Decomposition of the linking number of a closed ribbon: A problem from molecular biology

(writhing number/DNA/nucleosome)

F. BROCK FULLER

California Institute of Technology, Sloan 253-37, Pasadena, California 91125

Communicated by F. H. C. Crick, May 12, 1978

ABSTRACT A closed duplex DNA molecule relaxed and containing nucleosomes has a different linking number from the same molecule relaxed and without nucleosomes. What does this say about the structure of the nucleosome? A mathematical study of this question is made, representing the DNA molecule by a ribbon. It is shown that the linking number of a closed ribbon can be decomposed into the linking number of a reference ribbon plus a sum of locally determined "linking differences."

1. Introduction

This paper has its origin in a problem from molecular biology. It has been observed that in a duplex DNA molecule a section about 140 base pairs long may attach itself to a group of protein molecules (histones) to form a compact structure called a nucleosome. An experiment that may give some information about the structure of the nucleosome is to compare the linking number of the two strands of a closed DNA molecule relaxed and containing nucleosomes with the linking number of the same molecule relaxed and without nucleosomes. One can then calculate the "linking per nucleosome" by dividing the difference of the two linking numbers by the number of nucleosomes. It is now tempting to say that this value of "linking per nucleosome" is a property of the nucleosome. That this reasoning rests on a false assumption is shown in Fig. 1, where the DNA molecule is represented by a ribbon (1, 2).

The mathematical reason why the linking number is not additive is that the linking number of two disjoint curves a and b has the properties of a product ab. By the linking number of curves that are not closed is meant their Gauss integral (3). Suppose, as in Fig. 1, that a ribbon A is cut into sections A_i, A = \Sigma A_i, and that a = \Sigma a_i and a' = \Sigma a_i' are the corresponding decompositions of the edges. Then

\[ \text{Lk}(A) = (\sum a_i)(\sum a_i') = \sum a_i a_i' = \sum a_i a_i' + a_i a_i'. \]

The additivity is upset by the mixed terms a_i a_j + a_i' a_j, i \neq j, which involve the "distant interactions" of A_i with A_j.

The objective of this paper is to see to what extent it is possible, in spite of the failure of additivity, to isolate the effect of a section of a closed ribbon on its linking number. It is shown in Section 8 that the linking number can be expressed as the linking number of a "reference" ribbon plus a sum of locally determined linking differences, each of which expresses the effect of altering a section of the reference ribbon.

Sections 2-8 summarize the properties of Lk, Tw, and Wr, including some properties not given in earlier papers (1, 2), in particular the direct interpretation of Wr in terms of spherical area in Section 6. Section 7 introduces the notion of a cord, a generalization of the notion of ribbon. Section 8 defines the linking difference \delta Lk and shows how it can be used to decompose the Lk of a closed cord.

There are some changes in terminology from my earlier paper (1): strip becomes ribbon, the arc length s is replaced by a general parameter t, central curve becomes axis curve, twist \omega_1 becomes angular twist rate \omega_1, and total twist number Tw becomes twist Tw. Differences from Crick's paper (2) are: Tw instead of T, Lk instead of L, Wr instead of W, and the linking number of two curves has the opposite sign convention. For ribbons this opposite convention is compensated by Crick's reversed orientation of one edge, so that the definitions of the linking number of a ribbon agree.

I wish to take the opportunity here to point out that the basic relation Lk = Tw + Wr and a number of the results of my earlier paper (1) had previously appeared in a paper by J. H. White (4).

2. Representation of a polymer by a ribbon

A stiff polymer, such as two-stranded DNA, can be represented by a ribbon (1, 2). The axis of the ribbon is a smooth non-self-intersecting space curve following the axis of the polymer. Analytically the axis curve can be written X(t), where X is a vector from the origin to the curve and t is a real parameter. It is assumed that the tangent velocity vector (d/dt)X is never zero, so that a unit tangent vector T(t) is defined along the curve. The twisting of the polymer about its axis is represented by unit normal vectors U(t) everywhere perpendicular to the axis curve and pointing to selected reference points on the polymer. The axis curve X(t) and the normals U(t) together constitute a ribbon.

The choice of the reference points on the polymer to which the normal vectors U(t) point is, in fact, arbitrary. One choice for two-stranded DNA would be pointing from the axis to one of the two strands. Another choice would be to choose the reference points in such a way that U(t) would be constant along a straight unstrained section. For the three quantities Lk, Tw, and Wr, a different choice of U(t) will change Lk and Tw, but will not change Wr, nor will it change the differences between the values of Lk or Tw for two different positions of the same polymer.

