

A simple topological method for describing stereoisomers of DNA catenanes and knots

(topoisomers/linking number/trefoils/sign and handedness/two-bridge links)

JAMES H. WHITE* AND NICHOLAS R. COZZARELLI†

*Department of Mathematics, University of California, Los Angeles, CA 90024; and †Department of Molecular Biology, University of California, Berkeley, CA 94720

Communicated by Robley C. Williams, February 13, 1984

ABSTRACT Although linking number is an effective topological invariant for describing supercoiled DNA, it is inadequate for the additional interwinding in catenated or knotted DNA. We explain how the two-bridge theory of Schubert provides a powerful yet simple method for analyzing these forms by associating them with two integral invariants, α and β , that measure their geometric complexity. These integers can either be determined graphically or computed with the aid of standard tables, and they allow tabulation of all possible stereoisomers of a given knot or catenane. A complete classification can then be made via a simple theorem. Stereoisomers of representative knots and catenanes are tabulated for easy reference. There are four stereoisomers of regularly interlocked catenanes that we designate right-handed parallel, right-handed antiparallel, left-handed parallel, and left-handed antiparallel according to the helical intertwining of the rings. The biological processes that form catenanes—replication, recombination, and topoisomerase action—predict distinctly different isomers.

All natural populations of DNA rings are to some extent interlocked as catenanes. The fraction is small for most viral and plasmid DNAs, but in some instances catenanes are the predominant species (1, 2). The ubiquity of catenanes arises from the multiplicity of processes that produce them. Replication or recombination of DNA rings in purified enzyme systems usually generates catenanes as the sole or major product (3–6). Moreover, all cells contain enzymes called topoisomerases that can form and resolve catenanes (7).

Interest in catenanes stems not just from their widespread occurrence; knowledge of their structure has provided critical insight into the mechanism of the processes that generate them. They have been important in establishing how the two types of topoisomerases pass DNA segments through each other, in suggesting how topological problems at the termination of DNA replication can be bypassed, and in showing how recombination enzymes in a unique fashion synapse and rearrange sites extensively intertwined by supercoiling (5, 8–11).

The need for rigorous classification of catenanes has become more pressing with improved analytical techniques. Agarose gel electrophoresis resolves even complex mixtures of dimeric DNA catenanes into discrete bands according to the number of interlocks (11). An improved electron microscopic technique for viewing DNA has recently provided the first reliable method for determining the complete structure of knots and catenanes (10). What is not widely known among biologists are the topologically distinct ways rings can be interlocked.

The formulation for the linking number of two curves by White in 1969 (12) has been of great value in analyzing DNA

structure. The linking number of closed circular duplex DNA is the sum of the writhing of the DNA axis in space and the twisting of either strand of the double helix about this axis (13, 14). Unlike linking number, neither of these features need be invariant or integral for a given DNA molecule. Linking number, however, is not very useful for describing the additional interwinding of DNA strands in catenanes, because many topologically different catenanes have the same linking number (15).

Fortunately the two-bridge description of curves by Schubert (16) distinguishes in a simple way the stereoisomers of many catenanes by associating them with two integers, α and β . We show how a catenane can be diagrammed in the Schubert two-bridge fashion that presents α and β graphically. This procedure may then be used to find α and β for the stereoisomers of the catenane with orientations reversed or reflected in a mirror. Via a simple theorem, the identity of these curves can be analyzed for a complete classification. We determine two important corollaries concerning α and β that can be used to tabulate all stereoisomers of a given type of catenane. A table of representative examples is presented for reference apart from the mathematics. We describe an alternative to two-bridge diagrams, called grid diagrams, that are easier to construct once α and β are known. Finally, we apply the two-bridge treatment to DNA knots.

Catenane Description in the Two-Bridge Fashion. We begin by applying the theory of two-bridge links to two kinds of DNA catenanes. First, we classify all possible pairs of closed circular DNAs interlocked with a regular right-handed or left-handed intertwining. In the bottom row of Figs. 1 and 2, single and double interlinks are illustrated. Second, we present the case of figure-8 catenanes that contain an entrapped supercoil (Fig. 3). These two cases represent the only types of links between DNA rings that have been established thus far (1, 10).

