth State, K_∞

The Jones polynomial in one variable t , denoted $VK(t)$, may also be used to classify knots and catenanes. This polynomial has properties similar to that of $PK(l, m)$. Indeed, $VK(t)$ can be obtained from $PK(l, m)$ by setting $l = it^{-1}$ and $m = -i(t^{\frac{1}{2}} - t^{-\frac{1}{2}})$, where $i = \sqrt{-1}$. Thus, for example, $PK(l, m)$ for the (-) trefoil is:

$$-2l^2 - l^4 + l^2m^2,$$

and, hence, $VK(t)$ for this knot is:

$$\begin{aligned} -2(it^{-1})^2 - (it^{-1})^4 + (it^{-1})^2(-i(t^{\frac{1}{2}} - t^{-\frac{1}{2}}))^2 \\ = -t^{-4} + t^{-3} + t^{-1}. \end{aligned}$$

$VK(t)$ also satisfies the two properties of the main theorem of section 1. Thus, if any two objects have their polynomials $VK(t)$ different in any coefficient, they are topologically distinct. Furthermore, they satisfy a formula similar to equation (1). Making the substitution for l and m indicated above, we obtain:

$$t^{-1}VK_+(t) - tVK_-(t) - (t^{\frac{1}{2}} - t^{-\frac{1}{2}})VK_0(t) = 0 \quad (2)$$

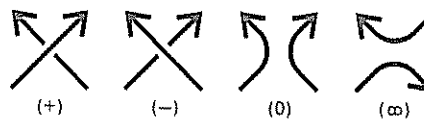


Figure 6. Comparison of the 4 types of nodes as they exist in a knot. The arrows represent the crossing segments of a knot and illustrate the (+), (-), (0) and (∞) nodes

Using this relationship and the fact that $VU(t) = 1$, i.e. the Jones polynomial for the trivial knot equals 1, one may derive the polynomial $VK(t)$ for any knot or catenane.

Because $VK(t)$ has only a single variable, it is easier to use for classification purposes than $PL(l, m)$. Another use of $VK(t)$ is to understand the structure of a fourth state, denoted K_∞ , which we show below is important in describing recombination enzymes. K_∞ is defined by the presence of the (∞) node which, like the (0) node, is actually a particular deletion of a node. Two separate cases of K_∞ need to be distinguished. In case 1, K_∞ is generated from a (+) node in K_+ , a knot, and in case 2, K_∞ is generated from a (+) node in a catenane.

In case 1, in which K_+ (and hence K_-) is a knot, the (∞) node is defined as shown in Figure 6. The top segment of the (∞) node is oriented from right to left and the bottom segment is oriented from left to right. One may choose the opposite orientation for each segment, because K_∞ is also a knot and the structure of knots, in general, does not depend on orientation. The orientation in Figure 6 is the one conventionally used. An example of the four states of a curve is shown in Figure 7; K_+ is the (+) trefoil, K_- is a trivial knot, K_0 is the (+) singly linked catenane, and K_∞ is a trivial knot. Note that K_∞ can be obtained by exchange from any of the other three states *only* if their orientation is violated; this is why K_∞ is not part of the state set described in section 2 for oriented curves.

If K_+ is a knot, K_0 is a catenane and hence has a linking number, $Lk(K_0)$, equal to the algebraic sum of the nodes interlocking the rings[†]. $Lk(K_0)$ and the

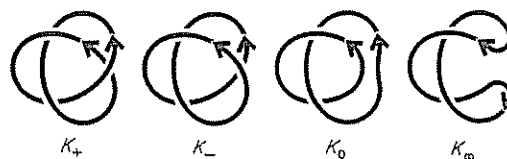


Figure 7. Illustration of a state set for the 4 node types. K_+ is a (+) trefoil and K_- is obtained by segment passage at the node where the orientation arrows cross. K_0 and K_∞ are obtained by segment exchange at the same node to give the forms shown in Fig 6

[†] The values for linking numbers in this paper are for single curves. They can be considered formally as the linking numbers of the axes of the DNA double helices. Thus, $Lk(K_0)$ is one-half the value of Ca as defined by Cozzarelli et al (1984).

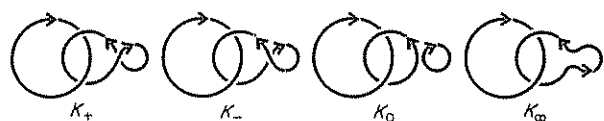


Figure 8. Formation of an (∞) node within 1 ring of a catenane. The state set shown is generated by conversion of the $(+)$ node made by the crossing arrows to the $(-)$, (0) and (∞) nodes as shown in Fig 6

polynomials $VK_+(t)$, $VK_-(t)$, and $VK_\infty(t)$ are related by the following equation (Lickorish, 1986):

$$VK_+(t) - tVK_-(t) - (1-t)t^{3Lk(K_0)}VK_\infty(t) = 0. \quad (3)$$

Thus, if $Lk(K_0)$ and the polynomials for K_+ and K_- are known, $VK_\infty(t)$ can be computed. Conversely, if K_∞ and $Lk(K_0)$ are known, then knowing the polynomials of K_+ will yield that of K_- , and vice versa.

Case 2, in which K_+ and K_- are dimeric catenanes, itself falls into two subcases, depending on the site of exchange. If this site is formed by a crossing of one of the catenane rings with itself, the (∞) crossing is defined just as in case 1. The K_0 state will have three rings, two of which are formed from the ring containing the exchange site. An example is shown in Figure 8. These two rings can be designated the left-hand and right-hand rings, depending on what side of the (0) node they are. If we define $Lk(K_0)$ in this case to be the linking number of the right-hand ring with the other two components, then the polynomial for K_∞ satisfies a similar equation to equation (3):

$$VK_+(t) - tVK_-(t) - (1-t)t^{3Lk(K_0)}VK_\infty(t) = 0 \quad (4)$$

Thus, for example, in Figure 8, K_+ is the $(+)$ singly linked catenane, K_- is topologically the same as K_+ , K_0 is a $(+)$ singly linked catenane plus a trivial knot, and K_∞ is the same as K_+ and K_- . The right-hand component of K_0 is a trivial knot that does not link the other two components, so $Lk(K_0) = 0$. Hence, the above equation becomes:

$$VK_+(t) - tVK_-(t) - (1-t)VK_+(t) = 0,$$

which is indeed correct, since the terms on the left-hand side cancel.

The second subcase of case 2 occurs when the exchange site is at the crossing of one ring of the catenane with the other. The (∞) crossing is now defined with both segments oriented in the same direction, left to right in our convention (Fig 9), and not the antiparallel orientation as in the other cases. In this subcase, the polynomial for K_∞ is easier to compute, for it depends only on K_+ and K_- . If $Lk(K_+)$ equals the linking number of K_+ , then:

$$VK_+(t) - tVK_-(t) - (1-t)^{3(Lk(K_+) - 1)}VK_\infty(t) = 0. \quad (5)$$

This remarkable formula shows that K_∞ is determined by K_+ and K_- only. Several examples of this subcase are given in the following section.

The proper drawing of the (∞) node in each

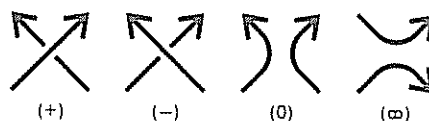


Figure 9. Comparison of the 4 types of nodes generated from a $(+)$ node interlocking the 2 rings of a catenane. The arrows of the $(+)$ node represent oriented crossing segments from each ring of a dimeric catenane, and the $(-)$, (0) and (∞) nodes are generated by segment passage and segment exchange. Note that the $(+)$, $(-)$ and (0) nodes are exactly the same as in knots (Fig 6), but that the upper limb of the (∞) node points in the opposite direction.

instance can be obtained by the following procedure. Start with a $(+)$ node as shown in Figure 4. Draw the lower limb of the (∞) node starting from C and proceeding up toward the crossing but then coming back down to A. The orientation of K_∞ is not set. Next draw the upper limb similarly but proceed from B or D depending on orientation.

4. $PK(l, m)$ and $VK(t)$ Polynomials of Special Examples

In this section we shall show how to use the theorems of the previous sections to find polynomials of some well-known catenanes and knots, thereby illustrating how these polynomials can be computed easily and what types of information are contained in their nature.

As our first example, we find the polynomials of two copies of the trivial knot. If we set this equal to K_0 , the four states shown in Figure 10 are obtained. Because K_+ , K_- and K_∞ are all trivial knots, their polynomials are all equal to 1. We can then calculate the two-variable polynomial for K_0 using equation (1):

$$l(1) + l^{-1}(1) + mPK_0(l, m) = 0,$$

$$PK_0(l, m) = -m^{-1}(l^{-1} + l)$$

Equation (2) gives us the corresponding Jones polynomial:

$$t^{-1}(1) - t(1) + (t^{\frac{1}{2}} - t^{-\frac{1}{2}})VK_0(t) = 0,$$

$$VK_0(t) = -(t^{\frac{1}{2}} + t^{-\frac{1}{2}}).$$

If we set μ equal to $-m^{-1}(l^{-1} + l)$ or $-(t^{\frac{1}{2}} + t^{-\frac{1}{2}})$, it is easy to see by iteration of the above method that polynomials of c copies of the unknot are equal to μ^{c-1} . An important advantage of polynomial invariants is that they can be calculated readily for catenanes containing more than two rings, and such complex forms are found in nature; for example, in kinetoplasts (Englund *et al.*, 1982).



Figure 10. State set for 2 copies of the unknot. K_0 is 2 copies of the unknot.

We next verify equation (3) for the computation of $VK_\infty(t)$. Because the two curves are unlinked, $Lk(K_0) = 0$. Thus, equation (3) becomes:

$$1 - t - (1 - t)VK_\infty(t) = 0,$$

which shows, as expected, that $VK_\infty(t) = 1$.