A ribbon is closed if X(t) and U(t) are periodic functions of t with the same period. Note that there may be "circular" or closed polymers that are not represented by closed ribbons, especially in view of the arbitrariness in the choice of U(t). However, these closed polymers will be represented by closed cords, as defined in Section 7.

3. Twist (Tw) of a ribbon

For any ribbon the rate at which U(t) revolves about T(t) in radians per units of t is the angular twist rate \omega_1:

\[ \omega_1 = \frac{d}{dt} U \times U \cdot T \]
The integral of the angular twist rate $\omega_1$, divided by $2\pi$, is the 
\textit{twist} $Tw$:

$$Tw = \frac{1}{2\pi} \int \omega_1 \, dt.$$  \hfill [3.2]

A change of parameter changes $\omega_1$ but does not change $Tw$. The twist $Tw$ counts the net number of right-handed turns that a ribbon makes about its axis.

Properties of $Tw$ are: (i) $Tw$ is invariant under rigid motions or dilations of the space containing the ribbon. It does not depend on which direction the ribbon is traversed. Its sign is changed by reflections in a plane or by reflection in a sphere (inversion by reciprocal radii) provided the axis curve avoids the center of the sphere. The twist $Tw$ is thus a conformal invariant (5).

$Tw$ is not a topological invariant even for a closed ribbon. However, the following is true. Suppose the entire axis curve and the values of the normal vectors $U$ at the two ends of the curve are held fixed. Then the value of $Tw$ is determined modulo 1 (up to an integer) and hence is unchanged by a deformation of the values of $U$ on the curve between the endpoints.

(ii) $Tw$ is additive: $Tw(\alpha + \beta) = Tw(\alpha) + Tw(\beta)$. The meaning of the sum $\alpha + \beta$ of two ribbons $\alpha$ and $\beta$ may be described as follows. Suppose the ribbon $\alpha + \beta$ has the parameter $t$, $a \leq t \leq b$. Then the strip $\alpha$ is the part $a \leq t \leq c$ of $\alpha + \beta$ for some $c$ between $a$ and $b$ and the strip $\beta$ is the remaining part $c \leq t \leq b$.

(iii) If the axis curve of a ribbon is passed through itself, as in Fig. 2, the value of $Tw$ changes continuously without any jump. The reason for this is that the definition of $Tw$ does not require that the axis curve be non-self-intersecting.

4. Linking number ($Lk$) of a closed ribbon

For a closed ribbon the edges $X(t) = U(t)$ are, for all sufficiently small positive $\epsilon$, disjoint closed curves and the linking number $Lk$ of the closed ribbon is defined as the topological linking number of the two edges $(1, 2)$ or, equivalently, of one edge and the axis curve $X(t)$.

Properties of $Lk$ are: (i) $Lk$ is an integer, unchanged by continuous deformations of the ribbon. It does not depend on which direction the ribbon is traversed.

(ii) $Lk$ is unchanged by an orientation-preserving topological
transformation of the containing space into itself. Its sign is changed by an orientation-reversing transformation, such as reflection in a plane.

(iii) If a ribbon is passed through itself, the value of $Lk$ jumps by $\pm 2$ according to Fig. 2.

(iv) The two edges of a closed ribbon can be disentangled if, and only if, $Lk = 0$ and the axis is an unknotted space curve. This is a consequence of Dehn's lemma (6, 7).

5. Writhing number ($Wr$) of a closed non-self-intersecting space curve

A closed ribbon has both a linking number $Lk$ and a twist $Tw$. In general, these numbers are not equal. Their difference, $Lk - Tw$, is called the \textit{writhing number} $Wr$, so that

$$Lk = Tw + Wr.$$ \hfill [5.1]

Properties of $Wr$ are: (i) $Wr$ of a ribbon depends only on its axis curve. The writhing number is thus defined for any smooth, closed, non-self-intersecting space curve. It does not depend on which direction the space curve is traversed.

(ii) $Wr$ is invariant under rigid motions or dilations of the space containing the curve. Its sign is changed by reflection in a plane or by reflection in a sphere, provided the curve avoids the center of the sphere. A consequence of these last two properties is that $Wr = 0$ for any closed curve in a plane or on a sphere; hence $Lk = Tw$ for a closed ribbon whose axis curve lies in a plane or on a sphere. The writhing number $Wr$ is a conformal invariant. It is not a topological invariant.