In the middle row of Figs. 1–3, these DNAs are drawn in the two-bridge fashion. The reason for the name two-bridge is easily explained. In Fig. 1*b*, the black curve crosses over, or bridges, the grey curve on the left side of the figure, and the grey curve bridges the black curve on the right side. The horizontal parts of the curves are the bridges. In Fig. 2*b*, the black curve bridges the grey and black curves a total of three times, as does the grey curve. In Fig. 3*b*, the grey and black curves each bridge the curves seven times. In each case, there are two and only two bridges.

Schubert has shown that associated with any two-bridge link are integers α and β with the properties that $|\beta| < \alpha$, their greatest common divisor is one, α is even, and β is odd (16). α is one plus the number of times either bridge crosses over the curves. Thus, α is 2 for the single interlock, 4 for the double, and 8 for the figure-8 catenane. For two curves that interlock in the regular way n times, α is equal to $2n$ —i.e., the number of intermolecular nodes.

To describe the determination of β , we must introduce the concepts of orientation and oriented crossings. An orienta-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

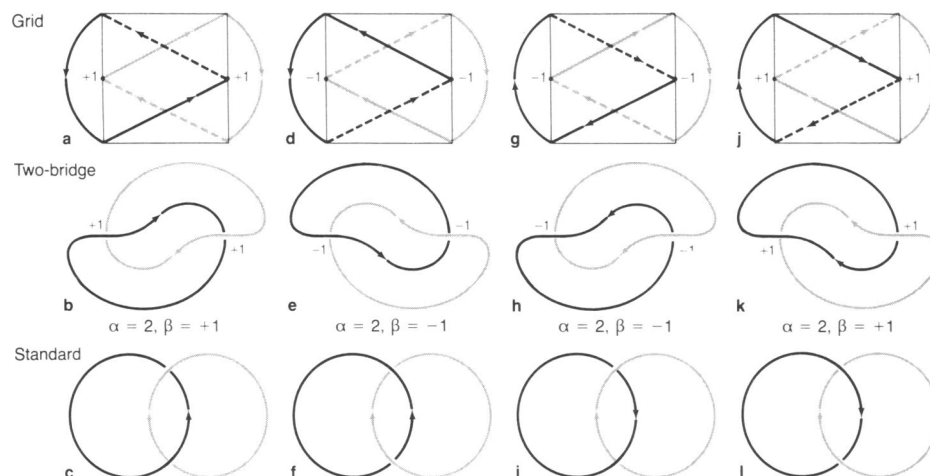


FIG. 1. Representations of singly interlocked regular dimeric catenanes. The black and grey curves represent interlocked DNA rings. The catenanes are shown in three topologically equivalent representations: the grid form (top row), the Schubert two-bridge form (middle row), and the manner standard in biology (bottom row). The orientation of curves is shown by arrows. The underlying of crossing segments is interrupted in two-bridge and standard representations. This convention is not followed in grid plots, where underlying portions of curves are dashed. The numbers are assigned to crossings according to the Schubert scheme.

tion of a curve is a choice of direction in which to travel along the curve. If an oriented curve crosses itself or another oriented curve, the crossing or node may be given a sign according to the convention shown in Fig. 4. This is the common mathematical convention and the one used in some (10) but not all (14) previous discussions of DNA topology. The orientation of linked curves can be the same or opposite and each of these forms has a mirror image. Thus, four two-bridge curves are generated for each of the three catenanes considered. The next step in determining β is to number the bridge crossings as follows. Move along the bridge in the direction of its orientation. Once the far end of the bridge is reached, back up, and number sequentially each crossing 1, 2, 3, . . . with sign dictated by the convention. Finally, start on one bridge and proceed along the curve in the direction of its orientation until it crosses under the other bridge for the first time. The signed number attached to that crossing is β .

If the orientation of one of the curves is reversed, then its bridge crossings are counted in reverse order and β is changed to $\alpha - \beta$ in absolute value. Furthermore, since orientation has been reversed, the sign of the crossing is also changed. Thus, our first corollary is that under a reversal of

orientation, β is changed to $-(\alpha - \beta)$. In Figs. 1-3, the catenanes in *h* and *k* have the reverse orientation of those in *b* and *e*, respectively. If the entire catenane is reflected in a mirror, then all crossing signs are changed but the orientations are preserved. Thus, β is changed to $-\beta$ and this is our second corollary. In Figs. 1-3, *b* and *e* are mirror images, as are *h* and *k*.

