Studying Linking Number in Book Embeddings

Eric Burkholder, Emma Harrell, Elijah Leake

Summer 2021

Contents

1 Introduction

In past decades, knot theory has frequently been used to further our understanding and knowledge of biology and molecules. DNA molecules are especially inclined to this sort of study. Biologists and chemists use knot theory to study the topological and geometric properties of DNA. DNA is composed of two long strands called backbones (made of sugar and phosphate molecules) which twist around one another to form a double helix [6] [10]. DNA resides inside the nucleus of a cell, but a DNA molecule is 1,000 to 500,000 times the length of the diameter of the nucleus [5]. With such a long molecule stuffed into such a compact space, it is unsurprising that DNA can become knottted, tangled, and linked [6]. In order for cell replication to occur, DNA must unknot itself. To help DNA unknot itself, a special enzyme known as topoisomarase cuts through the knotted parts of the DNA molecule without changing any intrinsic part of the DNA, and reconnects any loose ends [6] [8]. Problems can easily arise during cellular replication, especially if topoisomarase enzymes do not work properly. By comparing the topological invariants of DNA before and after enzymes act on it, we can learn more about mechanisms of these enzymes and their effects on the structure of DNA [7]. In order to distinguish the different topological forms (knots) that DNA takes, we consider invariants, such as the linking number. The intersection of knot theory and biology is a very active field. For example, this field can be extended to cancer research and the development of Type I topoisomerase inhibiting treatments. Although Type I topoisomerase helps during cellular replication, these enzymes can also help cancerous cells to grow, which is very problematic [5]. Mathematicians and biologists are frequently making new discoveries towards cancer research and other areas that stand at the intersection of biology and knot theory.

Spatial embeddings of graphs have been studied since the 1980s. Of particular influence is Conway and Gordon's 1983 paper in which they proved that any spatial embedding of K_6 contains a non-trivial link, and any spatial embedding of K_7 contains a non-trivial knot [3].

Various models of random knots and links have been developed. Arsuaga et al. studied the mean squared linking number of two random polygons generated by the Uniform Random Polygon within a confined space in their paper [1]. Additionally, Even-Zohar et al. have studied the Petaluma model, and have found a distribution for the linking number [4].

The linking number of book embeddings has also been studied. Rowland classified all possible links that could appear in book embeddings of K_6, as well as the connection between linking number and sheet number [9].

We expand on these results by studying the distribution of the linking number in randomly generated book embeddings of K_n. We begin by introducing book embeddings, and provide our exact and approximate linking number calculations for book embeddings. We then show that the mean squared linking number of two polygons of length n in a book embedding is $\frac{1}{2}n^2q$, where q is a constant. Finally, we produce calculations for the relative frequency of a linking number between two monotonic polygons, and investigate the maximum linking number between two polygons.

2 Key Concepts

Definition 1. *A graph $G = (V, E)$ is a collection of V vertices and E edges connecting them.*

Definition 2. *A complete graph K_n is a graph with n vertices such that each vertex shares an edge with every other vertex.*

Definition 3. *A spatial embedding of a graph places the graph into a 3-dimensional space such that the vertices are points in $\mathbb{R}^3$ and the edges are non-intersecting curves.*

Definition 4. *A link L is two or more disjoint components or loops in 3-dimensional space.*

The two components of a link may cross each other.

Definition 5. *A crossing is positive if, from the perspective of the top strand, the bottom strand moves from right to left.*

Definition 6. *A crossing is negative if, from the perspective of the top strand, the bottom strand moves from left to right.*

A positive and negative crossing are illustrated in Figure 1.

The linking number, ℓk, of a link is an invariant which measures how intertwined two components of a link are.

Definition 7. *The linking number is one half the number of positive crossings between two components minus the number of negative crossings between the components, or*

$$\ell k(L) = \frac{1}{2}((\textit{number of positive crossings}) - (\textit{number of negative crossings}))$$

For Definitions 8, 9, and 10, assume that we place our vertices on a circle, and label the vertices $1, 2, 3, \ldots$ in a clockwise direction.

Definition 8. *Given a polygon with edges* $\{\overrightarrow{a_1a_2}, \overrightarrow{a_2a_3}, ..., \overrightarrow{a_{n-1}a_n}\}$*, such that when* a_1 *is the smallest vertex in the polygon, that polygon is strictly increasing if for all* $i \in n$*,* $a_i < a_{i+1}$*.*

Essentially, a strictly increasing polygon whose edges are always oriented from a smaller vertex to a larger vertex.

Definition 9. *Given a polygon with edges* $\{\overrightarrow{a_1a_2}, \overrightarrow{a_2a_3}, ..., \overrightarrow{a_{n-1}a_n}\}$*, such that when* a_1 *is the largest vertex in the polygon, that polygon is strictly decreasing if for all* $i \in n$*,* $a_i > a_{i+1}$*.*

Definition 10. *A polygon is monotonic if it is either strictly increasing or strictly decreasing.*

For example, in Figure 2, the quadrilateral on the left is monotonic, because if we list the vertices in order of orientation, starting from 1, we get 1234, which is strictly increasing. However, if we did the same for the quadrilateral on the right, we get 1324, which is not monotonic, since 1 to 3 is increasing, 3 to 2 is decreasing, and 2 to 4 is increasing, making it impossible to always be increasing or decreasing no matter the starting vertex. Another way to think about it is that a monotonic polygon will never cross itself.

Figure 1: A positive crossing (left) and a negative crossing (right)

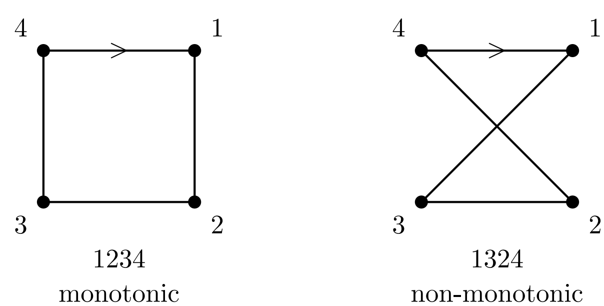

Figure 2: Monotonic (left) and Non-Monotonic (right) Polygons

3 Book Embeddings

A book embedding places the vertices of a graph into a spine and the edges of a graph into discrete sheets such that if two edges are on the same sheet, they do not cross [2] . The spine of a book embedding is classically a straight line, the pages half-planes, and edges semi-circles. However, book embeddings have also been studied as circles, by way of curving the spine such that it forms a circle. Then, the spine becomes a circle, the pages discs, and the edges straight lines. In this paper, we will consider circular book embeddings. We focus on book embeddings of K_n. For our model, n vertices will be placed equidistantly on a circle and labeled from 1 to n clockwise. The perimeter of the circle will form the edges between vertex i and $i+1$ for all $i \in n$ (take the labels modulo n). These are exterior edges, and they do not cross any other edge. The remaining edges are drawn inside the circle. These are interior edges. Since there are $\binom{n}{2}$ edges in a complete graph K_{2n}, there are $\binom{n}{2} - n$ interior edges. To generate a random spatial embedding of K_n, we place each of the $\binom{n}{2} - n$ interior edges into its own sheet, and then generate a random permutation π to determine the heights of the sheets.

Once a book embedding has been generated, we consider pairs of cycles within the graph. For example, there are $\frac{1}{2}\binom{6}{3} = 10$ pairs of 3-cycles within K_6 [3].

Theorem 1. *The number of pairs of cycles in a book embedding of K_n which use all n vertices is given by*

$$f_n = \frac{1}{2}\sum_{i=3}^{n-3}\binom{n}{i}\frac{(i-1)!}{2}\cdot\frac{(n-i-1)!}{2}$$

Proof. Say we have n vertices and we wish to draw Cycle A and Cycle B with these these vertices. So, we will have i vertices in Cycle A and $n-i$ vertices in Cycle B. There are $\binom{n}{i}$ ways to choose which of the vertices will be in Cycle A.

Then, there are different ways to order the vertices of both cycles A and B. If we have a vertices, it doesn't matter which vertex we start drawing the cycle with since all will be included, and choosing the direction of the vertices also has no affect. So, there are $\frac{(a-1)!}{2}$ ways to draw a cycle with a vertices.

Therefore, there are $\frac{(i-1)!}{2}$ ways to draw Cycle A and $\frac{(n-i-1)!}{2}$ ways to draw Cycle B.

Then, we need to iterate over the different sizes of polygons we can have. We must have that both polygons have at least three vertices, so we iterate from $i=3$ to $n-3$. Then, notice that we are double counting. When $i \neq n-i$, for a size of a cycle s, we will consider this both when $i=s$ and when $n-i=s$. However, since the cycles are arbitrarily labelled, these will count the same cases. As well, when $i=n-i$, we are only interested in the ways to select the vertices for half of the possible cycles, as the other half of the cycles will be defined by not being selected. So, we half the entire summation. ☐

Each of these pairs of cycles can be considered as a link, with each cycle forming one of the components. Note that since the edges in book embeddings consist of straight lines, we may refer to these cycles as polygons. Then we can calculate the linking number of all the pairs of links within the embedding. We can repeat this process for all the possible different orderings of the pages of the book embedding. For a graph K_n, there are $(\binom{n}{2}-n)!$ different permutations of the interior edges. We seek to find the linking number distribution for all of these cases.