The most commonly found catenanes are regularly interlocked. Known technically as members of the torus family (Rolfsen, 1976), these can be described analogously to the classes of DNA double helices, because the two rings helically intertwine. The helix can be right-handed or left-handed, and the orientation of the two rings can be parallel or antiparallel. The resulting four classes are right-handed parallel, left-handed parallel, right-handed antiparallel, and left-handed antiparallel. Examples of the left-handed forms are shown in Figure 15. For the special case of singly interlocked catenanes, the first and the fourth classes are equivalent, as are the second and third; these have been designated as (+) (Fig 11) and (-) (Fig 12), respectively, according to the sign of the nodes. Our next examples arise from these classes of catenanes. They will show how chirality and orientation affect the polynomials.

(a) *The singly interlocked right-handed parallel catenane, R_p*

To compute the polynomials for this catenane, we use the state set shown in Figure 11. K_- is two copies of the trivial knot and its polynomial is therefore μ . Since K_0 is the trivial knot, its polynomial is 1. Therefore, applying equations (1) and (2), we obtain the polynomials of R_p :

$$\begin{aligned} PR_p(l, m) &= PK_+(l, m) = -l^{-1}(l^{-1}\mu + m) \\ &= (l^{-3} + l^{-1})m^{-1} - l^{-1}m, \\ VR_p(t) &= VK_+(t) = t[t\mu + (t^{\frac{1}{2}} - t^{-\frac{1}{2}})] \\ &= -t^{\frac{1}{2}} - t^{\frac{3}{2}} \end{aligned}$$

We next verify equation (5) for $VK_\infty(t)$, an unknot. In this case, the linking number of K_+ equals +1. Hence, equation (5) yields:

$$\begin{aligned} -t^{\frac{1}{2}} - t^{\frac{3}{2}} - t[-(t^{\frac{1}{2}} + t^{-\frac{1}{2}})] \\ - (1 - t)t^{3(1 - \frac{1}{2})}VK_\infty(t) = 0, \\ -t^{\frac{1}{2}} + t^{\frac{3}{2}} - (t^{\frac{1}{2}} - t^{\frac{3}{2}})VK_\infty(t) = 0, \end{aligned}$$

which shows, as expected, that $VK_\infty(t) = 1$

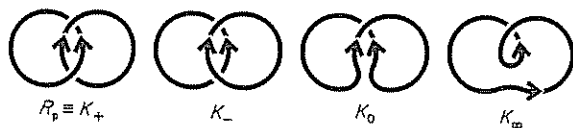


Figure 11. State set for the right-handed parallel singly linked catenane. This catenane, R_p , is the K_+ state and is interlocked by 2 (+) nodes. Handedness is defined by the sense of the helical wrapping of one curve around the other, and parallel by the relative orientation of adjacent segments from the 2 rings. The rest of the state set is generated as for Fig 9

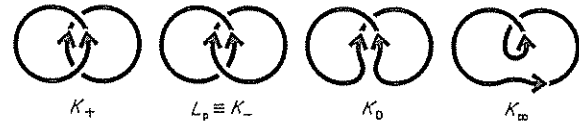


Figure 12. State set for the left-handed parallel singly linked catenane. This catenane, L_p , is the K_- state and is interlocked by 2 (-) nodes. The rest of the state set is generated as for Fig 9.

(b) *The singly interlocked left-handed parallel catenane, L_p*

To compute the polynomials for this catenane, we use the state set shown in Figure 12, where $K_- = L_p$. Clearly, K_+ is two copies of the trivial knot, and K_0 and K_∞ are just the trivial knot. Hence, employing equations (1) and (2):

$$\begin{aligned} PL_p(l, m) &= PK_-(l, m) = -l(l\mu + m) \\ &= (l + l^3)m^{-1} - lm, \text{ and} \\ VL_p(t) &= VK_-(t) = t^{-1}[t^{-1}\mu - (t^{\frac{1}{2}} - t^{-\frac{1}{2}})] \\ &= -t^{-\frac{1}{2}} - t^{-\frac{3}{2}} \end{aligned}$$

Note that only positive powers of the variable l occur in the two-variable polynomial for the left-handed catenane, whereas there are only negative powers of l in the polynomials for the right-handed catenane. In the one-variable polynomial for the left-handed catenane, only negative powers of t occur, whereas there are only positive values of t in the right-handed catenane. This brings us to the following fundamental corollary of the main theorem (Lickorish & Millett, 1987). If \bar{K} denotes the mirror image of K , then the polynomials of K and \bar{K} are related by the equation: $P\bar{K}(l, m) = PK(l^{-1}, m)$ and $V\bar{K}(t) = VK(t^{-1})$. For example, $L_p = \bar{R}_p$ and, by our calculations above, $PL_p(l, m) = PR_p(l^{-1}, m)$, and $VL_p(t) = VR_p(t^{-1})$, verifying the assertion of the corollary in this special case. Thus, we conclude that for a knot or catenane to be topologically equivalent to its mirror image, a property known as amphichirality, the associated polynomial must be symmetric in l and l^{-1} or t and t^{-1} , i.e. the polynomials must be left unchanged if we reverse the signs of the exponents of the l or the t terms. For example, the polynomials of the amphichiral four-noded knot, shown in Figure 2(b), are:

$$\begin{aligned} -l^{-2} - 1 - l^2 + m^2, \text{ and} \\ t^{-2} - t^{-1} + 1 - t + t^2. \end{aligned}$$

(c) *The positive trefoil knot, T_+*

To compute the polynomials for the positive trefoil, T_+ , we use the state set shown in Figure 7. K_- and K_∞ are the unknot and K_0 is R_p . Using equations (1) and (2), and our results for R_p , we obtain the polynomials of T_+ :

$$\begin{aligned} PT_+(l, m) &= -l^{-2} - l^{-1}(ml^{-3}m^{-1} \\ &\quad + ml^{-1}m^{-1} - ml^{-1}m) \\ &= -l^{-4} - 2l^{-2} + l^{-2}m^2 \end{aligned}$$

$$VT_+(t) = t^2 + t(t^{\frac{1}{2}} - t^{-\frac{1}{2}})(-t^{\frac{1}{2}} - t^{\frac{1}{2}}) \\ = -t^4 + t^3 + t$$

(d) *The negative trefoil, T_-*

The polynomials for the negative trefoil T_- are easy to compute from those for T_+ , because T_- is the mirror image of T_+ . Thus, we need only interchange l with l^{-1} and t with t^{-1} . Hence:

$$PT_-(l, m) = -l^4 - 2l^2 + l^2m^2, \\ VT_-(t) = -t^{-4} + t^{-3} + t^{-1}$$

We note that, since the polynomials of T_+ and T_- are distinct, T_+ and T_- are not equivalent and hence the trefoil is not amphichiral

(e) *The doubly interlocked right-handed parallel catenane, RD_p*

To compute the polynomials for the doubly interlocked right-handed parallel catenane, we use the state set shown in Figure 13. We observe that K_- is the singly linked right-handed parallel catenane, R_p , and that K_0 is the positive trefoil, T_+ . Thus, the polynomials of RD_p are:

$$PRD_p(l, m) = PK_+(l, m) = (-l^{-5} - l^{-3})m^{-1} \\ + (l^{-5} + 3l^{-3})m - l^{-3}m^3, \\ VRD_p(t) = VK_+(t) = -t^4 + t^{\frac{1}{2}} - t^{\frac{1}{2}} - t^{\frac{1}{2}}$$

Note that all the powers of l are negative and those of t are positive.

(f) *The doubly interlocked right-handed antiparallel catenane, RD_{ap}*

To compute the polynomial for the doubly interlocked right-handed antiparallel catenane, RD_{ap} , we use the state set shown in Figure 14. We observe that K_+ is the singly linked left-handed parallel catenane, L_p , K_0 is the trivial knot, and that K_∞ is the positive trefoil. Using equations (1) and (2), we obtain:

$$PRD_{ap}(l, m) = PK_-(l, m) \\ = (-l^3 - l^5)m^{-1} + (-l + l^3)m,$$

and

$$VRD_{ap}(t) = VK_-(t) = -t^{-\frac{1}{2}} - t^{-\frac{1}{2}} + t^{-\frac{1}{2}} - t^{-\frac{1}{2}}$$

Note that the powers of l are all positive and those of t are all negative, even though the catenane

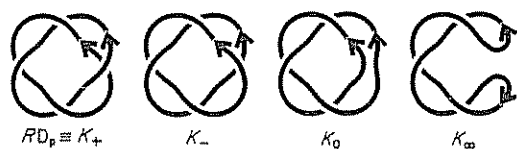


Figure 13. State set for the right-handed parallel doubly linked catenane. This catenane, RD_p , is the K_+ state and is interlocked by 4 (+) nodes. The rest of the state set is constructed as shown for Fig 9

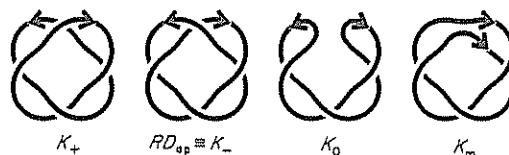


Figure 14. State set for the right-handed antiparallel doubly linked catenane. This catenane, RD_{ap} , is the K_- state and is interlocked by 4 (-) nodes. The remainder of the state set is generated as shown for Fig 9

is right-handed. This results from the antiparallel property of the catenane. This situation is analogous to that of the sign of the linking number between catenated rings; the change of either handedness or relative orientation changes sign, whereas the change of both maintains it.

We next verify equation (5) for $K_\infty(t)$. In this case, $Lk(K_+) = -1$. Thus, equation (5) yields:

$$(-t^{-\frac{1}{2}} - t^{-\frac{1}{2}}) - t(-t^{-\frac{1}{2}} - t^{-\frac{1}{2}} + t^{-\frac{1}{2}} - t^{-\frac{1}{2}}) \\ - (1-t)t^{-\frac{1}{2}}VK_\infty(t) = 0$$

and $VK_\infty(t)$ equals $-t^4 + t^3 + t$, the polynomial for the positive trefoil, T_+ .