In summary, the four catenanes obtained from a given catenane type have the same α and the possible β values are β , $-\beta$, $-(\alpha - \beta)$, and $\alpha - \beta$. These are designated as model, mirror, reversed, and mirror-reversed catenanes, respectively.

Classification of the Illustrated Catenanes. We now classify our catenanes using the theorem proven by Schubert (16). Let Ca_1 and Ca_2 be two catenanes with the same value of α , and let β_1 be the β of Ca_1 and β_2 the β of Ca_2 . Then Ca_1 and Ca_2 can be deformed one into the other without breakage if, and only if, either $\beta_1 \equiv \beta_2 \pmod{2\alpha}$ or $\beta_1\beta_2 \equiv 1 \pmod{2\alpha}$. Because a number mod 2α is the remainder after division by 2α , two numbers are equal mod 2α if their difference is divisible by 2α . Thus, $\beta_1 \equiv \beta_2 \pmod{2\alpha}$ if 2α divides $\beta_1 - \beta_2$, and $\beta_1\beta_2 \equiv 1 \pmod{2\alpha}$ if 2α divides $\beta_1\beta_2 - 1$.

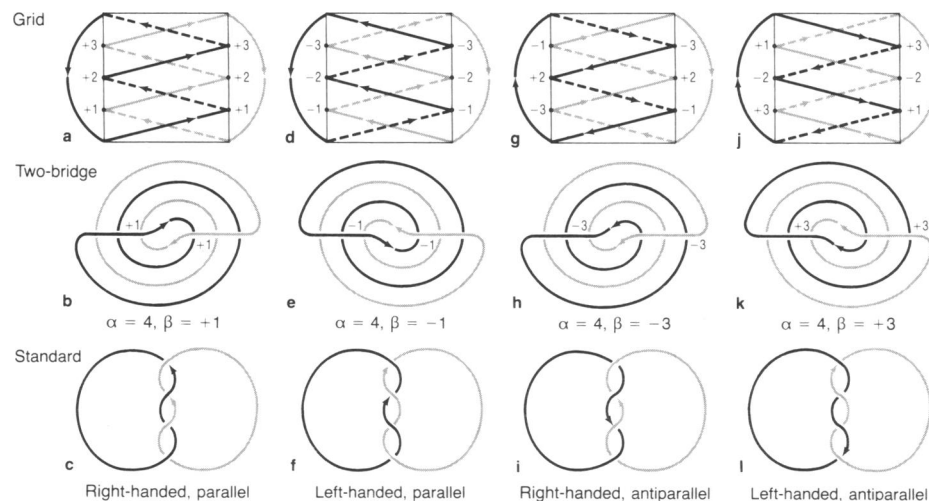


FIG. 2. Representations of doubly interlocked regular dimeric catenanes. The curves are drawn as described in Fig. 1 except, for clarity, only the β th value of bridge crossings are shown in the middle row.

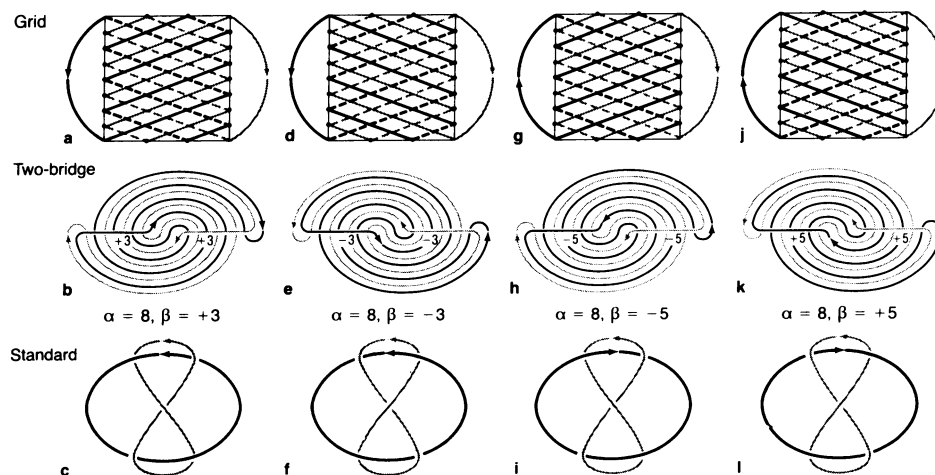


FIG. 3. Representations of figure-8 catenanes. The curves are drawn as described in Fig. 1, except only the β th value of bridge crossings are shown in the middle row and none are shown in the top row.

