The Gauss linking integral

1 Introduction

The linking number of two oriented closed curves in 3-space can be computed by projecting the two curves on a plane and counting their intercrossings with appropriate sign (see knot textbook). This interpretation of the linking number is a simplification of the original definition provided by Gauss as a double integral over the two curves. The original Gauss definition can be applied also to measure how *open* curves interwind around each other. In terms of sums of signs in a projection, the original definition requires doing it for all possible projection directions and taking an average. For polygonal curves, one can avoid doing the average over projections, by, equivalently, using a closed formula of the integral derived by Banchoff.

All this is made precise in the following sections.

We use the Banchoff method to compute the linking of parts of proteins [5, 1].

1.1 The Gauss linking integral

A measure of the degree to which curves interwind is the Gauss linking integral:

(Gauss Linking Number). The Gauss *Linking Number* of two disjoint (closed or open) oriented curves l_1 and l_2, whose parametrizations are $\gamma_1(t), \gamma_2(s)$ respectively, is defined as the following double integral over l_1 and l_2 [3]:

$$L(l_1, l_2) = \frac{1}{4\pi} \int_{[0,1]} \int_{[0,1]} \frac{(\dot{\gamma}_1(t), \dot{\gamma}_2(s), \gamma_1(t) - \gamma_2(s))}{||\gamma_1(t) - \gamma_2(s)||^3} dtds, \tag{1}$$

where $(\dot{\gamma}_1(t), \dot{\gamma}_2(s), \gamma_1(t) - \gamma_2(s))$ is the *scalar triple product* of $\dot{\gamma}_1(t), \dot{\gamma}_2(s)$ and $\gamma_1(t) - \gamma_2(s)$.

For closed curves, the Gauss linking integral is equal to half the algebraic sum of inter-crossings in the projection of the two curves in any projection direction, it is an integer and a topological invariant of the link.

For open curves, the Gauss linking integral is equal to the average of half the algebraic sum of inter-crossings in the projection of the two curves over all possible projection directions. It is a real number and a continuous function of the curve coordinates.

1.1.1 Finite form of the Gauss linking integral

In [2], a finite form for the Gauss linking integral of two edges was introduced, which gives a finite form for the Gauss linking integral over two polygonal curves.

Let E_n, R_m denote two polygonal curves of edges $e_i, i = 1, \ldots, n, r_j, j = 1, \ldots, m$, then

$$L(E_n, R_m) = \sum_{i=1}^{n} \sum_{j=1}^{m} L(e_i, r_j) \tag{2}$$

where $L(e_i, r_j)$ is the Gauss linking integral of two edges. Let e_i be the edge that connects the vertices $\vec{p}_i, \vec{p}_{i+1}$ and r_j be the edge that connects the vertices $\vec{p}_j, \vec{p}_{j+1}$ (see Figure 1 for an illustrative example). In [2] it was shown that $L(e_i, r_j) = \frac{1}{4\pi} Area(Q_{i,j})$, where Q_{ij} for $i < j$ denotes the quadrangle defined by the faces of the quadrilateral formed by the vertices $\vec{p}_i, \vec{p}_{i+1}, \vec{p}_j, \vec{p}_{j+1}$. This area can be computed by adding the dihedral angles of this quadrilateral. The faces of this quadrangle have normal vectors $\vec{n}_i, i = 1, \ldots 4$, defined as follows [4]:

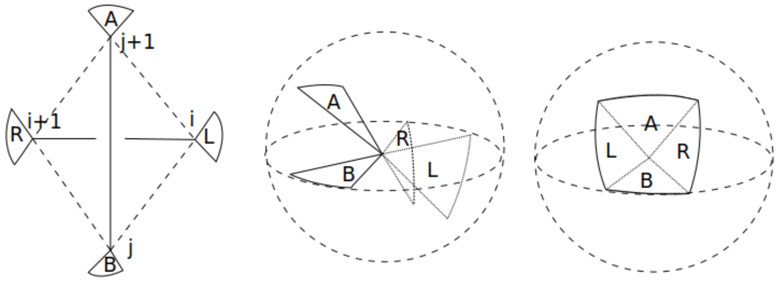

Figure 1: The area of the quadrangle is bounded by the great circles with normal vectors $\vec{n}_1, \vec{n}_2, \vec{n}_3, \vec{n}_4$, determined by the faces of the quadrilateral. In fact, the quadrangle is formed by gluing together, with correct orientation the tiles A, B, R, L. The vectors vectors $\vec{n}_1, \vec{n}_2, \vec{n}_3, \vec{n}_4$ are perpendicular to the tiles L, A, R and B respectively, pointing outwards of the tetrahedron for A, B and inwards for L, R. These tiles define a quadrangle with faces A, L, B, R in the counterclockwise orientation, with all the normal vectors pointing outside the quadrangle.

$$\vec{n}_1 = \frac{\vec{r}_{i,j} \times \vec{r}_{i,j+1}}{||\vec{r}_{ij} \times \vec{r}_{i,j+1}||}, \vec{n}_2 = \frac{\vec{r}_{i,j+1} \times \vec{r}_{i+1,j+1}}{||\vec{r}_{i,j+1} \times \vec{r}_{i+1,j+1}||}, \vec{n}_3 = \frac{\vec{r}_{i+1,j+1} \times \vec{r}_{i+1,j}}{||\vec{r}_{i+1,j+1} \times \vec{r}_{i+1,j}||}, \vec{n}_4 = \frac{\vec{r}_{i+1,j} \times \vec{r}_{i,j}}{||\vec{r}_{i+1,j} \times \vec{r}_{i,j}||}$$

where $\vec{r}_{ij} = \vec{p}_i - \vec{p}_j$, $\vec{r}_{i,j+1} = \vec{p}_i - \vec{p}_{j+1}$, $\vec{r}_{i+1,j} = \vec{p}_{i+1} - \vec{p}_j$, $\vec{r}_{i+1,j+1} = \vec{p}_{i+1} - \vec{p}_{j+1}$.

The area of the quadrangle Q_{ij} is: $Area(Q_{ij}) = \arcsin(\vec{n}_1 \cdot \vec{n}_2) + \arcsin(\vec{n}_2 \cdot \vec{n}_3) + \arcsin(\vec{n}_3 \cdot \vec{n}_4) + \arcsin(\vec{n}_4 \cdot \vec{n}_1)$.

2 The Writhe

A measure for the degree of intertwining of the chain around itself is the Writhe, which is defined by taking the Gauss linking integral over one curve (instead of two)

(Writhe). For a curve ℓ with arc-length parameterization $\gamma(t)$ is the double integral over l:

$$Wr(l) = \frac{1}{4\pi} \int_{[0,1]} \int_{[0,1]} \frac{(\dot{\gamma}(t), \dot{\gamma}(s), \gamma(t) - \gamma(s))}{||\gamma(t) - \gamma(s)||^3} dtds. \tag{3}$$

The Writhe is a continuous function of the chain coordinates for both open and closed chains.

By taking the absolute value of the integrand, we obtain the Average Crossing Number, ACN. This is another informative measure which shows how many crossings we see in a random projection of the protein.

References

[1] Q. Baldwin and E. Panagiotou. The local topological free energy of proteins. *J. Theor. Biol. (submitted)*, 2021.

[2] T. Banchoff. Self-linking numbers of space polygons. *Indiana Univ. Math. J.*, 25:1171–1188, 1976.

[3] K. F. Gauss. *Werke*. Kgl. Gesellsch. Wiss. Göttingen, 1877.

[4] K. Klenin and J. Langowski. Computation of writhe in modelling of supercoiled DNA. *Biopolymers*, 54:307–17, 2000.

[5] E. Panagiotou and K. W. Plaxco. A topological study of protein folding kinetics. *Topology and Geometry of Biopolymers, AMS Contemporary Mathematics Series*, 746:223–233, 2020.