4 Exact and Approximate Distributions

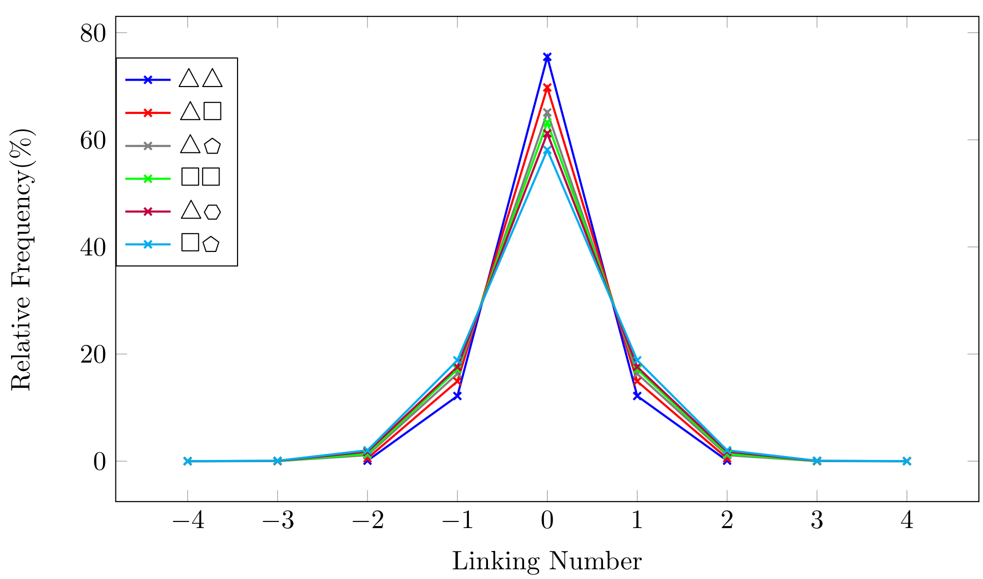

Figure 3: Linking Number Distribution for Two Triangles (K_6), One Triangle and One Quadrilateral (K_7), One Triangle and One Pentagon (K_8), Two Quadrilaterals (K_8), One Triangle and One Hexagon (K_9), and One Quadrilateral and One Pentagon (K_9)

Using code, we find the exact distribution of all possible links in the book embeddings of K_6, K_7, K_8, and K_9. The relative frequency of linking numbers in these cases is shown in Figure 3. Unfortunately, running this code for graphs of size K_{10} or larger is not feasible, due to the factorial nature of the computations.

4.1 Detailed Distribution of K_6

For K_6, there are 10 pairs of cycles (triangles) within the book embedding. The 10 pairs of cycles take one of 3 forms:

1. Two Exterior Edges, One Interior Edge in each triangle. 3 of the 10 pairs are of this type. See the first pair of triangles in Figure 12.
2. One Exterior Edge, Two Interior Edges in each triangle. 6 of the 10 pairs are this type. See the second pair of triangles in Figure 12.
3. Three Interior Edges in each triangle. 1 of the 10 pairs is this type. See the third pair of triangles in Figure 12.

For the first case, the linking number will be 0, since the triangles do not even cross.

For the second case, the linking number is 0 $\frac{2}{3}$ of the time and the linking number is $|1|$ $\frac{1}{3}$ of the time.

For the third case, the linking number is 0 $\frac{33}{60}$ of the time, the linking number is $|1|$ $\frac{26}{60}$ of the time, and the linking number is $|2|$ $\frac{1}{60}$ of the time.

Combining these 3 cases, we find the linking number distribution for a random pair of cycles within a book embedding of K_6. These values are shown in Table 1.

Linking Number	Relative Frequency (%)
-2	0.0833
-1	12.166
0	75.5
1	12.166
2	0.0833

Table 1: Relative Frequencies of Linking Numbers in K_6

4.2 Detailed Distribution for K_7

Within a book embedding of K_7, there are 12 different ways to draw a quadrilateral and a triangle (see Figure 4), and 105 total possible drawings.

Cases 1, 2, 4, 5, and 6 are topologically equivalent to Case 2 from K_6, and thus have a linking number distribution of $P(lk = 0) = \frac{2}{3}$ and $P(lk = |1|) = \frac{1}{3}$. These cases occur a total of 56 out of the 105 drawings.

Cases 7 and 8 are topologically equivalent to Case 3 from K_6, so they have a linking number distribution of $P(lk = 0) = \frac{33}{60}$, $P(lk = |1|) = \frac{26}{60}$, and $P(lk = |2|) = \frac{1}{60}$. These cases occur a total of 14 out of the 105 drawings.

Case 4 has a linking number distribution of $P(lk = 0) = \frac{8}{15}$, $P(lk = |1|) = \frac{6}{15}$, and $P(lk = |2|) = \frac{1}{15}$. This case occurs 7 out of the 105 drawings.

Case 9 has a linking number distribution of $P(lk = 0) = \frac{41}{90}$, $P(lk = |1|) = \frac{44}{90}$, and $P(lk = |2|) = \frac{5}{90}$. This case occurs 7 out of the 105 drawings.

Cases 10, 11, and 12 must all have $lk = 0$, since the polygons do not even cross. These cases occur a total of 21 our of the 105 drawings.

Combining these cases, we find the linking number distribution between a random triangle and quadrilateral within a book embedding of K_7. These values are shown in Table 2.

4.3 Formulas for Calculating Linking Number

The code uses the following formula to calculate the linking number.

Lemma 2. *Let there be two oriented edges, edge a denoted* $\overrightarrow{a_i a_{i+1}}$ *and edge b denoted* $\overrightarrow{b_j b_{j+1}}$*, which cross. Define* $E = \{\overrightarrow{a_i a_{i+1}}, \overrightarrow{b_i b_{i+1}}\}$ *Let* $\pi : E \to S_2$ *be a function which takes the edge* $\overrightarrow{v_i v_{i+1}}$ *to its height, where 2 is the greatest height and 1 is the smallest height. When* $\pi(a_i a_{i+1}) > \pi(b_j b_{j+1})$*, then the sign of their crossing is*

$$\chi = sign((a_i - a_{i+1})(b_j - b_{j+1})((a_i + a_{i+1}) - (b_j + b_{j+1})))$$

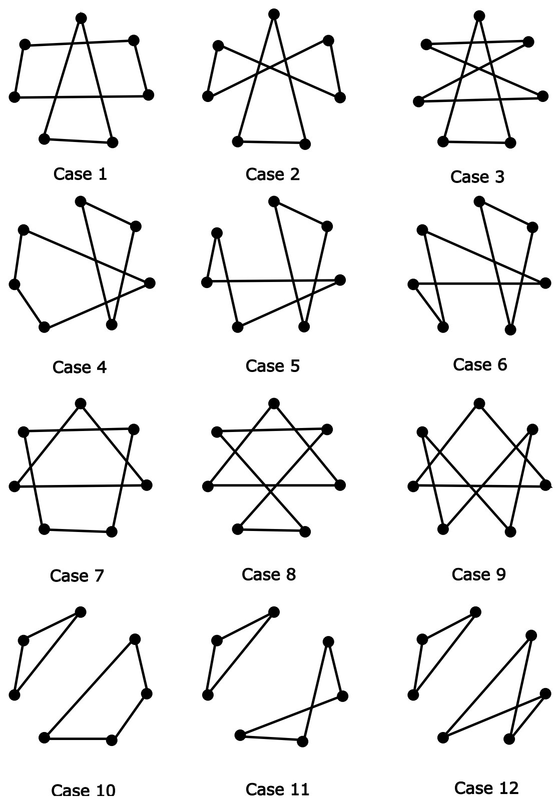

Figure 4: 12 Cases of a Triangle and Quadrilateral

Linking Number	Relative Frequency (%)
-2	0.5185
1	14.5925
0	69.777
1	14.5925
2	0.5185

Table 2: Relative Frequencies of Linking Numbers Between a Triangle and a Quadrilateral

Proof. To see this is true, there are 8 cases we have to consider, depending on the ordering of the endpoints.

From Figure 5, we can see that when if either a_i or a_{i+1} is the largest vertex and $a_i - a_{i+1}$ and $b_j - b_{j+1}$ have the same signs, then the crossing is positive, and if $a_i - a_{i+1}$ and $b_j - b_{j+1}$ have different signs, then the crossing is negative.

However, if either b_j or b_{j+1} is larger than both a_i and a_{i+1}, then the crossing is the opposite sign of $sign((a_i - a_{i+1})(b_j - b_{j+1}))$.

We can account for this by multiplying by the value of $(a_i + a_{i+1}) - (b_j + b_{j+1})$. We can do this because we assume the edges are crossing, meaning that the ordering of the values of the vertices must alternate between the edge a and the edge b, as seen in the different cases of Figure 5. So, if one of the a vertices is the largest, then $a_i + a_{i+1} > b_j + b_{j+1}$, so we'll get a positive value, but if one of the b vertices is the largest, $a_i + a_{i+1} < b_j + b_{j+1}$, so we'll get a negative value. □

Lemma 3. *The cross ratio, CR, of two edges $a_i a_{i+1}$ and $b_j b_{j+1}$ has a negative value when the two edges intersect and a positive value when they do not.*

$$CR(a_i, a_{i+1}, b_j, b_{j+1}) = \frac{b_j - a_i}{b_j - a_{i+1}} \cdot \frac{b_{j+1} - a_{i+1}}{b_{j+1} - a_i}$$

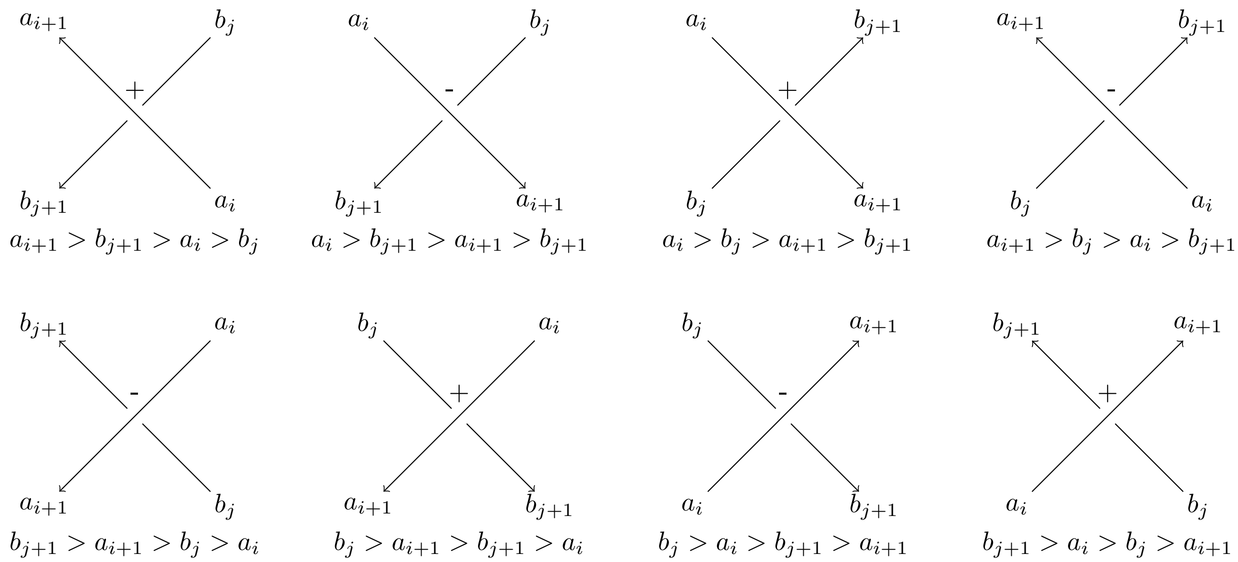

Figure 5: The 8 possible orderings of a_i, a_{i+1}, b_j and b_{j+1}.