To determine whether two catenanes are equivalent—i.e., whether one may be continuously deformed into the other—we need only to determine if any of the pairwise products of the numbers β , $-\beta$, $-(\alpha - \beta)$, and $\alpha - \beta$ is equal to 1 (mod 2α). There are four such products: $-\beta^2$, $-\beta(\alpha - \beta)$, $\beta(\alpha - \beta)$, and $-(\alpha - \beta)^2$. The last can be simplified mod 2α to $-\beta^2$ as follows. Since β is an integer, $2\alpha\beta \equiv 0 \pmod{2\alpha}$; since α is even for catenanes, $\alpha/2$ is integral and $\alpha^2 \equiv 0 \pmod{2\alpha}$ because $\alpha^2 = 2\alpha(\alpha/2)$. Thus the table of disjoint products is:

	(model)	(mirror)	(reversed)	(mirror reversed)
	β	$-\beta$	$-(\alpha - \beta)$	$\alpha - \beta$
(model)	β	$-\beta^2$	$-\beta(\alpha - \beta)$	$\beta(\alpha - \beta)$
(mirror)	$-\beta$		$\beta(\alpha - \beta)$	$-\beta(\alpha - \beta)$
(reversed)	$-(\alpha - \beta)$			$-\beta^2$

We need then only decide if $-\beta^2$, $\beta(\alpha - \beta)$, or $-\beta(\alpha - \beta)$ is equal to 1 (mod 2α).

We apply these criteria to our special cases. For singly interlocked catenanes, $\alpha = 2$, β and $(\alpha - \beta) = 1$, and $-\beta$ and $-(\alpha - \beta) = -1$. Therefore, mirror reversal of singly interlocked rings does not change the value of β . This equivalence can be seen in Fig. 1 because a simple rotation about a horizontal diameter converts *c* into *l* and *f* into *i*. Because $-\beta^2$ is not congruent to 1 (mod 4), there are two classes of singly linked catenanes and they are mirror images. For the

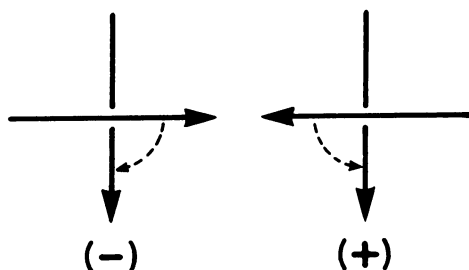


FIG. 4. Sign convention for crossings of oriented curves. The bold arrows are the tangent vectors for two oriented curves, and the curved dashed arrows show the direction the vector on top must be rotated (angle $< 180^\circ$) to be congruent with the underlying vector. Clockwise and counterclockwise motion define (-) and (+) crossings, respectively.

model doubly interlocked catenane, $\alpha = 4$, $\beta = 1$, and $\alpha - \beta = 3$. All four stereoisomers are independent because $-\beta^2$, $\beta(\alpha - \beta)$, and $-\beta(\alpha - \beta)$ equal -1 , -3 , and 3 , respectively, which are not congruent to 1 (mod 8).

For the general case of regular helically interlocked catenanes, $\alpha = 2n$ and $\beta = 1$ for the model catenane. The product table is:

	1	-1	$-(2n - 1)$	$2n - 1$
1		-1	$-(2n - 1)$	$2n - 1$
-1			$2n - 1$	$-(2n - 1)$
$-(2n - 1)$				-1

Since for $n > 1$, -1 , $-(2n - 1)$, and $2n - 1$ are not congruent to 1 (mod $4n$), all four classes are independent.