(iii) If the curve is passed through itself, the value of $Wr$ jumps by $\pm 2$ according to Fig. 3.

Note that each of the three quantities, $Lk$, $Tw$, and $Wr$, defined for a closed ribbon has a desirable property not shared by the other two: $Lk$ is topological, $Tw$ is additive, and $Wr$ depends only on the axis curve of the ribbon. The relation $Lk = Tw + Wr$ is illustrated in Fig. 4.

6. Interpretation of the writhing number $Wr$ in terms of spherical area

The definition of the writhing number $Wr$ of a closed non-self-intersecting curve involves something extraneous to the curve in that one constructs a closed ribbon having the curve as axis and computes $Wr = Lk - Tw$. Since the result does not depend on the ribbon chosen, one may ask whether it is possible to dispense with the ribbon and obtain $Wr$ directly from the curve. This question is answered in part by the following theorem.

\textbf{Theorem.} The \textit{writhing number} $Wr$ of a closed non-self-intersecting space curve has the following direct interpretation. The unit tangents $T(t)$ to the curve trace out, if their starting points are translated to the origin, a closed curve on the unit sphere. Let $A$ be the solid angle in steradians (area on the unit sphere) enclosed by this curve. Then

$$1 + Wr = \frac{1}{2\pi} A \mod 2.$$ \hfill [6.1]

\textbf{Discussion.} Eq. 6.1 is only good modulo 2 because the solid angle enclosed by the curve $T(t)$ is determined only up to integer multiples of $4\pi$ (the area of the unit sphere). However,

\begin{equation}
Wr \left( \begin{array}{c}
X
\end{array} \right) = Wr \left( \begin{array}{c}
X
\end{array} \right) + 2
\end{equation}

\textbf{Fig. 3.} How the value of $Wr$ jumps if a closed non-self-intersecting curve is passed through itself. The broken lines show which way the ends of the crossing solid lines are connected.
intersecting curves $X_1(t), 1 \leq \lambda \leq 2$, in such a way that $T_1(t)$ and $T_2(t)$ are never oppositely directed. Then the writhe numbers of the two space curves are related by

$$W_{r2} - W_{r1} = \frac{1}{2\pi} \int_0^\lambda T_1 \times T_2 \cdot \frac{d}{dt}(T_1 + T_2) dt. \quad [6.4]$$

Note that Eq. 6.4 is an equality, not a congruence modulo 2.

What the integral calculates is the area swept out by the unique shortest geodesic (great circle arc) from $T_1(t)$ to $T_2(t)$ as $t$ runs over a period.

A typical application of Eq. 6.4 would be to calculate the writhe number of a closed “bent helix,” taking $X_1(t)$ to represent the curve “axis” and $X_2(t)$ to represent the bent helix winding around it.

7. Cords

By a cord we shall mean a non-self-intersecting space curve $X(t)$—its axis curve—together with an angular twist rate $\omega(t)$ defined along the curve. Each ribbon thus determines a cord, while to each cord corresponds a family of ribbons with the same axis curve and such that any two ribbons of the family differ by a constant angle along the curve. The normals $U(t)$ for the ribbons corresponding to a given cord are the solutions of the differential equation, Eq. 3.1. The twist $Tw$ of a cord has the same definition and properties as the twist of a ribbon.

A closed cord is one whose axis curve is closed, represented by a periodic function $X(t)$ of $t$, and whose angular twist rate $\omega(t)$ is also periodic with the same period. Note that the ribbons corresponding to a closed cord need not be closed, since the normal $U(t)$ need not return to its initial value when the axis curve is traversed once.

The linking number $Lk$ of a closed cord is defined by the equation $Lk = Tw + Wr$, where $Wr$ is the writhing number of its axis curve. The linking number of a cord has the properties (1) (3) of the linking number of a ribbon, except that it is a real number, and not necessarily an integer.