Proof. If $a_i a_{i+1}$ and $b_j b_{j+1}$ do not intersect, then the order of the vertices could be any of the following:

1. $a_i > a_{i+1} > b_j > b_{j+1}$
2. $a_{i+1} > a_i > b_j > b_{j+1}$
3. $a_i > a_{i+1} > b_{j+1} > b_j$
4. $a_{i+1} > a_i > b_{j+1} > b_j$
5. $b_j > b_{j+1} > a_i > a_{i+1}$
6. $b_j > b_{j+1} > a_{i+1} > a_i$
7. $b_{j+1} > b_j > a_i > a_{i+1}$
8. $b_{j+1} > b_j > a_{i+1} > a_i$
9. $b_{j+1} > a_i > a_{i+1} > b_j$
10. $b_j > a_i > a_{i+1} > b_{j+1}$
11. $b_{j+1} > a_{i+1} > a_i > b_j$
12. $b_j > a_{i+1} > a_i > b_{j+1}$
13. $a_{i+1} > b_j > b_{j+1} > a_i$
14. $a_i > b_j > b_{j+1} > a_{i+1}$
15. $a_{i+1} > b_{j+1} > b_j > a_i$
16. $a_i > b_{j+1} > b_j > a_{i+1}$

Regardless of the case, $b_j - a_i$ and $b_j - a_{i+1}$ will have the same sign, making $\frac{b_j - a_i}{b_j - a_{i+1}} > 0$, and $b_{j+1} - a_{i+1}$ and $b_{j+1} - a_i$ will have the same sign, making $\frac{b_{j+1} - a_{i+1}}{b_{j+1} - a_i} > 0$. The product of two positives will also be positive. So if $a_i a_{i+1}$ and $b_j b_{j+1}$ do not intersect, $CR(a_i, a_{i+1}, b_j, b_{j+1}) > 0$.

If $a_i a_{i+1}$ and $b_j b_{j+1}$ do intersect, then the order of the vertices follows some alternating structure:

1. $a_i > b_j > a_{i+1} > b_{j+1}$:

   $b_j - a_i < 0$ and $b_j - a_{i+1} > 0$, so $\frac{b_j - a_i}{b_j - a_{i+1}} < 0$, and $b_{j+1} - a_{i+1} < 0$ and $b_{j+1} - a_i < 0$, so $\frac{b_{j+1} - a_{i+1}}{b_{j+1} - a_i} > 0$. Thus $CR(a_i, a_{i+1}, b_j, b_{j+1}) < 0$.

2. $a_{i+1} > b_j > a_i > b_{j+1}$:

   $b_j - a_i > 0$ and $b_j - a_{i+1} < 0$, so $\frac{b_j - a_i}{b_j - a_{i+1}} < 0$, and $b_{j+1} - a_{i+1} < 0$ and $b_{j+1} - a_i < 0$, so $\frac{b_{j+1} - a_{i+1}}{b_{j+1} - a_i} > 0$. Thus $CR(a_i, a_{i+1}, b_j, b_{j+1}) < 0$.

3. $a_i > b_{j+1} > a_{i+1} > b_j$:

   $b_j - a_i < 0$ and $b_j - a_{i+1} < 0$, so $\frac{b_j - a_i}{b_j - a_{i+1}} > 0$, and $b_{j+1} - a_{i+1} > 0$ and $b_{j+1} - a_i < 0$, so $\frac{b_{j+1} - a_{i+1}}{b_{j+1} - a_i} < 0$. Thus $CR(a_i, a_{i+1}, b_j, b_{j+1}) < 0$.

4. $a_{i+1} > b_{j+1} > a_i > b_j$:

   $b_j - a_i < 0$ and $b_j - a_{i+1} < 0$, so $\frac{b_j - a_i}{b_j - a_{i+1}} > 0$, and $b_{j+1} - a_{i+1} < 0$ and $b_{j+1} - a_i > 0$, so $\frac{b_{j+1} - a_{i+1}}{b_{j+1} - a_i} < 0$. Thus $CR(a_i, a_{i+1}, b_j, b_{j+1}) < 0$.

5. $b_j > a_{i+1} > b_{j+1} > a_i$:

   $b_j - a_i > 0$ and $b_j - a_{i+1} > 0$, so $\frac{b_j - a_i}{b_j - a_{i+1}} > 0$, and $b_{j+1} - a_{i+1} < 0$ and $b_{j+1} - a_i > 0$, so $\frac{b_{j+1} - a_{i+1}}{b_{j+1} - a_i} < 0$. Thus $CR(a_i, a_{i+1}, b_j, b_{j+1}) < 0$.

6. $b_j > a_i > b_{j+1} > a_{i+1}$:

   $b_j - a_i > 0$ and $b_j - a_{i+1} > 0$, so $\frac{b_j - a_i}{b_j - a_{i+1}} > 0$, and $b_{j+1} - a_{i+1} > 0$ and $b_{j+1} - a_i < 0$, so $\frac{b_{j+1} - a_{i+1}}{b_{j+1} - a_i} < 0$. Thus $CR(a_i, a_{i+1}, b_j, b_{j+1}) < 0$.

7. $b_{j+1} > a_{i+1} > b_j > a_i$:

   $b_j - a_i > 0$ and $b_j - a_{i+1} < 0$, so $\frac{b_j - a_i}{b_j - a_{i+1}} < 0$, and $b_{j+1} - a_{i+1} > 0$ and $b_{j+1} > a_i < 0$, so $\frac{b_{j+1} - a_{i+1}}{b_{j+1} - a_i} > 0$. Thus $CR(a_i, a_{i+1}, b_j, b_{j+1}) < 0$.

8. $b_{j+1} > a_i > b_j > a_{i+1}$:

   $b_j - a_i < 0$ and $b_j - a_{i+1} > 0$, so $\frac{b_j - a_i}{b_j - a_{i+1}} < 0$, and $b_{j+1} - a_{i+1} > 0$ and $b_{j+1} - a_i > 0$, so $\frac{b_{j+1} - a_{i+1}}{b_{j+1} - a_i} > 0$. Thus $CR(a_i, a_{i+1}, b_j, b_{j+1}) < 0$.

In each of these cases where a_ia_{i+1} and b_jb_{j+1} intersect, $CR(a_i, a_{i+1}, b_j, b_{j+1}) < 0$.

Thus we have shown that $CR(a_i, a_{i+1}, b_j, b_{j+1}) > 0$ when a_ia_{i+1} and b_jb_{j+1} do not intersect, and $CR(a_i, a_{i+1}, b_j, b_{j+1}) < 0$ when a_ia_{i+1} and b_jb_{j+1} do intersect. □

Theorem 4. *Let there be two oriented polygons, A and B, in a book embedding, such that polygon A contains vertices* $\{a_1, a_2, ..., a_n\}$ *and edges* $\{\overrightarrow{a_1a_2}, \overrightarrow{a_2a_3}, ...\overrightarrow{a_{n-1}a_n}, \overrightarrow{a_na_1}\}$ *and polygon B contains vertices* $\{b_1, b_2, ..., b_m\}$ *and edges* $\{\overrightarrow{b_1b_2}, \overrightarrow{b_2b_3}, ...\overrightarrow{b_{m-1}b_m}, \overrightarrow{b_mb_1}\}$*. For* $E = \{\overrightarrow{a_1a_2}, \overrightarrow{a_2a_3}, ...\overrightarrow{a_{n-1}a_n}, \overrightarrow{a_na_1}\} \cup \{\overrightarrow{b_1b_2}, \overrightarrow{b_2b_3}, ...\overrightarrow{b_{m-1}b_m}, \overrightarrow{b_mb_1}\}$*, let the function* $\pi : E \to S_{m+n}$ *take the edge* $\overrightarrow{v_iv_{i+1}}$ *to it's height, such that* $m+n$ *is the highest height and* 1 *is the lowest height. The linking number between two polygons is given by*

$$\ell k(AB) = \frac{1}{2} \sum_{\overrightarrow{a_ia_{i+1}} \in A} \; \sum_{\overrightarrow{b_jb_{j+1}} \in B, \pi(\overrightarrow{a_ia_{i+1}}) > \pi(\overrightarrow{b_jb_{j+1}})} \chi(1 - sign(CR(a_i, a_{i+1}, b_j, b_{j+1})))$$

Proof. By Lemma 3, we know that the cross ratio is negative when two edges cross and is positive when two edges do not cross. Therefore, the value $(1 - sign(CR(a_i, a_{i+1}, b_j, b_{j+1})))$ will be 2 whenever two edges cross and will be 0 whenever two edges do not cross. So, the summation is only affected by edges that cross each other.

Then, when a crossing does occur, by Lemma 2, χ gives us the sign of the crossing between edges a_ia_{i+1} and b_jb_{j+1}, assuming $\pi(a_ia_{i+1}) > \pi(b_jb_{j+1})$. Since our summation includes the same condition that $\pi(a_ia_{i+1}) > \pi(b_jb_{j+1})$, we only count crossings in which χ is accurate.

Then, we multiply the summation by $\frac{1}{2}$, since the quantity $(1 - sign(CR(a_i, a_{i+1}, b_j, b_{j+1})))$ will double count whatever the crossing number actually is.