Thus, there are exactly four classes of regularly interlocked catenanes. We can describe these analogously to the classes of DNA double helices, because the two rings helically intertwine. The helix can be left-handed or right-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, as shown in the bottom row of Fig. 2. Right-handed parallel and left-handed antiparallel catenanes have only positive nodes, whereas left-handed parallel and right-handed antiparallel catenanes have negative nodes. One might have expected only two classes of catenanes corresponding to right- and left-handed or positive and negative nodes, but because handedness and node sign assort independently, there are four.

It is important to note that the various processes for generating catenanes predict different structures. Catenanes resulting at the termination of DNA replication (11) preserve the positive nodes and the right-handed intertwining of the double helix and are therefore parallel. During intramolecular recombination, catenanes that result from the intertwining of substrate plectonemic supercoils (8) must have negative nodes and yet be right-handed, as are such supercoils, and therefore they must be antiparallel. *Escherichia coli* topoisomerase I generates a racemic mixture of duplex DNA knots (10) and should produce a mixture of catenanes. Racemic catenanes should also result from cyclization of linear DNA in the presence of DNA rings (17).

Strikingly, there are only two stereoisomers of figure-8 catenanes (Fig. 3) even though they are more elaborate than

regular catenanes. For figure-8 catenanes, $\alpha = 8$ and the possible values of β are 3 for the model, -3 for the mirror, -5 for the reversed, and 5 for the mirror reversed. The possible products are then -9 , -15 , and 15. Now, $-15 \equiv 1 \pmod{16}$. Hence, by Schubert's theorem, the $\beta = 3$ and $\beta = -5$ catenanes are equivalent, as are their mirror images. This means that reversal of orientation of one of the two DNA rings does not alter the catenane, but that mirror images are not superimposable. We note that the intermolecular linking number must be zero, because linking number changes sign under reversal of one of the curves (12). Since separation of the two rings of a figure-8 catenane requires a double-strand break, a pair of DNA rings that are topologically linked can nonetheless have a linking number of zero.

Classification of Additional Examples. To classify catenanes, it is possible to bypass the Schubert two-bridge drawings and to compute α and β with the aid of a standard table. Rolfsen's useful table (15) illustrates a single unoriented example of many catenanes along with a descriptive notation developed by Conway (18). The Conway notation is a single integer for curves that can be drawn in the two-bridge fashion; this includes all catenanes in the Rolfsen table that have up to 7 nodes and many more complex links. The continued fraction of the Conway integer is equal to $-\beta/\alpha$ and the continued fraction of a number abc is

$$a + \frac{1}{b + \frac{1}{c}}$$

A negative sign results because Conway used the opposite sign convention from ours. The complete set of allowed values of β can then be calculated from the corollaries above and the Schubert classification theorem applied. Because the stereoisomers of curves are not given in Rolfsen's table, we have compiled in Table 1 the values of α , β , the Conway notation, and the notation used by Rolfsen for all possible stereoisomers of some common catenanes. Comparison of Table 1 with the Rolfsen table will make it easier to classify additional examples.

Grid plots. The Schubert two-bridge representations are valuable for a topological description of a catenane but can be difficult to construct. An alternative graphical presentation, called the grid or rectangular plot, is much easier to draw. The grid cannot be used to determine the value of α and β , because its construction requires them. The grids are useful, however, in perceiving the topology of a curve and its relationship to other curves. Examples of grids of our catenane types are shown in the top row of Figs. 1-3.

We explain step-by-step how to construct grids for doubly interlocked (Fig. 2a) and figure-8 (Fig. 3a) catenanes. For the former, $\alpha = 4$ and $\beta = 1$; for the latter, $\alpha = 8$ and $\beta = 3$. Start with a rectangle whose vertical sides are divided into α segments and whose horizontal sides are divided into β segments. Begin at the lower left-hand corner and draw a straight line to the β th dividing point on the right vertical wall. Think of the rectangle as a solid plate with the line segment on top. Next, draw a line connecting the β th dividing point on the right vertical wall to the 2β th divisor on the left wall on the back side of the rectangle. Repeat this procedure until the upper left hand corner or the upper horizontal top is reached. In the $\alpha = 4$ and $\beta = 1$ model, the upper left-hand corner is reached in four steps. Now, connect this corner to the lower left hand corner by an arc. In the $\alpha = 8$ and $\beta = 3$ model, after the first two steps you are searching for the 9th dividing point on the right-hand wall, but there are only 8 segments, so that the top wall is reached first. The next step is to count down the right-hand wall one notch to divider point number 7 and draw on the back side of the plate a line