The deformation invariance and the topological invariance properties of the linking number of a cord are not obvious from its definition. To see these we give an alternative description of $Lk$ in purely topological terms. Let $U(t), -\omega < t < \omega$, be a solution of the differential equation, Eq. 3.1. The curves $X(t) + \epsilon U(t)$, for various solutions $U(t)$ and various sufficiently small positive $\epsilon$, are the trajectories of a flow on a tubular neighborhood of the axis curve $X(t)$. The linking number $Lk$ of the cord is the average rate at which one of the trajectories links the axis curve, in the following sense. A trajectory $X(t) + \epsilon U(t)$ for $0 \leq t \leq np$, where $p$ is the period of $X(t)$, can be completed to a closed curve $c_n$ by adding to it a circular arc of radius $\epsilon$. Then

$$Lk = \lim_{n \to \infty} \frac{1}{n} Lk(c_n, \text{axis curve}). \quad [7.1]$$

The limit [7.1] will exist for flows topologically equivalent to the given one, as follows from the theory of differential equations on a torus (8); $Lk$ is the Poincaré rotation number of the flow on the torus of points at a distance $\epsilon$ from the axis curve, using as reference for the angles a closed curve $X(t) + \epsilon V(t)$, $V(t)$ a unit vector perpendicular to $T(t)$ that is not linked with the axis curve.

The limit [7.1] is not an efficient way to calculate $Lk$. To do that one can use the definition $Lk = Tw + Wr$ in the following way: After $\omega(t)$, and hence $Tw$, in such a way that the altered cord is represented by a closed ribbon. Then

$$Lk(\text{cord}) = Lk(\text{closed ribbon}) - (\text{change in } Tw). \quad [7.2]$$
The linking number of closed ribbon can be obtained by counting overcrossings (refs. 1 and 2 and Fig. 1).

8. The linking difference $\delta Lk$

We have seen in Section 1 that the linking number $Lk$ of a closed cord cannot be obtained by adding up contributions from pieces of the cord. But we shall see that it is possible to isolate the relative contribution to $Lk$ of a piece of the closed cord, in the following sense. Suppose a section of a closed cord is altered. Then the resulting change in $Lk$ is the sum of a "local" term, the linking difference $\delta Lk$, calculated from the old and new sections without reference to the rest of the cord, plus an "interaction" term, expressing how the rest of the cord threads its way between the old and new sections. In most cases the calculation can be arranged so that the interaction term is zero.

Definition: Let $\alpha$ and $\beta$ be two cords whose ends agree in position and direction. Let $\gamma$ be a cord that completes $\alpha$ and $\beta$ to closed cords $\alpha + \gamma$ and $\beta + \gamma$ (see Fig. 5). Then the linking difference $\delta Lk(\alpha, \beta; \gamma)$ of $\alpha$ and $\beta$ relative to $\gamma$ is defined by

$$\delta Lk(\alpha, \beta; \gamma) = Lk(\beta + \gamma) - Lk(\alpha + \gamma).$$  \[8.1\]

Remark: The linking difference $\delta Lk(\alpha, \beta; \gamma)$ is defined for ribs $\alpha$ and $\beta$ if their ends agree in position ($X$), direction ($T$), and normal direction ($U$). In this case it is an integer.

At first glance the definition (Eq. 8.1) does not appear to be useful because of the dependence of $\delta Lk(\alpha, \beta; \gamma)$ on $\gamma$. The point is that the dependence is quite crude. To begin with, we note that $\delta Lk(\alpha, \beta; \gamma)$ depends only on the axis curve $c$ of $\gamma$, since any change in the twist of $\gamma$ affects $Lk(\beta + \gamma)$ and $Lk(\alpha + \gamma)$ equally and so cancels out. We may thus write $\delta Lk(\alpha, \beta; c)$, where $c$ is a curve that completes the axis curves $a$ and $b$ of $\alpha$ and $\beta$ to closed non-self-intersecting curves $a + c$ and $b + c$. Second, a deformation of $c$ does not change $\delta Lk(\alpha, \beta; c)$ so long as $c$ does not cross $a$ or $b$ (if $c$ crosses itself, the resulting $\pm 2$ jump affects $Lk(\beta + \gamma)$ and $Lk(\alpha + \gamma)$ equally). If the effect of $c$ crossing $a$ or $b$ is analyzed, one arrives at the following theorem.

Theorem: Let $\alpha$ and $\beta$ be two cords whose endpoints agree in position and direction (see Fig. 6). Let $a$ and $b$ be their axis curves and let $c$ be a curve that completes $a$ and $b$ to closed non-self-intersecting curves $a + c$ and $b + c$. Let $c'$ be another such curve. Then

$$\delta Lk(\alpha, \beta; c') = \delta Lk(\alpha, \beta; c) + 2Lk(b - a, c' - c).$$  \[8.2\]

The expression $Lk(b - a, c' - c)$ requires some explanation, since the closed curve $b - a$ meets the closed curve $c' - c$ at the two endpoints of $a$ and $b$ and the linking number is defined only for two disjoint closed curves. Here, however, the two curves $b - a$ and $c' - c$ lie on opposite sides of a plane element in the vicinity of each of the points where they meet. There is thus a well-defined way to pull them apart near these two points and $Lk(b - a, c' - c)$ is defined to be the linking number of the separated curves.