By only summing this value when the pairs of edges meet the condition $\pi(\overrightarrow{a_i a_{i+1}}) > \pi(\overrightarrow{b_j b_{j+1}})$, we calculate the linking number without needing to multiply by $\frac{1}{2}$. It is known that there must be an even number of crossings between two polygons. So, say that as we move along polygon A, we label the crossings from 1 to $2k$. Then, say the crossings are paired, such that crossings 1 and 2 are a pair, crossings 3 and 4 are a pair, and so on. Then, we consider the different cases of these pairs of crossings.

Whenever a pair of crossings both have polygon A on top or both have polygon B on top, one of the crossing signs will be positive and the other crossing sign will be negative. In either case, we will either not count either crossing, contributing nothing to the summation, or count both crossings, in which we would contribute a total of 0 to the summation.

If one of the crossings has polygon A on top and the other crossing has polygon B on top, then the crossing sign of the two would be the same. We will count exactly one of these crossings. Then, we need not multiply by $\frac{1}{2}$, since we are only going to count half of the positive crossings from this case and half of the negative crossings from this case. □

4.4 Random Code Results

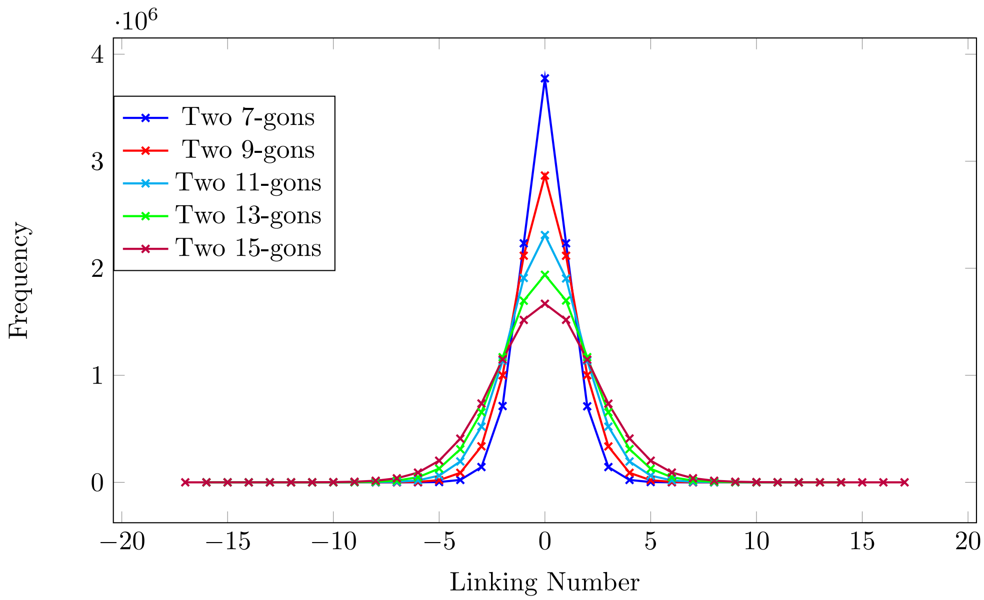

Figure 6: Linking Number Distribution for Two 7-gons, Two 9-gons, Two 11-gons, and Two 13-gons over 10,000,000 Iterations of Random Code

Using random code gives us an approximation of the linking number distribution between m- and n-gons in book embeddings for large values of m and n where our systematic code is not applicable. Figure 6 shows the linking number distribution between two 7-gons, 9-gons, 11-gons, 13-gons, and 15-gons over 10,000,000 iterations of this random code. Observe how the frequency of $lk = 0$ decreases and the frequency of higher linking numbers increases as the polygons get bigger.

4.5 Systematic and Random Code Structures

As mentioned, using the above formulas we developed two codes to discover linking number distributions for different sizes of polygons. We developed both a systematic and randomized simulation.

Systematic Code

For the systematic code, we generate all the the different pairs of disjoint sets of the $m + n$ vertices. Then, for each of these disjoint sets, we generate the possible drawings of these polygons by generating the distinct cycles. Next, for each possible pair of drawings, we generate all of the edges and the distinct edge permutations. Then, for each edge permutation, we consider every possible pair of edges, one from each polygon. Whenever the height of the edge from polygon A is greater than the height of the edge from polygon B and the cross ratio of the two edges is negative, then we calculate the crossing sign between the two edges. We complete this for all of the possible cases, and record cases accordingly. This code may be found at `https://colab.research.google.com/drive/1D4UdBq63p3KYpdoeWdXyQN3p9HwYrX7j?usp=sharing`

Randomized Code

For the randomized code, we divide the $m+n$ vertices randomly between two disjoint sets, and the shuffle the order of these vertices in order to have a random drawing. Then, we generate the edges of these polygons and randomize the order of these edges. Next, we consider every possible pair of edges, one from each polygon. Whenever the height of the edge from polygon A is greater than the height of the edge from polygon B and the cross ratio of the two edges is negative, then we calculate the crossing sign between the two edges. Then, we record the linking number accordingly. We iterate this process 10000000 times in order to get a reasonable sample of the possible polygons and their edge permutations. This code may be found at `https://colab.research.google.com/drive/1eDSgtD_qlUXaTsl04guQv12722cvyGbp?usp=sharing`

5 Mean Squared Linking Number

5.1 Expected Value of Crossing Products

Borrowing ideas from the work of Arsuaga et al. [1], we set out to find the mean squared linking number for book embeddings for two polygons. First, consider the case of two random oriented disjoint edges e_1 and e_2. Since the two edges are disjoint, the probability that these two edges intersect in the projection is a positive number we call $2p$. Like Arsuaga et al., we define a random variable ϵ such that $\epsilon = 0$ when e_1 and e_2 do not intersect, $\epsilon = -1$ when e_1 and e_2 form a negative crossing, and $\epsilon = 1$ when e_1 and e_2 form a positive crossing. Positive and negative crossings are defined in Figure 1.

Since the edges are randomly oriented, we know that $P(\epsilon = 1) = P(\epsilon = -1) = p$. Therefore, we know that $E(\epsilon) = 0$, and $E(\epsilon^2) = (1)^2 \cdot p + (-1)^2 \cdot p = 2p$.

Next, we will consider how a set of four edges e_1, e_2, e_1', and e_2' will interact. Specifically, we are interested in the expected product of the crossing sign of two different intersections. Like Arsuaga et al., we will define ϵ_1 as the crossing sign between edge e_1 and edge e_1' and ϵ_2 as the crossing sign between edge e_2 and edge e_2'. Note that we will consider cases when some of these edges may be equal or share a vertex. The first case of our lemma will be when $e_1 = e_2$ and $e_1' = e_2'$, the case we considered previously.

Lemma 5. *(1) If $e_1 = e_2$ and $e_1' = e_2'$, then $E(\epsilon_1\epsilon_2) = 2p > 0$.*

(2) If e_1, e_2, e_1', and e_2' are all disjoint edges, $E(\epsilon_1\epsilon_2) = 0$.

(3) If $e_1 = e_2$ and e_1' and e_2' are disjoint edges, then $E(\epsilon_1\epsilon_2) = 0$.

(4) If e_1 and e_2 are disjoint and e_1' and e_2' share a vertex, then $E(\epsilon_1\epsilon_2) = 0$.

(5) If e_1 and e_2 share a vertex and e_1' and e_2' share a vertex, then $E(\epsilon_1\epsilon_2) = 0$.

(6) If $e_1 = e_2$ and e_1' and e_2' share a vertex, then $E(\epsilon_1\epsilon_2) = u < 0$. More specifically, if we assume $\epsilon_1 \neq 0$ and $\epsilon_2 \neq 0$, then $E(\epsilon_1\epsilon_2) = -\frac{1}{3}$.

Proof. For all of these cases, if either $\epsilon_1 = 0$ or $\epsilon_2 = 0$, then $\epsilon_1\epsilon_2 = 0$. So, unless otherwise noted, only consider cases in which $\epsilon_1 \neq 0$ and $\epsilon_2 \neq 0$.

(1) In this case, $\epsilon_1 = \epsilon_2$. So, $E(\epsilon_1\epsilon_2) = E(\epsilon_1^2)$. Then, we know that the crossing number squared must be positive. So, if $P(\epsilon_1 = 1) = P(\epsilon_1 = -1) = p > 0$, then $E(\epsilon_1^2) = 2p > 0$.

(2) Consider the current permutation of the edges e_1, e_2, e_1', and e_2' in the book embedding. Then, consider the equally likely permutation in which we switch the edges e_1 and e_1'. Then, ϵ_1 will switch signs, so these two permutations will have opposite values of $\epsilon_1\epsilon_2$. Since this can be done for all permutations of these edges, $E(\epsilon_1\epsilon_2) = 0$.

(3) Consider the orientations of the edges $e_1 = e_2$, e_1', and e_2' in the current embedding. Then, consider the equally likely orientation when e_1' is directed the opposite way. Then, ϵ_1 will switch signs, so these two permutations will have opposite values of $\epsilon_1\epsilon_2$. Since this can be done for all orientations of these edges, $E(\epsilon_1\epsilon_2) = 0$.

(4) Consider the orientations of the edges e_1, e_2, e_1', and e_2' in the current embedding. Then, consider the equally likely orientation when e_1 is directed the opposite way. Then, ϵ_1 will switch signs, so these two permutations will have opposite values of $\epsilon_1\epsilon_2$. Since this can be done for all orientations of these edges, $E(\epsilon_1\epsilon_2) = 0$.

(5) Consider the current permutation of the edges e_1, e_2, e_1', and e_2' in the book embedding. Then, consider the equally likely permutation in which we switch the edges e_1 and e_1'. Then, ϵ_1 will switch signs, so these two permutations will have opposite values of $\epsilon_1\epsilon_2$. Since this can be done for all permutations of these edges, $E(\epsilon_1\epsilon_2) = 0$.