Table 1. Stereoisomers of DNA catenanes and knots

Rolfsen* notation	Nodes†	α	β	Conway* notation	Fig.
Catenanes					
2_1^2	2+	2	+1		1 c and l
2_1^2	2-	2	-1	2	1 f and i
$2n_1^2$	2n+	2n	+1		2c
$2n_1^2$	2n-	2n	-1	2n	2f
$2n_1^2$	2n-	2n	$-(2n-1)$		2i
$2n_1^2$	2n+	2n	$+(2n-1)$		2l
5_1^2	3+, 2-	8	+3		3 c and i
5_1^2	3-, 2+	8	-3	212	3 f and l
6_2^2	6+	10	+3		
6_2^2	6-	10	-3	33	
6_2^2	6-	10	-7		
6_2^2	6+	10	+7		
6_3^2	6+	12	+5		
6_3^2	6-	12	-5	222	
6_3^2	4-, 2+	12	-7		
6_3^2	4+, 2-	12	+7		
Knots					
$(2m+1)_1$	$(2m+1)+$	$2m+1$	+1		5
$(2m+1)_1$	$(2m+1)-$	$2m+1$	-1	$2m+1$	
4_1	2+, 2-	5	+3	22	
5_2	5+	7	+3		
5_2	5-	7	-3	32	
6_1	4+, 2-	9	+5		
6_1	4-, 2+	9	-5	42	
6_2	4+, 2-	11	+3		
6_2	4-, 2+	11	-3	312	
6_3	3-, 3+	13	+5	2112	

*The first number in the notation used by Rolfsen (15) is the number of nodes in the simplest representation of the curve. The superscript is the number of catenated rings, and the subscript distinguishes curves with the same number of nodes. Because the curves in Rolfsen's table are not oriented, stereoisomers have the same Rolfsen notation. The Conway notation is given for an oriented version of a curve depicted in Rolfsen's table. For amphichiral knots, the $-\beta$ value is not listed because it is equivalent to the $+\beta$ knot. $n = \text{integer} > 1$; $m = \text{integer} > 0$.

†Node composition in simplest representation of the curve.

from where you left off at the top wall to this point. The same procedure is now used going down and back up this grid. The procedure is always used a total of β times to reach the upper left-hand corner, which is then connected to the lower left-hand corner by an arc. The final step is to repeat the process for the grey curve of both catenanes but starting at the lower right-hand corner and on the back of the plate.

The relationship between grid and Schubert presentations is as follows. The solid parts of the curves on the top side of the grid correspond to the parts of the curves above the bridges. The dashed parts of the curves on the back side of the grid correspond to the parts of the curves below the bridges, and the connecting arcs correspond to the bridges. Bridge crossings correspond to the point where a curve changes sides of the rectangular plate. The handedness of a curve is more apparent in the grid plots. The grid, two-bridge, and standard representations are topologically equivalent; a grid constructed using a model for DNA, such as tubing, can be easily converted to the standard form by removing the rectangular plate support.

Knots. DNA knots are less well-studied than catenanes but are also important products of recombination *in vitro* (5), and topoisomerases readily tie and untie knots (7, 19, 20). Knotted DNA has been identified in bacteriophage capsids (21) and in preparations of bacterial plasmids.

Linking number is undefined for knots. Fortunately, the Schubert theory also applies to knots, because single curves

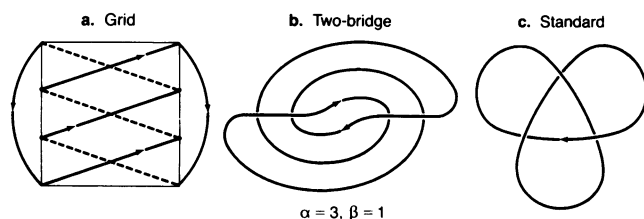


FIG. 5. Representations of a trefoil knot. See Fig. 1 legend for explanation.

can be drawn in the two-bridge fashion. For knots, both α and β are odd. We present just two examples, trefoil and figure-8 knots, because the treatment is analogous to that used for catenanes. For our model trefoil or 3-noded knot, $\beta = 1$ and $\alpha = 3$. Since only a single curve is involved, the curve must cover all dividing points of both the left and right walls of the grid model (Fig. 5a). In the Schubert representation, the same curve serves as both bridges (Fig. 5b). Reversal of orientation does not change β for a knot, but reflection changes the sign of β . β is then -1 for the mirror image of the model trefoil. Mod 6, $-1 \cdot 1 \equiv -1$. Thus, there are two, and only two, trefoil stereoisomers.