(g) *The doubly linked left-handed parallel catenane, LD_p*

The polynomials for LD_p , pictured in Figure 15(a), can be obtained from the polynomial of RD_p by using the fact that LD_p is the mirror image of RD_p . Thus, we need only interchange l^{-1} for l in the $PRD_p(l, m)$ and t for t^{-1} in $VRD_p(t)$. We find, then, that:

$$PLD_p(l, m) = (-l^5 - l^3)m^{-1} + (l^5 + 3l^3)m - l^3m^3$$

and

$$VLD_p(t) = -t^{-4} + t^{-\frac{1}{2}} - t^{-\frac{1}{2}} - t^{-\frac{1}{2}}$$

(h) *The doubly interlinked left-handed antiparallel catenane, LD_{ap}*

The polynomials for LD_{ap} , pictured in Figure 15(b), can be obtained from the polynomials of RD_{ap} by using the fact that LD_{ap} is the mirror image of RD_{ap} . Thus:

$$PLD_{ap}(l, m) = (-l^{-3} - l^{-5})m^{-1} + (-l^{-1} + l^{-3})m,$$

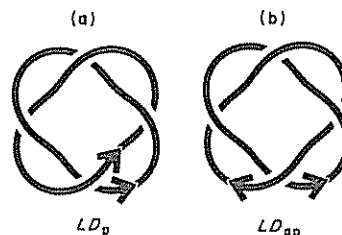


Figure 15. The 2 types of left-handed doubly linked catenanes (a) The doubly interlinked left-handed parallel catenane, LD_p , (b) the left-handed antiparallel catenane, LD_{ap} . LD_p , like RD_{ap} , is interlocked by 4 (-) nodes and LD_{ap} , like RD_p , is interlocked by 4 (+) nodes, but all 4 stereoisomers are distinct.

and

$$VLD_{ap}(t) = -t^3 - t^2 + t^1 - t^0$$

(i) *The four-crossing knot, F*

We compute the polynomials of the four-crossing knot using the state set shown in Figure 16. Clearly, K_- is the trivial knot, and K_0 is the singly interlocked left-handed parallel catenane, L_p . Thus, using equations (1) and (2) we obtain:

$$PF(l, m) = PK_+(l, m) = -l^{-2} - 1 - l^2 + m^2,$$

$$VF(t) = VK_+(t) = t^{-2} - t^{-1} + 1 - t + t^2.$$

(j) *Connected unions of knots and catenanes*

The processes of recombination and topoisomerase action can generate compound knots and catenanes (Wasserman & Cozzarelli, 1986). By iteration of the general formula, one may calculate the polynomials associated with a union of any number of knots or catenanes. This is true because the calculation is done for one of the constituents before applying the algorithm to another. The result is that one may simply factor out the contribution of the other constituents and continue the process.

We denote the two separate configurations, either knots or catenanes, K_1 and K_2 ; as examples, we show a (+) trefoil and an antiparallel right-handed doubly linked catenane (Fig 17). Next, one takes a small segment of either K_1 or K_2 and stretches it until it is contiguous to a segment of the other curve. These segments are broken and rejoined as shown in Figure 17. We denote this connected union by $K_1 \# K_2$. Its polynomial is (Lickorish & Millett, 1987):

$$P(K_1 \# K_2)(l, m) = PK_1(l, m)PK_2(l, m)$$

We compute similarly the polynomials associated with the "square" (SQ) and "granny" (GR) knots shown in Figure 18. The former is the union of a (+) trefoil with a (-) trefoil, and we use the union of two (+) trefoils as an example of the latter:

$$\begin{aligned} PSQ(l, m) &= PT_-(l, m)PT_+(l, m) \\ &= (-2l^2 - l^4 + l^2m^2)(-l^{-4} - 2l^{-2} + l^{-2}m^2) \\ &= (2l^{-2} + 5 + 2l^2) + (-l^{-2} - 4 - l^2)m^2 + m^4, \end{aligned}$$

whereas:

$$\begin{aligned} PGR(l, m) &= PT_+(l, m)PT_+(l, m) \\ &= (-2l^2 - l^4 + l^2m^2)^2 \\ &= (4l^4 + 4l^6 + l^8) + (-4l^4 - 2l^6)m^2 + l^4m^4. \end{aligned}$$

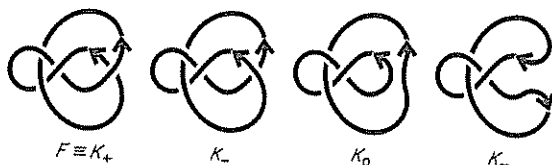


Figure 16. State set for the 4-noded knot. This knot is the K_+ state, and the remainder of the state set is generated as shown for Fig 6

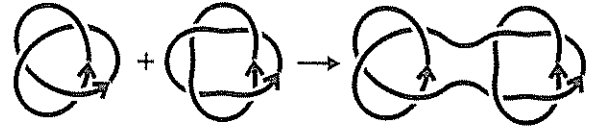


Figure 17. Joining 2 prime curves to form a compound curve. A (+) trefoil and an antiparallel right-handed doubly linked catenane are joined together by pulling out a segment of one curve until it is adjacent to a segment of the other and then exchanging the segments in a way that respects the orientation of each curve.

Clearly $PSQ(l, m) \neq PGR(l, m)$, and therefore the square and granny knots are not topologically equivalent. Note that $PSQ(l, m)$ is symmetric in l and l^{-1} ; thus the square knot could be amphichiral. Indeed, this can be shown by the identity of the mirror image of SQ (formed by reversing all the crossings) and a planar 180° rotation of SQ .

(k) *Summary and observations*

We close this section by reviewing and making additional observations. We showed that the positive and negative trefoils T_+ and T_- are topologically distinct but that the four-noded knot F is amphichiral. Analyses of the polynomials of RD_p, RD_{ap}, LD_p and LD_{ap} demonstrated that there are four topologically distinct doubly interlocked catenanes, a fact we proved earlier using Schubert's theory (White & Cozzarelli, 1984).

There are several main observations to be drawn from these examples. First, the polynomials of a mirror image of a catenane or knot can be computed from those of the catenane or knot by interchanging l and l^{-1} or t and t^{-1} . Hence, if a knot or catenane is amphichiral, its polynomial is symmetric in l and l^{-1} or t and t^{-1} .

A second observation about the polynomial $VK(t)$ we illustrated with doubly interlinked catenanes. The catenane RD_{ap} can be obtained from the catenane RD_p by simply reversing the orientation of one of its component rings. We recall that:

$$VRD_p(t) = -t^2 + t^3 - t^1 - t^0$$

and

$$VRD_{ap}(t) = -t^{-2} - t^{-1} + t^1 - t^0$$

Therefore,

$$VRD_{ap}(t) = t^{-6}VRD_p(t)$$

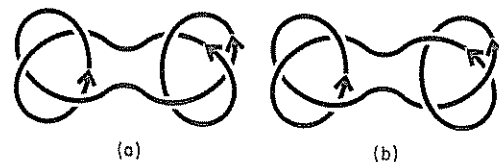


Figure 18. Two common compound knots. (a) The square knot is the union of a (+) and (-) trefoil whereas (b) the granny knot shown is the union of 2 (+) trefoils. The mirror-image of the latter would be a granny knot with only (-) nodes, whereas the square knot is amphichiral.

Now, the linking number of RD_p is $+2$, so that one can write this relationship as (Lickorish, 1986):

$$VRD_{ap}(t) = t^{-3Lk(RD_p)} VRD_p(t)$$

This formula is an example of a much more general formula. Let K be any catenane and \hat{K} be the same catenane, except that one of its rings has the reverse orientation. Then the polynomials $VK(t)$ and $V\hat{K}(t)$ satisfy the equation (Lickorish & Millett, 1986):

$$V\hat{K}(t) = -t^{-3Lk(K)} VK(t),$$

where $Lk(K)$ is the linking number of K . This important formula states that K and \hat{K} are topologically distinct when $Lk(K) \neq 0$. Furthermore, if one combines this result with the earlier statement about mirror images, one can arrive at the following conclusion. Let K and \hat{K} be as above. Let \tilde{K} be the mirror image of K and $\tilde{\hat{K}}$ be the mirror image of \hat{K} , the so-called mirror reversed catenane. Then the polynomials of \tilde{K} , $\tilde{\hat{K}}$ and \hat{K} can all be computed from the polynomial of K :

$$V\tilde{K}(t) = t^{-3Lk(K)} VK(t),$$

$$V\tilde{\hat{K}}(t) = VK(t^{-1}), \quad (6)$$

$$V\hat{K}(t) = V\tilde{\hat{K}}(t^{-1}) = t^{3Lk(K)} VK(t^{-1}).$$

Finally, we observe that the $Lk(K)$ is easy to compute by counting nodes. However, it is also possible to find $Lk(K)$ from the polynomial of K , as detailed elsewhere (Lickorish & Millett, 1987). Hence, all information necessary to compute $V\tilde{K}(t)$, $V\tilde{\hat{K}}(t)$ and $V\hat{K}(t)$ is contained in $VK(t)$. There is no similar set of relationships for the two-variable polynomials.

In this section we have shown how the polynomials of a complex knot or catenane can be computed from less complicated ones by means of equations (1) and (2). We used the polynomial of the unknot to compute the polynomials associated to multiple copies of the unknot. We used these results to compute the polynomials associated with trefoils and singly linked catenanes. In turn, we used these polynomials to compute those of the doubly interlocked catenanes and the four-crossing knot. In practice, this recursive procedure is even easier to use because the polynomials of many prime knots or catenanes can be found in mathematical tables. The most extensive are by Thistlethwaite (1985, 1987). These papers assume a knowledge of topology, and for the convenience of the reader we have in section 6 prepared a Table of invariants for the simplest prime knots and catenanes, and a standard drawing of these curves.

5. Applications to DNA

The polynomial method provides a rigorous and logical way for classifying DNA knots and catenanes, which is essential for the orderly description of these forms and analysis of their cellular role. For example, several cellular processes produce DNA catenanes or knots that belong to the

torus family. The polynomials allow one to prove that for this family there are four, and only four, stereoisomers for catenanes containing more than four or more nodes, but only two isomers of knots and singly linked catenanes. These distinctions are important because different cellular processes generate different types of torus forms (Wasserman & Cozzarelli, 1986). The catenanes produced by DNA replication are expected to be right-handed and parallel, but those produced by recombination by enzymes such as Int are also right-handed but antiparallel. The knots made by Int are right-handed, whereas those tied by topoisomerase I of *Escherichia coli* are equally likely to be right-handed as left-handed.