![Fig. 5. Example of the linking difference for two ribbons $\alpha$ and $\beta$ whose ends agree (after ref. 2, figure 4b). Using the $\gamma$ shown, $\delta Lk(\alpha, \beta; \gamma) = Lk(\beta + \gamma) - Lk(\alpha + \gamma) = (-1) - (0) = -1$. Since $\delta Lk(\alpha, \beta; \gamma)$ depends only on the axis curve $c$ of $\gamma$, it can as well be written $\delta Lk(\alpha, \beta; c)$.](image)

![Fig. 6. How the linking difference $\delta Lk(\alpha, \beta; c)$ depends on $c$, illustrated for ribbons $\alpha$ and $\beta$. $\delta Lk(\alpha, \beta; \gamma) = 1 - 0 = 1$; $\delta Lk(\alpha, \beta; \gamma') = 3 - 0 = 3$. $Lk(c' - c, b - a) = 1$; Eq. 8.2 is here $3 = 1 + 2 \times 1$. In this example the completing curves $c$ and $c'$ thread their way around and between $\alpha$ and $\beta$ by essentially different paths. The linking difference $\delta Lk(\alpha, \beta; c)$ is the same for all $c$ that stay outside of the interior of a topological ball containing $\alpha$ and $\beta$.](image)

We see from Eq. 8.2 that $\delta Lk(\alpha, \beta; c) = \delta Lk(\alpha, \beta; c')$ if $Lk(b - a, c' - c) = 0$. An important case where this happens is if $\alpha$ and $\beta$ are contained in a topological ball whose interior excludes $c$ and $c'$.

To prove Eq. 8.2 one first notes that the equation is correct if $c'$ = $c$. Then one shows that the difference between the two sides of the equation does not change if $c'$ is deformed, for the terms in the equation are not changed by a deformation of $c'$ unless $c'$ crosses $a$ or $b$. Then a change of $\pm 2$ in $\delta Lk(\alpha, \beta; c)$ is balanced by a change of $\pm 1$ in $Lk(b - a, c' - c)$. Finally, since any $c'$ can be deformed into $c$, the equation must hold for any $c$ and $c'$.

Application of the linking difference is just a matter of interpreting the definition and Eq. 8.2. Suppose a section of a closed cord is changed; let $\alpha$ be the old section, $\beta$ the new, and let $c$ be the rest of the cord. The cord is thus changed from $\alpha + c + \beta + c$ to $\beta + c + \alpha + c$. Calculate $\delta Lk(\alpha, \beta; c)$ for any convenient $c$. This quantity is "local" in the sense that it involves only $\alpha$ and $\beta$ and does not involve the rest of the cord. Then, from Eq. 8.2, the replacement of $\alpha$ by $\beta$ changes the linking number according to

$$Lk(b + r) - Lk(\alpha + r, \beta) = \delta Lk(\alpha, \beta; c) + 2Lk(b - a, r - c),$$

where $r$ is the axis curve of $p$. The "interaction" term $2Lk(b - a, r - c)$ is zero unless $r$ and $c$ thread their way around or between $a$ and $b$ by essentially different paths. In particular, as remarked above, if $\alpha$ and $\beta$ are in a topological ball whose interior excludes both $r$ and $c$, then the interaction term is zero.

Decomposition of the linking number of a closed cord can be described as follows (see Fig. 7). Suppose a closed cord contains a number of mysterious sections (nucleosomes) each contained in a topological ball, and suppose that the interior of each ball excludes the rest of the cord and the other balls. Inside each ball replace each mysterious section by a standard section; one then obtains a "reference" cord. Let $Lk_i$ be the linking difference between the mysterious and the standard section in the $i$th ball, relative to a curve outside the interior of the ball. Then, applying Eq. 8.2 repeatedly, one has the decomposition

$$Lk(\text{original cord}) = Lk(\text{reference cord}) + \sum \delta Lk_i.$$
I wish to acknowledge the benefit of correspondence and conversations with F. H. C. Crick.