For the 4-noded or figure-8 knot, $\alpha = 5$ and $\beta = \pm 3$. However, mod 10, $3 \cdot -3 \equiv 1$. Thus, surprisingly, the mirror images are superimposable and there is only one 4-noded knot; i.e., it is amphichiral. Additional examples of stereoisomers of some common knots are indexed in Table 1, and Rolfsen's table illustrates a single stereoisomer of knots that contain ten or fewer nodes. As in the case of catenanes, the continued fraction of the Conway number is $-\beta/\alpha$, with one complication. If the numerator of the fraction is odd, it is indeed equal to $-\beta$. However, if the numerator is even, then β is the numerator minus α or its equivalent via the Schubert theorem. β is always the smallest positive odd integer, which represents the knot type. For example, for the 4_1 knot, the continued fraction of 22 is $2/5$, so that $\beta = 2 - 5 = -3$. For the 5_2 knot, β is shown to be -3 as follows. The continued fraction of 32 is $2/7$ and $2 - 7 = -5$, which in turn equals -3 in the Schubert system because $-3 \cdot -5 = 15 \equiv 1 \pmod{14}$.

We are indebted to M. Krasnow for stimulating discussions and for the nomenclature for catenane stereoisomers. This work was supported by National Institutes of Health Grants GM 31655 and GM 31657.

1. Kasamatsu, H. & Vinograd, J. (1974) *Annu. Rev. Biochem.* **43**, 695-719.
2. Englund, P. T., Hajduk, S. L. & Marini, J. C. (1982) *Annu. Rev. Biochem.* **51**, 695-726.
3. Sakakibara, Y., Suzuki, K. & Tomizawa, J.-I. (1976) *J. Mol. Biol.* **108**, 569-582.
4. Wold, M. S., Mallory, J. B., Roberts, J. D., LeBowitz, J. H. & McMacken, R. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 6176-6180.
5. Nash, H. A. (1981) *Annu. Rev. Genet.* **15**, 143-167.
6. Reed, R. A. (1981) *Cell* **25**, 713-719.
7. Cozzarelli, N. R. (1980) *Cell* **22**, 327-328.
8. Mizuuchi, K., Gellert, M., Weisberg, R. A. & Nash, H. A. (1980) *J. Mol. Biol.* **141**, 485-494.
9. Krasnow, M. A. & Cozzarelli, N. R. (1983) *Cell* **32**, 1313-1324.
10. Krasnow, M. A., Stasiak, A., Spengler, S. J., Dean, F., Koller, Th. & Cozzarelli, N. R. (1983) *Nature (London)* **304**, 559-560.
11. Sundin, O. & Varshavsky, A. (1981) *Cell* **25**, 659-669.
12. White, J. H. (1969) *Am. J. Math.* **91**, 693-728.
13. Crick, F. H. C. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 2639-2643.
14. Bauer, W. R., Crick, F. H. C. & White, J. H. (1980) *Sci. Am.* **243** (7), 118-133.
15. Rolfsen, D. (1976) *Knots and Links* (Publish or Perish Inc., Wilmington, DE).
16. Schubert, H. (1956) *Math Z.* **65**, 133-170.
17. Wang, J. C. & Schwartz, H. (1967) *Biopolymers* **5**, 953-966.
18. Conway, J. H. (1970) in *Computational Problems in Abstract Algebra*, ed. Leech, J. (Pergamon, Oxford), pp. 329-358.
19. Liu, L. F., Depew, R. E. & Wang, J. C. (1976) *J. Mol. Biol.* **106**, 439-452.
20. Liu, L. F., Liu, C.-C. & Alberts, B. M. (1980) *Cell* **19**, 697-707.
21. Liu, L. F., Perkocha, L., Calendar, R. & Wang, J. C. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 5498-5502.