The attractiveness of the polynomials in describing the structure of DNA is perhaps even more apparent in its description of the actions of topoisomerases and, particularly, recombination enzymes. The type 2 topoisomerases are enzymes that promote a double-stranded DNA passage, i.e. the passage of the axis of a double helix through itself (Wang, 1985). Similarly, these enzymes can pass the axis of one DNA through the axis of another. The first operation can tie or untie knots and the second can catenate or decatenate rings. The fundamental question is, given a particular knot or catenane, what is the structure of the product DNA after a double-stranded passage? This question is answered, at least in part, by the main theorems that relate the substrate DNA to the product DNA in terms of another DNA, in which a (+) or (-) node is replaced by a (0) or (∞) node.

This can be illustrated using the torus DNA catenanes. A right-handed parallel torus catenane that intertwines n times has $2n$ (+) nodes. A strand passage at one of these nodes changes it to a (-) node, and the resulting product is topologically the same as a right-handed parallel catenane that intertwines $(n-1)$ times. In this case, K_+ is the catenane intertwined n times and K_- is the catenane intertwined $(n-1)$ times. The state set in which $n=2$ is shown in Figure 7. The same Figure shows that K_∞ is the trivial knot. In fact, no matter what value n assumes, K_∞ is always the unknot. Thus, if the Jones polynomial of the $(n-1)$ helically intertwined catenane is known, one can use equation (5) to compute the polynomial of the n -catenane, since $VK_\infty(t) = 1$. By recursion, one can find the polynomial of the n -catenane, knowing only the polynomial of the singly interlinked catenane (or of the 2 copies of the unknot). Then equations (6) can be used to generate the polynomial for the three stereoisomers of the right-handed parallel n -catenane. In summary, the Jones polynomial can be used to classify all torus DNA catenanes knowing only the polynomial of the singly interlinked catenane.

The second basic operation that can change the structure of DNA is exchange in which two pieces of DNA are broken and the ends switched. This is the process of interconverting a (0) or (∞) node and a (+) or (-) node. Once again, the fundamental

question is: given a particular DNA knot or catenane, what is the end product after exchange? In order to show the utility of the new polynomial theories in answering this question, we develop two specific applications.

The first will show how one can predict all of the products of successive reactions of Tn3 resolvase with only a knowledge of the first product. A complete exposition of the scheme for resolvase-mediated recombination has been given (Cozzarelli *et al.*, 1984; Wasserman *et al.*, 1985). The conclusion was that after formation of the synaptic intermediate in which the sites to be recombined are aligned on the enzyme, the process of exchange usually occurs once but is occasionally repeated. Determination of the structure of the knotted and catenated products of these extra rounds of recombination was critical in elucidating the mechanism of the reaction. The resolvase synaptic intermediate has three essential (-) nodes and each round of recombination introduces a (+) node†. In Figure 19, the scheme is shown for three rounds of recombination.

Let S be the substrate, and K_1 , K_2 and K_3 be the first, second and third recombination products, respectively. The resolvase mechanism generates a series of (+) nodes along a contiguous segment of the DNA, one for each round of recombination (Fig. 19). If K_n is the product after n rounds, we will show that K_n is determined completely by S and K_1 .

We consider the state set for the second round of recombination in which the (0) state is K_1 , and therefore the (+) state is K_2 (Fig. 20). We note that K_- is topologically equivalent to the original substrate, S . The polynomials of these molecules are related by the equation:

$$lPK_+(l, m) + l^{-1}PK_-(l, m) + mPK_0(l, m) = 0,$$

or:

$$lPK_2(l, m) + l^{-1}PS(l, m) + mPK_0(l, m) = 0,$$

since $PK_-(l, m) = PS(l, m)$. Therefore, one can

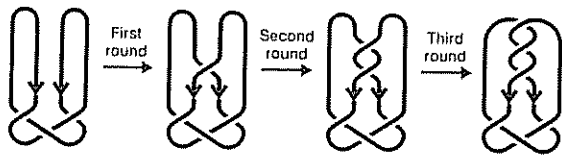


Figure 19. Three successive rounds of recombination by resolvase. The curve on the left represents the resolvase substrate with the 2 recombination sites (arrows) synapsed in parallel and 3 (-) supercoils trapped in the synaptic complex. In each of the 3 successive recombination rounds shown, exchange at the arrows generates a (+) node. The products are, successively, the (-) singly linked catenane, the 4-noded knot, and the (+) figure-8 catenane.

† This treatment neglects the other changes in DNA by recombination that do not affect product knot or catenane structure (Cozzarelli *et al.*, 1984; Benjamin & Cozzarelli, 1986).

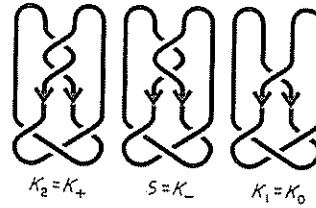


Figure 20. State set for the 2nd round of iterative recombination by resolvase. In the 2nd round of recombination by resolvase shown in Fig. 19, K_1 , the (-) singly linked catenane, is converted to K_2 , the 4-noded knot. In the state set shown, K_1 is K_0 , K_2 is K_+ , and K_- is the substrate for the 1st round of recombination, the unknot with 3 (-) supercoils.

write the polynomial for K_2 in terms of the polynomials of S and K_1 . In fact, one obtains:

$$PK_2(l, m) = l^{-1}(-mPK_1(l, m) - l^{-1}PS(l, m))$$

Hence, the topology of K_2 is known in terms of the topology of K_1 and S .

Next, we consider the state set for the third round in which the (0) state is K_2 , the (+) state is the third-round product K_3 , and K_- is topologically equivalent to K_1 , the first-round product (Fig. 21). Their polynomials are related by the equation:

$$lPK_+(l, m) + l^{-1}PK_-(l, m) + mPK_0(l, m) = 0,$$

or:

$$lPK_3(l, m) + l^{-1}PK_1(l, m) + mPK_0(l, m) = 0$$

Therefore, one can write $PK_3(l, m)$ in terms of $PK_1(l, m)$ and $PK_2(l, m)$ and get:

$$PK_3(l, m) = l^{-1}(-mPK_2(l, m) - l^{-1}PK_1(l, m))$$

But $PK_2(l, m)$ is known in terms of $PK_1(l, m)$ and $PS(l, m)$ and hence:

$$\begin{aligned} PK_3(l, m) &= l^{-1}[-ml^{-1}(-mPK_1(l, m) \\ &\quad - l^{-1}PS(l, m)) - l^{-1}PK_1(l, m)] \\ &= (+m^2l^{-2} - l^{-2})PK_1(l, m) + ml^{-2}PS(l, m) \end{aligned}$$

Hence, the topology of K_3 is known in terms of the topology of S and K_1 . If this procedure is repeated, one obtains that K_n is known if K_1 and S are known.

Matrix notation is very helpful in writing down the general result†. The two equations:

$$\begin{aligned} PK_1 &= 0PS + 1PK_1 \\ PK_2 &= -l^2PK_0 - l^{-1}mPK_1 \end{aligned}$$

† The matrix notation:

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

means that

$$\begin{aligned} c_1 &= a_{11}b_1 + a_{12}b_2 \\ c_2 &= a_{21}b_1 + a_{22}b_2, \end{aligned}$$

where i and j in a_{ij} refer to the row and column, respectively, of the matrix elements.

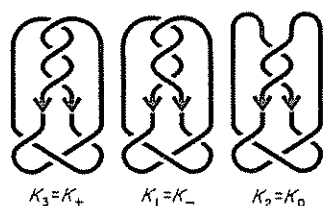


Figure 21. State set for the 3rd round of iterative recombination by resolvase. In the 3rd round of recombination by resolvase, shown in Fig 19, K_2 , the 4-noded knot, is converted to the (+) figure-8 catenane. In the state set shown, K_2 is K_0 , K_3 is K_+ , and K_1 is K_- .

in matrix notation can be written:

$$\begin{bmatrix} PK_1 \\ PK_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -l^{-2} & -l^{-1}m \end{bmatrix} \begin{bmatrix} PS \\ PK_1 \end{bmatrix}$$

where we have abbreviated $PK(l, m)$ as PK

The second set of equations:

$$PK_2 = 0PK_1 + 1PK_2$$

$$PK_3 = -l^{-2}PK_1 - l^{-1}mPK_2$$

can be expressed:

$$\begin{bmatrix} PK_2 \\ PK_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -l^{-2} & -l^{-1}m \end{bmatrix} \begin{bmatrix} PK_1 \\ PK_2 \end{bmatrix}$$

Combining equations,† we obtain:

$$\begin{bmatrix} PK_2 \\ PK_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -l^{-2} & -l^{-1}m \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -l^{-2} & -l^{-1}m \end{bmatrix} \begin{bmatrix} PS \\ PK_1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 \\ -l^{-2} & -l^{-1}m \end{bmatrix}^2 \begin{bmatrix} PS \\ PK_1 \end{bmatrix}$$

After n such reiterations, we obtain:

$$\begin{bmatrix} PK_n \\ PK_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -l^{-2} & -l^{-1}m \end{bmatrix}^n \begin{bmatrix} PS \\ PK_1 \end{bmatrix}$$

By making the appropriate substitutions, $l = it^{-1}$ and $m = -i(t^{\frac{1}{2}} - t^{-\frac{1}{2}})$, one can also write the general relationship in terms of the $VK(t)$ polynomial. In fact, $-l^{-2} = t^2$ and $-l^{-1}m = i(t^{\frac{1}{2}} - t^{-\frac{1}{2}})$, so that:

$$\begin{bmatrix} VK_n \\ VK_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ t^2 & i(t^{\frac{1}{2}} - t^{-\frac{1}{2}}) \end{bmatrix}^n \begin{bmatrix} VS \\ VK_1 \end{bmatrix}$$

Thus, $VK_n(t)$ is known in terms of $VK_0(t)$ and $VK_1(t)$

† The multiplication of the 2 matrices:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ and } \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

yields a matrix:

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

where $c_{ij} = a_{11}b_{1j} + a_{12}b_{2j}$

There is another property of this type of reaction. Starting with the original substrate S as the trivial knot, K_2, K_4, \dots, K_{2m} for any positive integer m are knots, whereas $K_1, K_3, K_5, \dots, K_{2m-1}$ are catenanes. One can, *via* the polynomial theory, determine all the products of the resolvase reaction if one knows any two successive products, or any two successive knots or catenanes. This can be seen from the general equation:

$$PK_n = -l^{-2}PK_{n-2} - l^{-1}mPK_{n-1}$$

For, given any two of the three polynomials, one can solve for the third, and by reiteration find the polynomials for any K in the sequence. Finally, and most generally, if one knows *any two* products of the resolvase reaction and their place in the rounds of recombination, or *the substrate* and *any one* product, then one can determine all products by solving the simultaneous equation (1) for all rounds between the two products.