(6) Notice that if we change the orientation of any edge in this case, both ϵ_1 and ϵ_2 change signs, as $e_1 = e_2$ and e_1' and e_2' are adjacent. Consider the different permutations of the book embedding of the three edges. Of these six permutations, two of the permutations have $\epsilon_1\epsilon_2 = 1$, and four of the permutations have $\epsilon_1\epsilon_2 = -1$. So, when we assume $\epsilon_1 \neq 0$ and $\epsilon_2 \neq 0$, we have $E(\epsilon_1\epsilon_2) = \frac{2}{3}(-1) + \frac{1}{3}(1) = -\frac{1}{3}$. Therefore, since we know this is a possible configuration of edges, we may say that $E(\epsilon_1\epsilon_2) = u < 0$.

□

Interestingly, for the uniform random polygon (URP) model, case 5 did not have an expected value of 0. Instead, the approximated value provided for this case was $E(\epsilon_1\epsilon_2) = 0.012 \pm 0.005$. The reason why this case is nonzero for the URP model and not for the book embedding is the case given by Figure 7. While this figure is possible for the URP model, it is not possible for the book embedding model. Consider trying to write the permutation of edges for these four edges. Start with edge e_1. Underneath edge e_1 is edge e_1'. Underneath edge e_1' is edge e_2. Underneath edge e_2 is edge e_2'. Underneath edge e_2' is edge e_1. But this means that edge e_1 is underneath itself, which is impossible in a book embedding permutation.

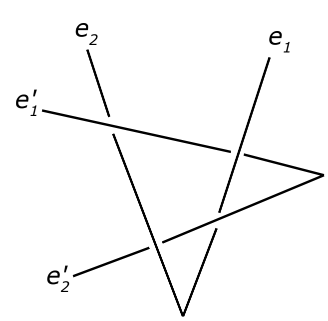

Figure 7: Embedding that is possible for URP model, but impossible for the Book Embedding model

5.2 Mean Squared Linking Number

Let's consider the case of two random polygons R_1 with m edges and R_2 with n edges. Like Arsuaga et al., name the edges of R_1 as $e_1, e_2, \cdots, e_m$, and label the edges of R_2 as $e_1', e_2', \cdots, e_n'$ such that the order of the edges matches the orientation of the polygon. We will say that ϵ_{ij} is the crossing sign between edge e_i and edge e'_j. So, the linking number between the two polygons is $\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}\epsilon_{ij}$. This provides us with the following result.

Theorem 6. *The mean squared linking number between two polygons R_1 with m edges and R_2 with n edges in a book embedding is $\frac{1}{2}nmq$, where $q = p + 2u$.*

Proof. Name the edges of R_1 as $e_1, e_2, \cdots, e_m$, and label the edges of R_2 as $e'_1, e'_2, \cdots, e'_n$ such that the order of the edges matches the orientation of the polygon. Then, the mean squared linking number is given by

$$E\left(\left(\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}\epsilon_{ij}\right)^2\right) = \frac{1}{4}E\left(\left(\sum_{i=1}^{m}\sum_{j=1}^{n}\epsilon_{ij}\right)^2\right) = \frac{1}{4}\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{i'=m}^{n}\sum_{j'=1}^{n}E(\epsilon_{ij}\epsilon_{i'j'}).$$

This summation takes the sum of the all of the products of the ordered pairs of crossings. All of the products of two crossings can be described by the six cases of Lemma 5. So, we may investigate how this breaks down for each case. Indices are taken modulo m for R_1 and modulo n for R_2.

(1) This is the case in which $i = i'$ and $j = j'$. This case is analogous to the terms of the summation $\sum_{i=1}^{m}\sum_{j=1}^{n}E(\epsilon_{ij}^2)$. This summation is equal to $2pnm$ as there are nm terms all with an expected value of $2p$.

(2) This is the case in which $i - i' \geq 2$, $i' - i \geq 2$, $j - j' \geq 2$, and $j' - j \geq 2$, as both pairs of edges are disjoint. We know that the expected value of all of these crossings is 0.

(3) This is the case in which $i = i'$, $j - j' \geq 2$, and $j' - j \geq 2$. This also includes cases in which $i - i' \geq 2$, $i' - i \geq 2$, and $j = j'$. Here, one pair of edges is disjoint, while the other pair of edges are equivalent. We know that the expected values of these crossings is 0.

(4) This is the case in which either $i - i' = 1$ or $i' - i = 1$, $j - j' \geq 2$, and $j' - j \geq 2$. This also includes cases in which $i - i' \geq 2$, $i' - i \geq 2$, and either $j - j' = 1$ or $j' - j = 1$. Here, one pair of edges is disjoint and the other pair of edges share a vertex. We know that the expected value of all of these crossings is 0.

(5) This is the case in which either $i - i' = 1$ or $i' - i = 1$ and either $j - j' = 1$ or $j' - j = 1$. Here, both pairs of edges share a vertex. We know that the expected value of all of these crossings is 0.

(6) This is the case in which $i = i'$ and either $j - j' = 1$ or $j' - j = 1$. This also includes cases in which either $i - i' = 1$ or $i' - i = 1$ and $j = j'$. This case is thus analogous to the terms of the summation $2\sum_{i=1}^{m}\sum_{j=1}^{n}(E(\epsilon_{ij}\epsilon_{(i-1)j}) + E(\epsilon_{ij}\epsilon_{i(j+1)}))$. Thus, we have $2\sum_{i=1}^{m}\sum_{j=1}^{n}(E(\epsilon_{ij}\epsilon_{(i-1)j}) + E(\epsilon_{ij}\epsilon_{i(j+1)})) = 2\sum_{i=1}^{m}\sum_{j=1}^{n}(u+u) = 4unm$.

So, as all other cases have values of 0, we have that

$$\frac{1}{4}\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{i'=m}^{n}\sum_{j'=1}^{n}E(\epsilon_{ij}\epsilon_{i'j'}) = \frac{1}{4}\left(\sum_{i=1}^{m}\sum_{j=1}^{n}E(\epsilon_{ij}^2) + 2\sum_{i=1}^{m}\sum_{j=1}^{n}(E(\epsilon_{ij}\epsilon_{(i-1)j}) + E(\epsilon_{ij}\epsilon_{i(j+1)}))\right) = \frac{1}{4}(2pnm + 4unm) = \frac{1}{2}nm(p+2u) = \frac{1}{2}nmq.$$

□

5.3 Calculating q

Unlike other models, it is actually fairly easy to calculate the value of q in book embeddings.

In order to calculate the value for p, consider four vertices on a circle which define two lines. We only need to consider these four points, as any four vertices may be topologically deformed to have one of the three structures we will consider. Notice that in Figure 8, there are three ways to draw the two lines between these two points. In one of these ways, the two lines intersect. So, the expected probability of an intersection is $E(\epsilon^2 = 1) = 2p = \frac{1}{3}$, so $p = \frac{1}{6}$.

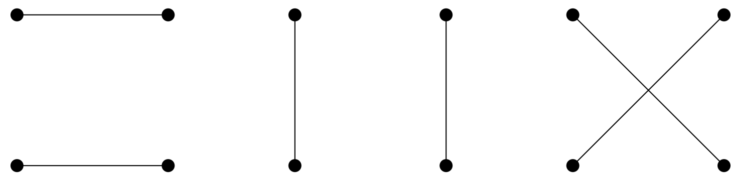

Figure 8: There are three ways to draw two lines given four defined points, one of which has a crossing

In order to calculate the value for u, remember from Lemma 5, Case 6 that when $\epsilon_1 \neq 0$ and $\epsilon_2 \neq 0$, then $E(\epsilon_1\epsilon_2) = -\frac{1}{3}$. Let's find the probability that $\epsilon_1 \neq 0$ and $\epsilon_2 \neq 0$. Consider 5 vertices on a circle like in Figure 9. These five vertices are enough to define an independent edge and two adjacent edges, so any case of an independent edge and two adjacent edges can be topologically deformed to one of the cases we will consider. First, draw the independent edge. As shown in the top left, there are five interior edges that could have crossings (blue) and five exterior edges that cannot have crossings (red). So, $\frac{1}{2}$ of the independent edges could have $\epsilon_1 \neq 0$ and $\epsilon_2 \neq 0$. Then, once we have seleced an interior edge, there are three ways to draw the adjacent edges with the remaining three vertices. One of these cases, the top right case has $\epsilon_1 \neq 0$ and $\epsilon_2 \neq 0$. So, $\frac{1}{3}$ of the choices for the adjacent edges have $\epsilon_1 \neq 0$ and $\epsilon_2 \neq 0$. So, the probability that $\epsilon_1 \neq 0$ and $\epsilon_2 \neq 0$ is $\frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$. Then, when $\epsilon_1 \neq 0$ and $\epsilon_2 \neq 0$ the expected value is $-\frac{1}{3}$, so $u = E(\epsilon_1\epsilon_2) = -\frac{1}{3} \cdot \frac{1}{6} = -\frac{1}{18}$.

Therefore, $q = \frac{1}{6} + 2(-\frac{1}{18}) = \frac{1}{18}$. This means that the mean squared linking number is given by $\frac{1}{2}n^2 \cdot \frac{1}{18} = \frac{1}{36}n^2$.

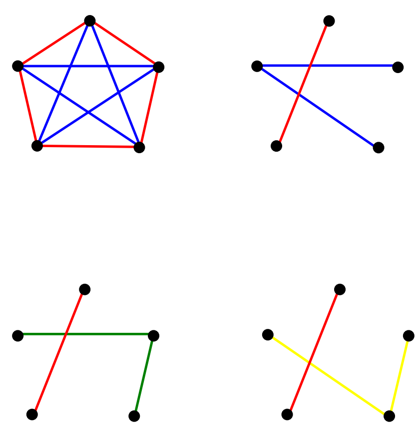

Figure 9: There are five independent edges that could have both crossings, and one of three ways to draw the adjacent edges that have both crossings

6 Linking Number Distribution of Two Monotonic Polygons

In order to calculate the linking number distribution for two monotonic polygons, we found two distributions. The first distribution inputs the number of crossings between the two polygons and outputs the linking number distribution. The second distribution inputs the size of both polygons and outputs the distribution of the number of crossings.

6.1 Linking Number Distribution for $2k$ Crossings

For the distribution from the number of crossings to the linking number distribution, we used Euler's Triangle. Euler's Triangle gives the number of permutations of length x with y ascents. An ascent in permutation π is when $\pi_i < \pi_{i+1}$. We say that $A(x, y)$ gives the value in the x^{th} row and the y^{th} column of Euler's Triangle.