The chief importance of this result is the mathematical rigor it brings to the study of recombination. For the first time the products of a particular mechanism of recombination can be predicted mathematically. The result can be generalized to any process that generates a series of (+) or (-) nodes along a continuous segment of DNA. From our earlier work we could predict, using mathematics alone, only the number of nodes in a product as a function of mechanism (Cozzarelli *et al.*, 1984). Moreover, previously one could never be sure that the same series of products could not be generated by a different mechanism. Now, in the case of resolvase, given the assumption of iteration of a single exchange mechanism, it is easy to show that the scheme in Figure 19 is unique and thus that topologically the resolvase system must have the postulated synaptic structure and exchange mechanism.

There are also two practical consequences. First, some limitations to the use of physical models for DNA to test recombination mechanisms are avoided by using predictions based on mathematics. It is easy to make errors in manipulating models. Often, strand rotations are introduced accidentally and it is hard to identify the structure of complex products. Recall that there are about 10^8 different 19-noded knots. Moreover, one is often not quite sure what assumptions have been made or are necessary. Thus, it can be difficult to convince anyone by an argument based solely on manipulation of a physical model or even to communicate it. Second, it is much harder experimentally to determine the topology of a catenane than a knot because it is necessary to determine the orientation of the two rings as well as the path of the DNA (Wasserman & Cozzarelli, 1984). Using polynomials, the recombination system can be specified solely from the structure of the knotted products without knowing the intermediate catenanes.

The second application of the polynomial invariants that we present in detail concerns a different type of site-specific recombination enzyme

and the utility of the fourth state, K_∞ . We must first distinguish local and global orientation. Global orientation has been defined as the choice of a direction in which to travel along a DNA circle. Once the global choice is made, it cannot be altered partway around the circle, and thus all segments of the ring have the same head-to-tail global orientation. Local orientation is the choice of a direction in which to travel along a small segment of the DNA circle, and thus local orientation need not be consistent with global orientation. Indeed, all site-specific recombinases recognize local orientation of the recombination sites, but only some, such as resolvase, respond to global orientation. Others, notably the bacteriophage λ Int system, recombine sites whose relative local orientation can either be head-to-tail (direct) (Fig 24) or head-to-head (inverse) (Fig 22; and see Nash, 1981). One of the inverse sites must have a local orientation opposite to the global orientation. The two-variable polynomial uses three states that are global orientation-specific. However, the Jones polynomial can be written in terms of the (∞) node, which, by definition, has no specific global orientation. This is exactly the desired mathematical property for enzyme systems such as Int that recombine sites without regard to global orientation.

To illustrate, the inversely repeated sites for Int are denoted by open arrows in Figure 22 and the global orientation is given by the filled arrows. Arbitrarily, four $(-)$ supercoils are shown trapped between the sites but this number varies (Spengler *et al.*, 1985). Like resolvase, Int aligns the sites in parallel locally (Griffith & Nash, 1985), with the result that an (∞) node is created. We note that the substrate DNA just before exchange (the synaptic intermediate) is a plectonemically interwound unknot. The parallel alignment of the inverse sites necessitates that the number of supercoil nodes trapped between the sites must be even, say $2m$, where m is an integer (Cozzarelli *et al.*, 1984). Recombination introduces a new $(+)$ node (Griffith & Nash, 1985) and changes the sign of the entrapped nodes, yielding a knot with $2m+1$ $(+)$ nodes (Cozzarelli *et al.*, 1984).

The new result is that we can obtain the structure of the knot by calculating its Jones polynomial by a simple recursion method, knowing



Figure 22. Recombination by Int between inversely repeated sites. The global orientation of the substrate is given by the filled arrows and the inversely repeated recombination sites by the open arrows. Recombination introduces a single $(+)$ node at the recombination sites and inverts the sequence between the sites. Because of this inversion, the 4 $(-)$ supercoil nodes in the substrate are converted to $(+)$ nodes. The product is then a knot of the torus family with 5 $(+)$ nodes.

only the number of trapped supercoil nodes in the synaptic intermediate substrate. The method is illustrated by the state sets in Figures 7 and 23. K_∞ is the substrate, and K_+ is the product of recombination. To calculate $VK_+(t)$ from equation (3), we must determine $Lk(K_0)$ and $VK_-(t)$. The former is easy to calculate. If the number of trapped nodes in the substrate is $2m$, then there are exactly $2m$ $(+)$ intercomponent nodes in K_0 , and hence $Lk(K_0) = m$. $VK_-(t)$ is also straightforward to determine. K_- is the product that would be obtained if the substrate had two fewer trapped nodes; i.e. $2m-2$ nodes. This is illustrated by Figures 7 and 23. K_+ in Figure 7, the positive trefoil, is the product obtained if two supercoil nodes are trapped, and is precisely K_- in Figure 23, in which four supercoil nodes are trapped. Let K_{2m+1} and K_{2m-1} denote, respectively, the products of recombination whose substrates have $2m$ and $2m-2$ trapped supercoil nodes. Then, using equation (3) and the analysis above, we obtain:

$$VK_{2m+1}(t) - tVK_{2m-1}(t) - (1-t)t^{2m} = 0.$$

If $m = 1$ (Fig 7), K_1 is the unknot, and hence $VK_1(t) = 1$. Thus, we immediately obtain a recursion formula that can be solved to give:

$$VK_{2m+1}(t) = t^m \left[1 + (1-t) \sum_{k=1}^m t^{2k} \right]$$

Thus, for $m = 1$,

$$VK_3(t) = -t^4 + t^3 + t,$$

which is the polynomial for the trefoil T_+ (see Fig. 7); and for $m = 2$:

$$VK_5(t) = -t^7 + t^6 - t^5 + t^4 + t^2,$$

which is the polynomial for the positive torus knot, denoted $\bar{5}_1$, in section 6. This calculation can be extended easily to substrates containing any even number of supercoils between the sites. The product is always a positive torus knot and the number of knot nodes increases with the number of trapped supercoils. This is precisely what is observed experimentally (Spengler *et al.*, 1985; Griffith & Nash, 1985). This is the first rigorous mathematical prediction of the result and again shows the power of the polynomial invariants.

One can also use $VK(t)$ to classify the products of Int action on directly repeated sites. In this case, the number of essential supercoil nodes trapped is odd, say $2m-1$, where m is an integer. An example of this is depicted in Figure 24, in which five $(-)$

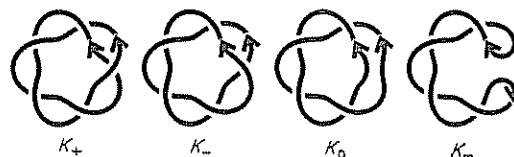


Figure 23. State set for the Int recombination reaction shown in Fig 22. The K_∞ form is the substrate, an unknot with 4 $(-)$ supercoils, and K_+ is the product, the $(+)$ 5-noded torus knot.

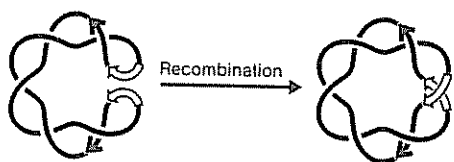


Figure 24. Recombination by Int between directly repeated sites. The filled arrows indicate the global orientation of the substrate and the open arrows indicate the directly repeated recombination sites. Recombination introduces a (-) node at these sites and converts the substrate into 2 catenated rings with 6 (-) nodes, 5 of them coming from the (-) supercoils in the substrate.

supercoils are trapped between the synapsed sites in the substrate, an unknot. In this case, a (0) node rather than an (∞) node is formed by a parallel orientation of the sites and recombination creates a (-) node (Griffith & Nash, 1985), yielding a catenane with $2m$ (-) nodes.

As described above for knots, we will show by a recursion method how to compute the Jones polynomial of the product, knowing only the number of trapped supercoil nodes. The method is illustrated in Figures 12, 14 and 25. K_0 is the substrate, and K_- is the product after recombination. To calculate $VK_-(t)$ using equation (2), we must determine $VK_+(t)$. K_+ is the product that would be obtained if the substrate had started with two fewer trapped nodes; i.e. $2m-3$ (-) nodes. For example, K_- in Figure 12, the negative singly interlinked catenane, is the product if one node is trapped. But this is precisely K_+ in Figure 14, in which three nodes are trapped. Similarly, K_- in Figure 14, RD_{ap} , is the product if three nodes are trapped, and is also K_+ in Figure 25, in which five nodes are trapped.