Theorem 7. *The frequency of linking number r of two strictly increasing polygons with $2k$ crossings when $k \geq 2$ is $A(2k-1, r+k-1) = \left\langle {2k-1 \atop r+k-1} \right\rangle$. The relative frequency of a linking number is given by $\frac{A(2k-1,r+k-1)}{(2k-1)!}$ since a row of Euler's triangle sums to $(2k-1)!$.*

Proof. Let A and B be two strictly increasing polygons with $2k$ crossings. If two monotonic polygons have $2k$ crossings, then each polygon has k interior edges, each of which has two crossings. Then we can consider a permutation of these $2k$ edges in order to determine the linking number. We only need to consider the permutation of the interior edges because the heights of the edges are only relevant so far as they give us information about the crossings between the components of the link. Since the exterior edges never cross any other edge, we may ignore their place in the permutation. The height of an exterior edge is irrelevant to the linking number, so it can be ignored. For a given permutation, label the highest/uppermost edge as $2k$, then moving counterclockwise, label the other edges in decreasing order, such that edge i for all $i \in 2k$ has a crossing with edges $i-1$ and $i+1$. (When $i = 2k$, let $i+1 = 1$.) Let $\pi : S_{2k} \to S_{2k}$ be a function that takes the elements of S_{2k} to their heights in the permutation, with $2k$ being the topmost edge, and 1 being the bottom-most edge. Then, the crossing sign between edge i and edge $i+1$ can be determined by $\pi(i)$ and $\pi(i+1)$. When $\pi(i) > \pi(i+1)$, there is a negative crossing. When $\pi(i) < \pi(i+1)$, there is a positive crossing, as seen in Figure 10.

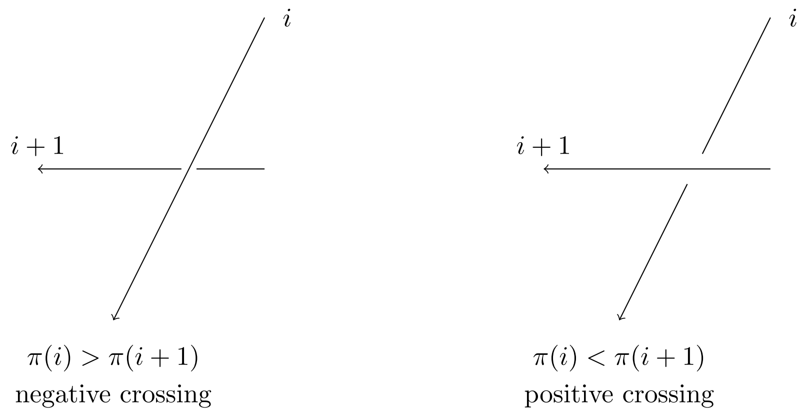

Figure 10: A negative crossing (left) and a positive crossing (right) in terms of $\pi(i)$ and $\pi(i+1)$

Then, the linking number for the link between polygon A and polygon B would be $\frac{1}{2}$ the number of times $\pi(i) < \pi(i+1)$ minus the number of times $\pi(i) > \pi(i+1)$ for all $i \in 2k$.

Observe that when $\pi(i+1) > \pi(i)$, this is equivalent to an ascension in the permutation π, and when $\pi(i+1) < \pi(i)$, this is equivalent to an descent in the permutation π. Since we also consider whether $\pi(2k) > \pi(1)$ or whether $\pi(2k) < \pi(1)$, this is essentially the distribution of the number of ascents in a permutation of length $2k+1$. However, two of these interactions between edges are not entirely independent. By virtue of labeling the top-most edge to be $2k$, that $\pi(2k) = 2k$. Then, regardless the heights of the other edges, $\pi(2k) > \pi(2k+1) = \pi(1)$ and $\pi(2k-1) < \pi(2k)$. Then, since the $\pi(2k)$ is always involved in one ascent and one descent, we know these two interactions will cancel each other out, so we do not have to consider any interactions with $2k$. (Topologically, this would be equivalent to using Reidemeister moves to remove the crossings with the top edge; since the edge is on top, it can be slid away from over everything else, removing its two crossings, which would have been one positive and one negative). Then we are left with $2k-1$ numbers in our permutation. The distribution of the number of ascents in a permutation of length $2k-1$ is given by row $2k-1$ of Euler's Triangle. Thus the linking number distribution of two strictly increasing polygons with $2k$ crossings will have the same distribution as the $2k-1^{\text{th}}$ row of Euler's Triangle.

Now that we have shown the linking number distribution for $2k$ crossings is equivalent to the $2k-1^{\text{th}}$ row of Euler's triangle, we must account for the fact that each row of Euler's triangle starts at a value of 0 (since 0 is the lowest possible number of ascents in a permutation), whereas the linking number distribution is symmetrical around 0 and includes negative values. To do this, we must shift the distribution leftward to the lowest possible linking number in each row. Since there are $2k-1$ entries in the $2k-1^{\text{th}}$ row, we need to shift everything leftwards by $k-1$. So if we want to know the frequency of linking number r between two polygons with $2k$ crossings, we would take the Euler triangle entry of $A(2k-1, r+k-1) = \left\langle {2k-1 \atop r+k-1} \right\rangle$. □

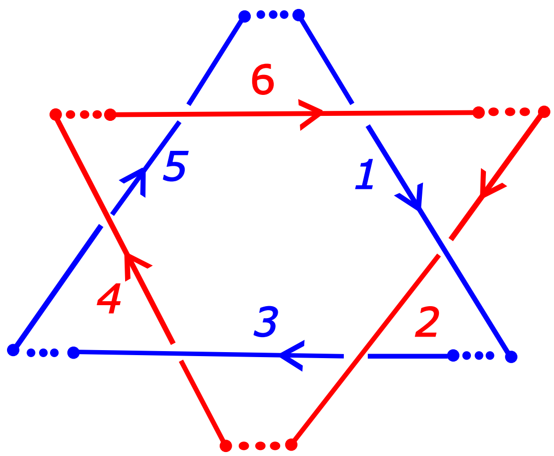

Figure 11: Example Pair of Polygons with 6 Crossings. Note that there are ellipses between vertices to illustrate that there could be any amount of exterior edges between the vertices, but all cases where there are 6 crossings are topologically equivalent to this structure.

S_i	$\pi(S_i)$	crossing sign between S_i and S_{i+1}	ascent or descent between $\pi(S_i)$ and $\pi(S_{i+1})$
1	5	$-$	descent
2	4	$-$	descent
3	3	$-$	descent
4	2	$-$	descent
5	1	$+$	ascent
6	6	$-$	descent

Table 3: Example Permutation of 6 edges

S_i	$\pi(S_i)$	crossing sign between S_i and S_{i+1}	ascent or descent between $\pi(S_i)$ and $\pi(S_{i+1})$
1	5	$-$	descent
2	4	$-$	descent
3	3	$-$	descent
4	2	$-$	descent
5	1		

Table 4: Example Permutation of 6 Edges, Ignoring Interactions With Edge 6

An example of the connection between ascents, descents, crossing signs, and linking number is shown in

Figure 11 and Tables 3 and 4. Observe in Table 3 that $\pi(5) < \pi(6) > \pi(1)$. Thus $\pi(5)$ will have an ascent and $\pi(6)$ will have a descent. Since these two values cancel out, we can consider Table 4, which ignores these values. Then we are left with 4 descents, 4 negative crossings, and a linking number of -2.

Table 5 shows the linking number distribution for 2 polygons with $2k$ crossings for $0 \le k \le 6$, as generated by Euler's Triangle.

Corollary 7.1. *The frequency of linking number r of two monotonic polygons with $2k$ crossings when $k \ge 2$ is $A(2k-1, r+k-1) = \left\langle {2k-1 \atop r+k-1} \right\rangle$. The relative frequency of a linking number is given by $\frac{A(2k-1,r+k-1)}{(2k-1)!}$ since a row of Euler's triangle sums to $(2k-1)!$.*

Proof. Very simply, by changing the orientation of a single polygon in the above proof, we make every positive crossing a negative crossing and vice versa. Therefore, since this will happen to every crossing, the actually linking number distribution will remain unchanged, meaning that our computation is true for monotonic polygons in general. ☐

$k \backslash lk$	-5	-4	-3	-2	-1	0	1	2	3	4	5
0						1					
2					1	4	1				
3				1	26	66	26	1			
4			1	120	1191	2416	1191	120	1		
5		1	502	14608	88234	156190	88234	14608	502	1	
6	1	2036	152637	2203488	9738114	15724248	9738114	2203488	152637	2036	1

Table 5: Linking Number Distribution for Two Polygons with $2k$ Crossings

6.2 Polygon Crossing Distributions

Since the linking number distribution depends solely on the number of crossings between two polygons, and not the number of edges in the polygons, it is pertinent to know how frequently two polygons, one with n vertices and one with m vertices, will have $2k$ crossings. There are two cases, depending on whether $n = m$.

Theorem 8. *The frequency with which a book embedding of two n-gons has $2k$ crossings where $k \ge 2$ is $\binom{n}{n-k}\binom{n-1}{n-k}$.*

Proof. The number of ways to draw two polygons of size n with $2k$ crossings is equivalent to the number of ways to place n identical balls into k of n distinct boxes.

First, we choose k of n boxes. Starting at vertex 1, draw the n edges of one of the polygons. Choose k of these edges to be interior edges, such that the polygon has a general structure, in which we know which edges will cross with the other polygon, but we do not know how the vertices of the other polygon will be distributed. There are $\binom{n}{k} = \binom{n}{n-k}$ ways to build this first polygon.

The k interior edges of the first polygon creates k spaces, or boxes, in which we will place the n vertices of the second polygon. None of the boxes (spaces) can be empty, as this would mean the edge we decided was an interior edge would be an exterior edge, and we would not obtain $2k$ crossings. So, to ensure no box is empty, first we place one ball (or vertex) into each box, or space. Then, we distribute the $n-k$ remaining identical balls/vertices into k distinct boxes, which is simply a stars and bars distribution, which gives us $\binom{(n-k)+k-1}{n-k} = \binom{n-1}{n-k}$.

Since the polygons are the same size, we do not need consider the order that we draw the polygons using this distribution, so combining the distribution, we have $\binom{n}{n-k}\binom{n-1}{n-k}$. ☐

Table 6 gives this crossing number distribution for $3 \le n \le 10$.

Theorem 9. *The frequency with which a book embedding of an m-gon and an n-gon has $2k$ crossings where $k \ge 2$ is*

$$\binom{m}{m-k}\binom{n-1}{n-k} + \binom{n}{n-k}\binom{m-1}{m-k}.$$

when $m \ne n$.

$n \setminus k$	0	2	3	4	5	6	7	8	9	10
3	3	6	1							
4	4	18	12	1						
5	5	40	60	20	1					
6	6	75	200	150	30	1				
7	7	126	525	700	315	42	1			
8	8	196	1176	2450	1960	588	56	1		
9	9	288	2352	7056	8820	4704	1008	72	1	
10	10	405	4320	17640	31752	26460	10080	1620	90	1

Table 6: Crossing Number Distribution for Two n-gons

Proof. We will break down this computation into two cases, depending on which polygon includes vertex 1.

Say we start drawing the polygon with m edges at vertex 1. Draw the m edges of this polygon, and choose k of the edges to be interior edges. such that the polygon has a general structure, in which we know which edges will cross with the other polygon, but we do not know how the vertices of the other polygon will be distributed. There are $\binom{m}{k} = \binom{m}{m-k}$ ways to build this first polygon.

The k interior edges of the first polygon creates k spaces in which we will place the n vertices of the second polygon. None of these spaces can be empty, as this would mean the edge we decided was an interior edge would be an exterior edge, and we would not obtain $2k$ crossings. So, to ensure no space is empty, first we place one vertex into each space. Then, we distribute the $n-k$ remaining vertices into k distinct spaces, which is simply a stars and bars distribution, which gives us $\binom{(n-k)+k-1}{n-k} = \binom{n-1}{n-k}$.

So, there are $\binom{m}{m-k}\binom{n-1}{n-k}$ ways to draw the polygon when the polygon with m vertices includes vertex 1.

We may again go through this process, but instead with the polygon with n vertices including vertex 1. This gives us $\binom{n}{n-k}\binom{m-1}{m-k}$ ways to draw the two polygons.

So, in total there are $\binom{m}{m-k}\binom{n-1}{n-k} + \binom{n}{n-k}\binom{m-1}{m-k}$ ways to draw the polygons of length m and n with $2k$ crossings. □

6.3 Linking Number Distribution Calculation

Using Sections 6.1 and 6.2, we can make a formula to find the linking number distribution based on the size of two polygons n and m. However, to do this, we will redefine our above calculations slightly. Rather than using number of crossings, $2k$, between the two polygons, we will use the number of interior edges $2h$, where each polygon has h interior edges. This is shown in Figure 12 for various numbers of buckets. Note that two polygons always have the same number of interior edges, and any drawings that have the same number of interior edges also have the same number of crossings.

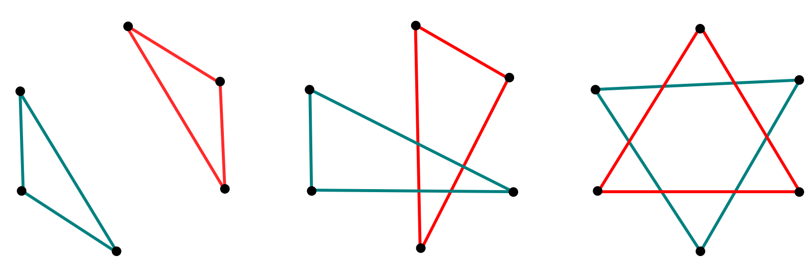

Figure 12: Different drawings of two triangles can have two (left), four (center), or six (right) interior edges

For $k \geq 2$, the number of crossings is equivalent to the number of interior edges. When there are $2k$

crossings, each of the polygons has k interior edges, and there are $2k$ total interior edges. Thus $2k = 2h$ when $k \geq 2$. This means that for our distributions in Theorems 7, 8, and 9 we may replace k with h.

However, when there are 0 crossings there are not 0 but rather 2 interior edges, one for each polygon; see Figure 12. So, for 0 crossings we have $h = 1$. Let's ensure this works for all of our distributions.

For Theorem 7, we want to know the distribution of the linking number when there are 0 crossings or 2 interior edges. The value $\frac{A(2h-1,r+h-1)}{(2h-1)!} = 1$ when $2h = 2$ and $r = 0$. In all other cases, this value is 0. Therefore, the first distribution works for this number of interior edges.

For Theorem 8, when $h = 1$, $\binom{n}{n-h}\binom{n-1}{n-h} = \binom{n}{n-1}\binom{n-1}{n-1} = n$. This is accurate, as when you draw two non-overlapping polygons of the same size, there are n distinct drawings.

For Theorem 9, when $h = 1$, $\binom{m}{m-h}\binom{n-1}{n-h} + \binom{n}{n-h}\binom{m-1}{m-h} = m + n$. This is accurate, as when you draw two non-overlapping polygons of the different sizes, there are $m + n$ distinct drawings.

So, all of the distributions will work appropriately using interior edges, covering all possible numbers of crossings. Note that 2 crossings is not possible for book embeddings.

Therefore, we may arrive at the following results.

Theorem 10. *For two polygons of n vertices, the relative frequency of linking number r is* $D(n,r) = \dfrac{\sum_{h=1}^{n} \frac{A(2h-1,r+h-1)}{(2h-1)!}\binom{n}{n-h}\binom{n-1}{n-h}}{\sum_{h=1}^{n}\binom{n}{n-h}\binom{n-1}{n-h}}$.

Proof. For each of the possible numbers of interior edges, there are $\binom{n}{n-h}\binom{n-1}{n-h}$ distinct drawings, and the relative frequency of linking number r of these drawings is given by $\frac{A(2h-1,r+h-1)}{(2h-1)!}$. Then, since polygons of n edges cannot have more than n interior edges, we sum from $h = 1$ to n. So, we sum this amount, and then divide by the total number of possible drawings, given in the denominator. □

This distribution is illustrated in Figure 13, with exact probabilities listed in Table 7.

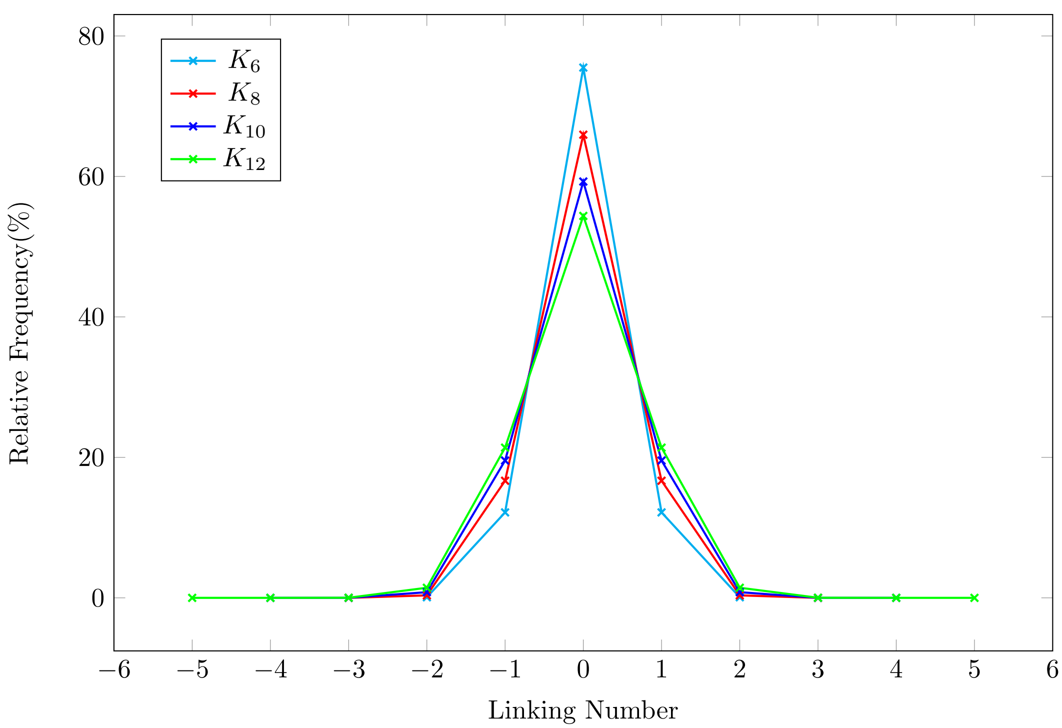

Figure 13: Relative Frequency of Linking Number for Two Monotonic n-gons in K_{2n}

$K_{2n} \backslash lk$	-5	-4	-3	-2	-1	0	1	2	3	4	5
K_6				0.0833	12.1666	75.5	12.1666	0.0833			
K_8			0.00056689	0.3537	16.67517	65.941	16.67517	0.3537	0.00056689		
K_{10}		7.874×10^{-6}	0.004247	0.8067	19.55	59.27	19.55	0.8067	0.004247	7.874×10^{-6}	
K_{12}	0	2.89×10^{-5}	0.01625	1.439	21.389	54.375	21.389	1.439	0.01625	2.89×10^{-5}	0

Table 7: Relative Frequency of Linking Number for Two Monotonic n-gons in K_{2n}. Note that the probability of obtaining a linking number of ± 5 within a book embedding of K_{12} is non-zero, just incredibly small (less than 1 in 18 billion).

Theorem 11. *For two polygons, one with m vertices and the other with n vertices such that $m \neq n$, the relative frequency of linking number r is* $D(m,n,r) = \dfrac{\sum_{h=1}^{n} \frac{A(2h-1,r+h-1)}{(2h-1)!} \left(\binom{m}{m-h}\binom{n-1}{n-h} + \binom{n}{n-h}\binom{m-1}{m-h} \right)}{\sum_{h=1}^{n} \left(\binom{m}{m-h}\binom{n-1}{n-h} + \binom{n}{n-h}\binom{m-1}{m-h} \right)}$.