Let K_{2m} and K_{2m-2} denote, respectively, the products of recombination whose substrates have $2m-1$ and $2m-3$ trapped nodes. Then, using equation (2) and the analysis above, we obtain:

$$t^{-1}VK_{2m-2}(t) - tVK_{2m}(t) - (t^{\frac{1}{2}} - t^{-\frac{1}{2}}) = 0.$$

In the case $m = 1$ (see Fig. 12), K_0 is two unlinked circles, so that $VK_0(t) = -(t^{\frac{1}{2}} + t^{-\frac{1}{2}})$. We obtain a recursion formula that can be solved to give:

$$VK_{2m}(t) = -t^{-2m}(t^{\frac{1}{2}} + t^{-\frac{1}{2}}) - (t^{\frac{1}{2}} - t^{-\frac{1}{2}}) \left[\sum_{k=1}^m t^{1-2k} \right]$$

Thus, for $m = 1$:

$$VK_2(t) = -t^{-1} - t^{-\frac{1}{2}},$$

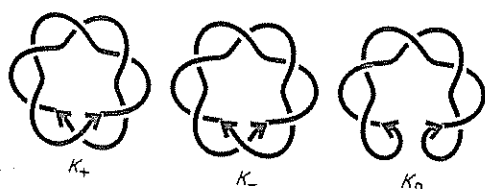


Figure 25. State set for the Int recombination reaction shown in Fig. 24. K_0 is the substrate, an unknot with 5 (-) supercoils, and K_- is the product, a right-handed, antiparallel, triply interlocked torus catenane.

which is the polynomial for L_p (see Fig 12); for $m = 2$:

$$VK_4(t) = -t^{-\frac{3}{2}} - t^{-1} + t^{-\frac{1}{2}} - t^{-\frac{1}{2}},$$

which is the polynomial for RD_{ap} (see Fig 14); and for $m = 3$:

$$VK_6(t) = -t^{-\frac{5}{2}} - t^{-2} + t^{-\frac{3}{2}} - t^{-1} + t^{-\frac{1}{2}} - t^{-\frac{1}{2}},$$

which Table 1 in section 6 shows is the polynomial of the triply interlocked right-handed antiparallel torus catenane, designated $\bar{6}_1^2$ (see Fig 25). This computation may be generalized to different odd numbers of supercoils between the sites in the substrate, and the products always belong to the same subfamily and differ only in the number of nodes. Experiments show that the products are indeed right-handed torus catenanes (Spengler *et al.*, 1985), and this is the first mathematical prediction of the result.

6. Tables of Polynomials

Table 1 gives the two-variable (Lickorish & Millett, 1987) and Jones (1985) polynomials associated with the prime (i.e. irreducible) knots containing up to eight nodes and the prime dimeric catenanes containing up to seven nodes.

For the knots, the first column gives the classical Alexander & Briggs (1927) notation n_i , where n is the number of nodes in a minimal plane presentation (such as shown), and i distinguishes curves with the same number of nodes. The second column gives the Conway (1969) notation, and the third column gives the two-variable polynomial above that of Jones. Drawings of the knots with the same chirality as described by Rolfsen (1976) are shown. Orientation is not indicated because reversal of orientation of the one-component knots shown leaves the knots and their associated polynomials unchanged. The mirror image of the knots are distinct, except for amphichiral knots, indicated by an asterisk next to the Alexander-Briggs notation. The polynomials for the mirror image curves are obtained simply by substituting t^{-1} for t and t^{-1} for t . The resulting knots can be denoted by a tilde over the Alexander-Briggs notation. For example, $\tilde{3}_1$ indicates the (-) trefoil shown and $\bar{3}_1$ indicates the (+) trefoil.

For catenanes, the orientation of the two rings must be specified, and oriented versions of the Rolfsen pictures are shown. The two rings are distinguished by thick and thin lines. The Rolfsen n_j^i nomenclature is in the first column, where n is the minimum number of nodes, i is the number of component rings (2 in this Table), and j distinguishes the catenanes with the same number of nodes. The Conway notation is in the second column, and the one and two variable polynomials are in the third. There are four possible stereoisomers for each catenane depicted. The polynomials for the catenanes obtained by reversing the orientation of the ring drawn with a thin line are given below the pair of polynomials for the depicted

Table 1
Polynomials


















A Knots Alexander- Briggs notation	Conway notation	Lickorish-Millett polynomial		Minimal plane projection	A A B n —
		Jones polynomial			
0 ₁	0	1	1		8
3 ₁	3	$-2l^2 - l^4 + m^2 l^2$ $-l^{-4} + l^{-3} + l^{-1}$			8
4 ₁	22	$-l^{-2} - 1 - l^2 + m^2$ $l^{-2} - l^{-1} + 1 - l + l^2$			8
5 ₁	5	$3l^4 + 2l^6 + m^2(-4l^4 - l^6) + m^4 l^4$ $-l^{-7} + l^{-6} - l^{-5} + l^{-4} + l^{-2}$			8
5 ₂	32	$-l^2 + l^4 + l^6 + m^2(l^2 - l^4)$ $-l^{-6} + l^{-5} - l^{-4} + 2l^{-3} - l^{-2} + l^{-1}$			8
6 ₁	42	$-l^{-2} + l^2 + l^4 + m^2(1 - l^2)$ $l^{-4} - l^{-3} + l^{-2} - 2l^{-1} + 2 - l + l^2$			8
6 ₂	312	$2 + 2l^2 + l^4 + m^2(-1 - 3l^2 - l^4) + m^4 l^2$ $l^{-5} - 2l^{-4} + 2l^{-3} - 2l^{-2} + 2l^{-1} - 1 + l$			8
6 ₃	2112	$l^{-2} + 3 + l^2 + m^2(-l^{-2} - 3 - l^2) + m^4$ $-l^{-3} + 2l^{-2} - 2l^{-1} + 3 - 2l + 2l^2 - l^3$			8
7 ₁	7	$-4l^6 - 3l^8 + m^2(10l^6 + 4l^8) + m^4(-6l^6 - l^8) + m^6 l^6$ $-l^{-10} + l^{-9} - l^{-8} + l^{-7} - l^{-6} + l^{-5} + l^{-3}$			8
7 ₂	52	$-l^2 - l^6 - l^8 + m^2(l^2 - l^4 + l^6)$ $-l^{-8} + l^{-7} - l^{-6} + 2l^{-5} - 2l^{-4} + 2l^{-3} - l^{-2} + l^{-1}$			8
7 ₃	43	$-2l^{-8} - 2l^{-6} + l^{-4} + m^2(l^8 + 3l^6 - 3l^4) + m^4(-l^6 + l^4)$ $l^2 - l^3 + 2l^4 - 2l^5 + 3l^6 - 2l^7 + l^8 - l^9$			8
7 ₄	313	$-l^{-8} + 2l^{-4} + m^2(l^{-6} - 2l^{-4} + l^{-2})$ $l - 2l + 3l^2 - 2l^4 + 3l^5 - 2l^6 + l^7 - l^8$			8
7 ₅	322	$2l^4 - l^8 + m^2(-3l^4 + 2l^6 + l^8) + m^4(l^4 - l^6)$ $-l^{-9} + 2l^{-8} - 3l^{-7} + 3l^{-6} - 3l^{-5} + 3l^{-4} - l^{-3} + l^{-2}$			8
7 ₆	2212	$1 + l^2 + 2l^4 + l^6 + m^2(-1 - 2l^2 - 2l^4) + m^4 l^2$ $-2l^{-7} + 3l^{-6} - 3l^{-4} + 6l^{-3} - 7l^{-2} + 5l^{-1} - 2 + l$			8
7 ₇	21112	$l^{-4} + 2l^{-2} + 2 + m^2(-2l^{-2} - 2 - l^2) + m^4$ $-l^{-3} + 3l^{-2} - 3l^{-1} + 4 - 4l + 3l^2 - 2l^3 + l^4$			8
8 ₁	62	$-l^{-2} - l^4 - l^6 + m^2(1 - l^2 + l^4)$ $l^{-6} - l^{-5} + l^{-4} - 2l^{-3} + 2l^{-2} - 2l^{-1} + 2 - l + l^2$			8
8 ₂	512	$-3l^2 - 3l^4 - l^6 + m^2(4l^2 + 7l^4 + 3l^6) + m^4(-l^2 - 5l^4 - l^6) + m^6 l^4$ $l^{-8} - 2l^{-7} + 2l^{-6} - 3l^{-5} + 3l^{-4} - 2l^{-3} + 2l^{-2} - l^{-1} + 1$			8

Table 1—continued

A. Knots Alexander- Briggs notation	Conway notation	Lickorish-Millett polynomial Jones polynomial	Minimal plane projection
8 ₃	44	$t^{-4} - 1 + t^4 + m^2(-l^{-2} + 2 - l^2)$ $t^{-4} - t^{-3} + 2t^{-2} - 3t^{-1} + 3 - 3t + 2t^2 - t^3 + t^4$	
8 ₄	413	$-2t^{-2} - 2 + t^4 + m^2(l^{-2} + 3 - 2l^2 - l^4) + m^4(-1 + l^2)$ $t^{-5} - 2t^{-4} + 3t^{-3} - 3t^{-2} + 3t^{-1} - 3 + 2t - t^2 + t^3$	
8 ₅	3.3.2	$-2t^{-6} - 5t^{-4} - 4t^{-2} + m^2(3t^{-6} + 8t^{-4} + 4t^{-2}) + m^4(-l^{-6} - 5t^{-4} - l^{-2}) + m^6t^{-4}$ $1 - t + 3t^2 - 3t^3 + 3t^4 - 4t^5 + 3t^6 - 2t^7 + t^8$	
8 ₆	332	$2 + l^2 - t^4 - t^6 + m^2(-1 - 2l^2 + 2l^4 + t^6) + m^4(l^2 - t^4)$ $t^{-7} - 2t^{-6} + 3t^{-5} - 4t^{-4} + 4t^{-3} - 4t^{-2} + 3t^{-1} - 1 + t$	
8 ₇	4112	$-2t^{-4} - 4t^{-2} - 1 + m^2(3t^{-4} + 8t^{-2} + 3) + m^4(-l^{-4} - 5l^{-2} - 1) + m^6t^{-2}$ $-l^{-2} + 2t^{-1} + 2 + 4t - 4t^2 + 4t^3 - 3t^4 + 2t^5 - t^6$	
8 ₈	2312	$-l^{-4} - l^{-2} + 2 + l^2 + m^2(l^{-4} + 2l^{-2} - 2 - l^2) + m^4(-l^{-2} + 1)$ $-l^{-3} + 2t^{-2} - 3t^{-1} + 5 - 4t + 4t^2 - 3t^3 + 2t^4 - t^5$	
8 ₉	3113	$-2t^{-2} - 3 - 2l^2 + m^2(3t^{-2} + 8 + 3l^2) + m^4(-l^{-2} - 5 - l^2) + m^6$ $t^{-4} - 2t^{-3} + 3t^{-2} - 4t^{-1} + 5 - 4t + 3t^2 - 2t^3 + t^4$	
8 ₁₀	3.21.2	$-3t^{-4} - 6t^{-2} - 2 + m^2(3t^{-4} + 9t^{-2} + 3) + m^4(-l^{-4} - 5l^{-2} - 1) + m^6t^{-2}$ $-l^{-2} + 2t^{-1} - 3 + 5t - 4t^2 + 5t^3 - 4t^4 + 2t^5 - t^6$	
8 ₁₁	3212	$1 - l^2 - 2t^4 - t^6 + m^2(-1 - l^2 + 2t^4 + t^6) + m^4(t^2 - t^4)$ $t^{-7} - 2t^{-6} + 3t^{-5} - 5t^{-4} + 5t^{-3} - 4t^{-2} + 4t^{-1} - 2 + t$	
8 ₁₂	2222	$t^{-4} + l^{-2} + 1 + l^2 + t^4 + m^2(-2l^{-2} - 1 - 2l^2) + m^4$ $t^{-4} - 2t^{-3} + 4t^{-2} - 5t^{-1} + 5 - 5t + 4t^2 - 2t^3 + t^4$	
8 ₁₃	31112	$-2l^2 - t^4 + m^2(-l^{-2} - 1 + 2l^2 + t^4) + m^4(1 - l^2)$ $-l^{-5} + 2t^{-4} - 3t^{-3} + 5t^{-2} - 5t^{-1} + 5 - 4t + 3t^2 - t^3$	
8 ₁₄	22112	$1 + m^2(-1 - l^2 + t^4 + t^6) + m^4(l^2 - t^4)$ $t^{-7} - 3t^{-6} + 4t^{-5} - 5t^{-4} + 6t^{-3} - 5t^{-2} + 4t^{-1} - 2 + t$	
8 ₁₅	21,21.2	$t^4 - 3t^6 - 4t^8 - t^{10} + m^2(-2t^4 + 5t^6 + 3t^8) + m^4(t^4 - 2t^6)$ $t^{-10} - 3t^{-9} + 4t^{-8} - 6t^{-7} + 6t^{-6} - 5t^{-5} + 5t^{-4} - 2t^{-3} + t^{-2}$	
8 ₁₆	2 20	$-2l^2 - t^4 + m^2(2 + 5l^2 + 2t^4) + m^4(-1 - 4t^2 - t^4) + m^6t^2$ $-t^{-6} + 3t^{-5} - 5t^{-4} + 6t^{-3} - 6t^{-2} + 6t^{-1} - 4 + 3t - t^2$	
8 ₁₇	2 2	$-l^{-2} - 1 - l^2 + m^2(2l^{-2} + 5 + 2l^2) + m^4(-l^{-2} - 4 - l^2) + m^6$ $t^{-4} - 3t^{-3} + 5t^{-2} - 6t^{-1} + 7 - 6t + 5t^2 - 3t^3 + t^4$	
8 ₁₈	8*	$l^{-2} + 3 + l^2 + m^2(l^{-2} + 1 + l^2) + m^4(-l^{-2} - 3 - l^2) + m^6$ $t^{-4} - 4t^{-3} + 6t^{-2} - 7t^{-1} + 9 - 7t + 6t^2 - 4t^3 + t^4$	
8 ₁₉	3.3.2.	$-l^{-10} - 5l^{-8} - 5l^{-6} + m^2(5l^{-8} + 10l^{-6}) + m^4(-l^{-8} - 6l^{-6}) + m^6t^{-6}$ $t^2 + t^5 - t^8$	

Table 1—continued

A. Knots			
Alexander-Briggs notation	Conway notation	Lickorish-Millett polynomial Jones polynomial	Minimal plane projection
8 ₂₀	3.21.2-	$-1 - 4l^2 - 2l^4 + m^2(1 + 4l^2 + l^4) - m^4l^2$ $-l^{-5} + l^{-4} - l^{-3} + 2l^{-2} - l^{-1} + 2 - l$	
8 ₂₁	21.21.2-	$-3l^2 - 3l^4 - l^6 + m^2(2l^2 + 3l^4 + l^6) - m^4l^4$ $l^{-7} - 2l^{-6} + 2l^{-5} - 3l^{-4} + 3l^{-3} - 2l^{-2} + 2l^{-1}$	
B. Catenanes			
Rolfsen notation	Conway notation	Lickorish-Millett polynomial Jones polynomial	Minimal plane projection
0 ₁ ²	0	$m^{-1}(-l^{-1} - l)$ $-l^{-1} - l^1$	
2 ₁ ²	2	$m^{-1}(l + l^3) - ml$ $-l^{-1} - l^1$ $m^{-1}(l^{-3} + l^{-1}) - ml^{-1}$ $-l^1 - l^1$	
4 ₁ ²	4	$m^{-1}(-l^3 - l^5) + m(3l^3 + l^5) - m^3l^3$ $-l^{-1} + l^{-1} - l^{-1} - l^{-1}$ $m^{-1}(-l^{-5} - l^{-3}) + m(l^{-3} - l^{-1})$ $-l^1 + l^1 - l^1 - l^1$	
5 ₁ ²	212	$m^{-1}(-l^{-1} - l) + m(l^{-1} + 2l + l^3) - m^3l$ $-l^{-1} + 2l^{-1} - l^{-1} - 3l^1 - l^1$	
6 ₁ ²	6	$m^{-1}(l^5 + l^7) + m(-6l^5 - 3l^7) + m^3(5l^5 + l^7) - m^5l^5$ $-l^{-1} + l^{-1} - l^{-1} + l^{-1} - l^{-1} - l^{-1}$ $m^{-1}(l^{-7} + l^{-5}) + m(-l^{-5} + l^{-3} - l^{-1})$ $-l^1 + l^1 - l^1 + l^1 - l^1 - l^1$	
6 ₂ ²	33	$m^{-1}(l^{-7} + l^{-5}) + m(-l^{-7} - 2l^{-5} + 2l^{-3}) + m^3(l^{-5} - l^{-3})$ $-l^1 + l^1 - 2l^1 + 2l^1 - 2l^1 + l^1 - l^1$ $m^{-1}(l^5 + l^7) + m(2l^3 - 2l^5 - l^7) + m^3(-l^3 + l^5)$ $-l^{-1} + l^{-1} - 2l^{-1} + 2l^{-1} - 2l^{-1} + l^{-1} - l^{-1}$	
6 ₃ ²	222	$m^{-1}(-l^3 - l^5) + m(2l^3 - l^5 - l^7) + m^3(-l^3 + l^5)$ $-l^{-1} + 2l^{-1} - 2l^{-1} + 2l^{-1} - 3l^{-1} + l^{-1} - l^{-1}$ $m^{-1}(-l^{-5} - l^{-3}) + m(2l^{-3} + l^{-1} + l) - m^3l^{-1}$ $-l^{-1} + 2l^{-1} - 2l^1 + 2l^1 - 3l^1 + l^1 - l^1$	
7 ₁ ²	412	$m^{-1}(l + l^3) + m(-3l - 4l^3 - 2l^5) + m^3(l + 4l^3 + l^5) - m^5l^3$ $l^{-1} - 2l^{-1} + 2l^{-1} - 3l^{-1} + 2l^{-1} - 2l^{-1} + l^{-1} - l^1$ $m^{-1}(l^{-3} + l^{-1}) + m(-l^{-3} - 2l^{-1} + l + l^3) + m^3(l^{-1} - l)$ $l^{-1} - 2l^{-1} + 2l^{-1} - 3l^{-1} + 2l^1 - 2l^1 + l^1 - l^1$	

Table 1—continued

B Catenanes			
Rolfsen Notation	Conway notation	Lickorish-Millett polynomial	Minimal plane projection
		Jones polynomial	
7_3^2	3112	$m^{-1}(l^{-3}+l^{-1})+m(-2l^{-3}-5l^{-1}-2l)+m^3(l^{-3}+4l^{-1}+l)-m^5l^{-1}$ $l^{-3}-2l^{-1}+2l^{-3}-4l^3+3l^3-3l^3+2l^3-l^3$ $m^{-1}(l+l^3)+m(l^{-1}-l^3-l^5)+m^3(-l+l^3)$ $l^{-3}-2l^{-1}+2l^{-3}-4l^{-1}+3l^{-1}-3l^{-1}+2l^3-l^3$	
7_3^2	232	$m^{-1}(-l^{-1}-l)+m(l^{-1}+l-l^3-l^5)+m^3(-l+l^3)$ $l^{-3}-2l^{-1}+2l^{-3}-3l^{-1}+3l^{-1}-3l^{-1}+2l^3-l^3$	
7_4^2	3.2.2	$m^{-1}(l^{-5}+3l^{-3}+2l^{-1})+m(-2l^{-5}-5l^{-3}-3l^{-1})+m^3(l^{-5}+4l^{-3}+l^{-1})-m^5l^3$ $-l^{-3}+l^3-3l^3+2l^3-3l^3+3l^3-2l^3+l^3$	
7_5^2	21.2.2	$m^{-1}(2l^5+3l^7+l^9)+m(l^3-4l^5-3l^7)+m^3(-l^3+2l^5)$ $l^{-3}-2l^{-1}+3l^{-3}-4l^{-3}+3l^{-1}-4l^{-1}+2l^{-1}-l^{-1}$ $m^{-1}(2l^{-3}+3l^{-1}+l)+m(-2l^{-3}-6l^{-1}-2l)+m^3(l^{-3}+4l^{-1}+l)-m^5l^{-1}$ $l^{-3}-2l^{-1}+3l^{-3}-4l^3+3l^3-4l^3+2l^3-l^3$	
7_6^2	2	$m^{-1}(-l^{-1}-l)+m(-l^{-1}-2l-l^3)+m^3(l^{-1}+3l+l^3)-m^5l$ $-l^{-1}+3l^{-1}-4l^{-1}+4l^{-1}-5l^{-1}+3l^3-3l^3+l^3$	
7_7^2	3.2.2-	$m^{-1}(l^{-9}+3l^{-7}+2l^{-5})+m(-4l^{-7}-6l^{-5})+m^3(l^{-7}+5l^{-5})-m^5l^{-5}$ $-l^{-1}-l^3-l^3+l^3$ $m^{-1}(l^{-1}+3l+2l^3)+m(-l^{-1}-4l-l^3)+m^3l$ $-l^{-1}-l^{-1}-l^{-1}+l^3$	
7_8^2	21.2.2-	$m^{-1}(2l+3l^3+l^5)+m(-2l-3l^3-l^5)+m^3l^3$ $l^{-3}+l^{-1}-5l^{-1}+4l^{-1}-l^{-1}-2l^{-1}$	

catenanes, except for the rare catenanes whose structure is orientation-independent. The polynomials for the mirror images of these forms can be obtained by substituting l^{-1} for l and l^{-1} for t . A compact nomenclature is n_j^i for the catenane shown, \hat{n}_j^i for the reverse catenane whose polynomial is also given, \tilde{n}_j^i for the mirror of the model catenane, and \check{n}_j^i for the mirror reversed form. Thus, the four doubly interlocked torus catenanes are coded 4_1^2 , $\hat{4}_1^2$, $\tilde{4}_1^2$, and $\check{4}_1^2$ for the left-handed parallel, left-handed anti-parallel, right-handed parallel, and right-handed anti-parallel forms, respectively.

Several additional generalizations can be made. First, reversing the orientation of the rings in a catenane changes the value of $VK(l)$ in a simple way, as described above. Second, the powers of the variables l and m are even if the number of components is odd, and odd if the number of components is even. Thus, in Table 1, the powers of l and m are even for knots because they have only a single component, and odd for catenanes because they have two components. The lowest power of m is 1 minus the number of catenated components, and is thus -1 for dimeric catenanes. Third, for

amphichiral curves, the polynomials are unchanged by reversal of the signs of the exponents of l and t . Fourth, the knots and catenanes that are two-bridge (Schubert, 1956) have a single Conway number without commas or decimal points. For these curves, the ratio of the Schubert invariants, β/α , can be calculated from the continued fraction of the Conway number, as explained elsewhere (White & Cozzarelli, 1984).

7. Conclusions

In the study of the topology of DNA, mathematical indices associated with knots and catenanes are important in two ways. They provide, first of all, a precise way of describing and classifying these forms that discriminate their essential features from incidental aspects of structure, thereby allowing the rigorous determination of whether two forms are identical or not, and how many (and which) related forms exist. Second, by delineating critical aspects of structure, they facilitate the determination of the mechanism of enzymes that form and unlink knots and catenanes. They allow calculation of key

reaction parameters that cannot be measured directly and provide, through their predicative and explicative power, rigorous tests of possible mechanisms.

We have described in the last three years, three indices of knots and catenanes: node number (Cozzarelli *et al.*, 1984), Schubert invariants (White & Cozzarelli, 1984), and, in this paper, polynomials. Although they are listed in ascending order of descriptive power, all have uses in the twin goals of classification and determination of enzyme mechanism.

The problem of classification is essentially solved for the known DNA knots and catenanes. Node number provides a ready first-level codification and is also critical for the structure of DNA in solution. The electrophoretic mobility of DNA knots and catenanes is generally proportional to node number, and is not strongly dependent on geometry, such that in the familiar electrophoretic ladders of DNA knots and catenanes each run can contain many distinct forms with the same node number (Dean *et al.*, 1985). Node counts also have the distinct advantage over the other two indices of describing other important structural features of DNA such as double-helical twist, supercoiling and linking number. The simple, integral Schubert invariants, β and α , have the advantage of classification uniqueness. Many different knots and catenanes have the same number of nodes and, extremely rarely, have the same polynomials. Nonetheless, it is the development of the polynomial invariants which basically solves the classification problem for biology at this time. These invariants can be attached to all knots and catenanes, and there are only a small number of duplications in the thousands of prime knots and catenanes containing up to 13 nodes (Thistlethwaite, 1987). We have endeavored to show that they can be calculated easily using recursive processes that can be facilitated using tables of polynomials. The application of the polynomials to the study of DNA brings these studies up to date with work at the forefront of topology.

In developing indices for determining the mechanism of enzymes of DNA metabolism, the ultimate goals are the description of all the rearrangements of DNA and the prediction of the exact structure of all possible products. Substantial gains have been made, such that these goals can be met completely in special cases and partially in all cases. Node counting has the important advantage of describing all the key topological and geometric properties of DNA that are altered by enzymes. Also, critical aspects of the mechanism of enzymes such as topoisomerases and recombinases can be encoded in two or three node-count indices (Cozzarelli *et al.*, 1984; Benjamin & Cozzarelli, 1986). These have been determined for several enzymes and allow the algebraic sum of nodes in products to be readily calculated. Often, the mechanism of a reaction specifies the topological family (e.g. torus) of the products and that makes

node number a much more precise predictor of product structure. The limitation, of course, is that node number by itself is not a strong index (and not even a true invariant (Thistlethwaite, 1987)), and thus its predictions and tests of mechanism suffer accordingly. The Schubert invariants by themselves are not useful in describing enzyme mechanism but when combined with the topology of tangled curves show great promise (Conway, 1969; D. Sumners, N. Cozzarelli and S. Spengler, personal communication). Of necessity, though, they apply only to two-bridge forms. The key biological limitation of invariants is that they are usually insensitive to writhe and double-helical twist because these are geometric and not topological properties. The polynomial of a DNA unknot is thus always equal to 1 irrespective of its supercoiling or helical form. We presented here, though, examples of site-specific recombination mechanisms for resolvase and Int, in which the supercoiling of the substrate that contributes to product knot and catenane structure was incorporated into a polynomial treatment. This was done by the comparison of the substrate and product in the state set. Thus, the structure of the products of different recombinational schemes could be predicted uniquely.

There are only two chief ways in biology for changing the higher-order structure of DNA, strand passage and strand exchange, the hallmarks of topoisomerases and recombinases, respectively. The main theorem of the polynomial methods depends precisely on just these two operations, thereby unifying the appropriate mathematics and biology.

We thank Vaughan Jones and Sylvia Spengler for many helpful discussions and John Jenkins for formatting the paper and checking the calculations. This work was supported by NSF grants DMS8503771 and DMS8613356 and NIH grants GM31655 and GM31657.

References

- Alexander, J. W. (1928) *Trans Amer Math Soc* **30**, 275-306
- Alexander, J. W. & Briggs, G. B. (1927) *Ann Math* **2**, 562-586
- Benjamin, H. W. & Cozzarelli, N. R. (1986) *Proceedings of The Robert A. Welch Foundation Conferences on Chemical Research, Genetic Chemistry: The Molecular Basis of Heredity*, vol. 29, pp. 107-126. Robert A. Welch Foundation, Houston, Texas.
- Benyajati, C. & Worcel, A. (1976) *Cell*, **9**, 393-407.
- Conway, J. H. (1969). In *Computational Problems in Abstract Algebra* (Leech, J., ed), pp 329-358, Pergamon Press, Oxford.
- Cozzarelli, N. R., Krasnow, M. A., Gerrard, S. P. & White, J. H. (1984). *Cold Spring Harbor Symp Quant Biol* **49**, 383-400.
- Dean, F., Stasiak, A., Koller, Th. & Cozzarelli, N. R. (1985) *J Biol Chem* **260**, 4975-4983.
- DiNardo, S., Voelkel, K. & Sternglanz, R. (1984) *Proc Nat Acad Sci. U.S.A* **81**, 2616-2620.
- Englund, P. T., Hajduk, S. L. & Marini, J. C. (1982) *Annu Rev Biochem*, **51**, 695-726.
- Freyd, P., Yetter, D., Hoste, J., Lickorish, W. B. R.,

- Millett, K & Ocneanu, A. (1985) *Bull A MS* 12, 239-246
- Griffith, J & Nash, H A (1985) *Proc Nat Acad Sci, USA* 82, 3124-3128
- Hoste, J (1986) *Pac J Math* 124, 295-320.
- Jones, V F R. (1985). *Bull A MS* 12, 103-111
- Kasamatsu, H & Vinograd, J (1974). *Annu Rev Biochem.* 43, 695-719.
- Krasnow, M, Stasiak, A., Spengler, S. J., Dean, F., Koller, Th & Cozzarelli, N R (1983) *Nature (London)*, 304, 599-560
- Lickorish, W. B R (1986) *Math Proc Camb Phil Soc* 100, 109-112
- Lickorish, W B R & Millett, K. C (1986) *Pac J Math* 124, 173-176.
- Lickorish, W. B R. & Millett, K. C (1987) *Topology*, 26, 107-141
- Nash, H A (1981) *Annu Rev Genet* 15, 143-167.
- Rolfson, D (1976) *Knots and Links*. Publish or Perish, Inc., Berkeley, CA
- Schubert, H. (1956) *Math Z* 65, 133-170
- Spengler, S J, Stasiak, A & Cozzarelli, N R (1985). *Cell*, 42, 325
- Stonington, G. O & Pettijohn, D E (1971) *Proc Nat Acad Sci, USA* 68, 6-9
- Sundin, O & Varshavsky, A (1980) *Cell*. 21, 103-114
- Sundin, O & Varshavsky, A (1981) *Cell*, 25, 659-669
- Thistlethwaite, M B (1985). In *Aspects of Topology* (James, I. M. & Kronheimer, E H, eds), vol 93, pp 1-76. London Math Soc Lecture Notes
- Thistlethwaite, M. B (1987) *Math. Comp.* In the press
- Uemura, T & Yanagida, M (1984) *EMBO J* 3, 1737-1744
- Wang, J C. (1985) *Annu Rev Biochem* 54, 665-697.
- Wasserman, S. A. & Cozzarelli, N R (1984) *Proc Nat Acad Sci, USA* 82, 1079-1083
- Wasserman, S A & Cozzarelli, N R (1986) *Science*. 232, 951-960
- Wasserman, S A, Dungan, J & Cozzarelli, N R (1985) *Science*, 229, 171-174
- White, J H & Cozzarelli, N R (1984) *Proc Nat Acad Sci, USA* 81, 3322-3326

Edited by A. Klug