Proof. Follow the same proof as for Theorem 10. However, instead there are $\binom{m}{m-h}\binom{n-1}{n-h} + \binom{n}{n-h}\binom{m-1}{m-h}$ possible drawings of the two polygons. As well, there can be no more than $\min(m,n)$ interior edges for either of the polygons. So, we can use n as our upper bound and sum from $h = 1$ to n. □

6.4 Investigation of Limits

We also are investigating the limits of linking number distributions for monotonic polygons, so we may have a more general idea about the expected distributions. While we know there exists some trends that a higher number of defined edges means a greater relative frequency of larger linking numbers, we are not entirely sure where these trends will lead.

One of the most fruitful of these investigations has been in regards to a normalized curve of the linking number distribution, mapping relative frequencies to their linking number divided by the size of the polygons of interest. We can see this in Figure 14. We notice that all of these lines seem to follow a similar structure. We hope to eventually show that this structure tends toward a specific limiting distribution.

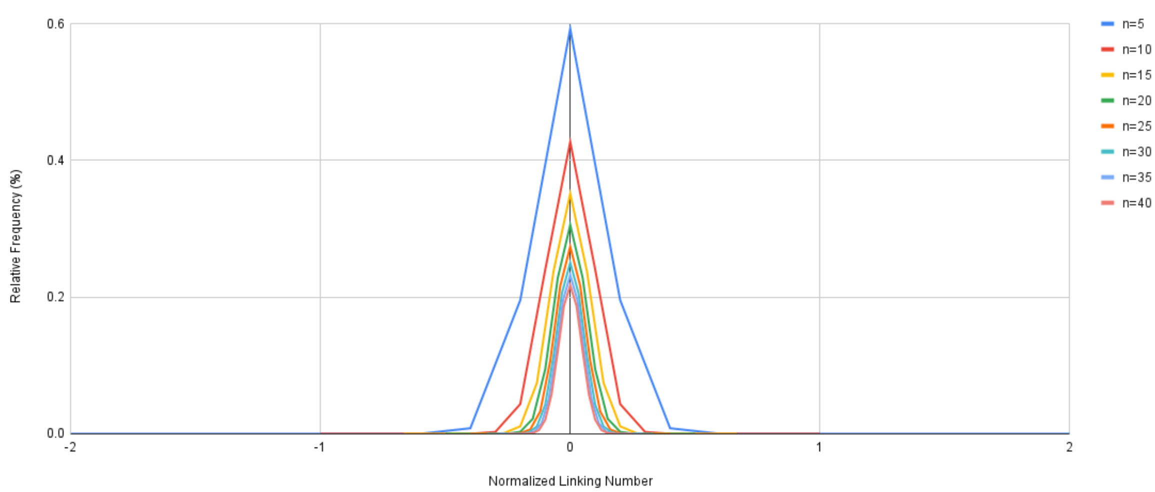

Figure 14: Relative Frequency of Linking Numbers -5 through 5 for Two n-gons

To generate this graph, we wrote code that used Theorem 10 to calculate the linking number distributions for a few linking numbers and sizes of n-gons. This code can be found at `https://colab.research.google.com/drive/1wFj43GavnSxwp7QpGGC8FTdFKSBcRd_l?usp=sharing`, but appears to be having trouble for some large values.

7 Maximum Linking Number

We have found the following bounds on the maximum linking number in a book embedding of a complete graph. We believe that the upper bound could be improved.

Lemma 12. *If two polygons have k crossings which alternate which polygon is the over strand, the linking number between the polygons is $\pm\frac{k}{2}$.*

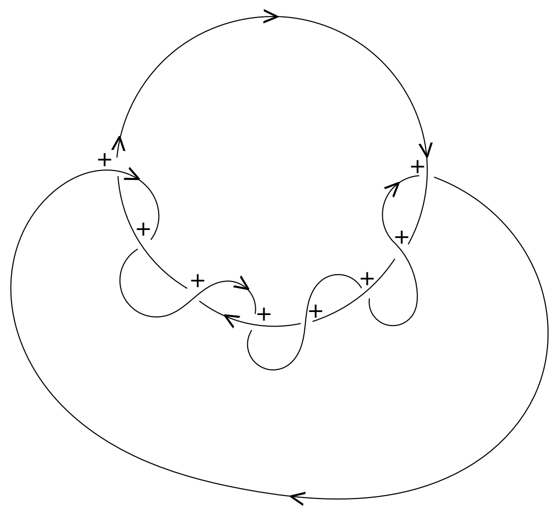

Figure 15: Link with alternating structure

Proof. Let polygon A and polygon B be the two components of a link such that the crossings between them alternate which polygon is the overstrand and which polygon is the understrand. So, consider a localized space of two adjacent crossings. Assume that the first crossing has a value of χ. If we move along polygon A, then at the next crossing, polygon B will now move in the opposite direction with respect to polygon A as before, meaning that the crossing sign would be $-\chi$. However, since the previous under strand is now the over strand, the crossing sign would be $-(-\chi) = \chi$. So the two crossings would have the same sign. We may do this process for any given pair of adjacent crossings, so all of the crossings must have the same value. Then, when we calculate the linking number, we get $\frac{1}{2}k \cdot \chi$, where $\chi = \pm 1$. Thus the linking number for this pair of polygons is $\pm\frac{k}{2}$. □

An example of Lemma 12 can be seen in Figure 15. At each crossing, the component whose strand is the overstrand alternates, causing all of the crossings to have the same sign. Since there are $k = 8$ crossings, all of which are positive, the linking number is $\frac{8}{2} = 4$.

Theorem 13. *The maximum linking number for a pair of cycles in a book embedding of K_n is at least $\left\lfloor \dfrac{n-2}{2} \right\rfloor$ and at most $\dfrac{n^2}{8}$.*

Proof. First we will show that the maximum linking number is at least $\left\lfloor \dfrac{n-2}{2} \right\rfloor$.

Label the vertices of K_n starting with 1 and moving clockwise up to n. Let the odd numbered vertices are part of polygon A, and the even numbered vertices are part of polygon B. Then, our book embedding will be the permutation of edges $13, 24, 35, 46, 57, \cdots$. An example of this structure is seen in Figure 16.

When n is odd, the edge between vertex n and vertex 1 is an exterior edge, and therefore irrelevant to the permutation of the book embedding. Therefore, the $\frac{n-1}{2}$ remaining edges of polygon A will each cross with two edges from polygon B. Therefore, there are $n-1$ crossings between the two polygons. Label the crossing of edges 13 and 24 as crossing c_1, and working clockwise, label the crossings up to c_{n-1}. Then we will have that the odd crossings will have an edge from polygon A going over an edge from polygon B, and the even crossings will have an edge from polygon B going over an edge from polygon A, except for the crossing c_{n-1}, in which edge 13 will go over the edge between vertex $n-1$ and vertex 2. Using Reidemeister Move 2, we may show that c_1 and c_{n-1} do not contribute to the linking number. Then, we are left with $n-3$ crossings which alternate the over and under components. By Lemma 12, we know all of these crossings will have the same sign. Therefore, since all of the crossings will be the same sign, the linking number will be $\frac{n-3}{2}$ which is equal to $\lfloor \frac{n-2}{2} \rfloor$ when n is odd.

When n is even, there are n crossings, as each of the $\frac{n}{2}$ edges from polygon A will cross with two of the edges from polygon B. Label the crossing of edges 13 and 24 as crossing c_1, and working clockwise, label the crossings up to c_n. Then the odd crossings will have an edge from polygon A going over an edge from polygon B, and the even crossings will have an edge from polygon B going over an edge from polygon A, except for the crossing c_n, in which edge 13 will go over the edge between vertex n and vertex 2. Using Reidemeister Move 2, we may show that c_1 and c_n do not contribute to the linking number. Then, we are left with $n-2$ crossings which alternate the over and under components. By Lemma 12, we know all of these crossings will have the same sign. Therefore, since all of the crossings will be the same sign, the linking number will be $\frac{n-2}{2}$.

We have shown that there is always a permutation of the edges of K_n, regardless of the parity of n, such that there exists a link with linking number $\left\lfloor \dfrac{n-2}{2} \right\rfloor$ within the book embedding of K_n. Thus the maximum linking number of a book embedding of K_n is at least $\left\lfloor \dfrac{n-2}{2} \right\rfloor$.

Now we will show that the maximum linking number of a book embedding of K_n is at most $\dfrac{n^2}{8}$.

The maximum linking number between two polygons A and B with a edges and b edges, respectively, would occur when there is the maximum number of crossings between A and B. Since A and B are cycles within a book embedding of K_n, $a+b=n$. The maximum number of crossings would occur when every edge in polygon A crosses every edge in polygon B, which would give a total of ab crossings. Observe that $ab = a(n-a)$ has a maximum when $a = b = \frac{n}{2}$. So the maximum number of crossings would be $\frac{n}{2} \times \frac{n}{2} = \frac{n^2}{4}$.

Given $\frac{n^2}{4}$ crossings, the maximum linking number would occur when all of the crossings had the same sign. This would give us a linking number of $\dfrac{\frac{n^2}{4}}{2} = \dfrac{n^2}{8}$.

Note that if n is odd, the most equitable distribution of edges would give one polygon $\frac{n-1}{2}$ edges and the other $\frac{n+1}{2}$ edges. Since $\frac{n-1}{2} \times \frac{n+1}{2} < (\frac{n}{2})^2$, our upper bound still holds.

Thus we have shown that the upper bound of the maximum linking number between a pair of cycles in a book embedding of K_n is $\dfrac{n^2}{8}$.

□

Theorem 14. *The maximum linking number of a pair of monotonic polygons within a book embedding of* K_n *is* $\lfloor \frac{n-2}{2} \rfloor$.

Proof. By Theorem 7, the linking number distribution for two monotonic polygons with $2k$ crossings is equivalent to Euler's triangle distribution for the row $2k-1$. The $2k-1$th row of Euler's triangle has $2k-1$ entries. One of these entries corresponds to $lk = 0$, then the remaining $2k-2$ entries would be symmetrical. So the maxiumum linking number would be $\frac{2k-2}{2} = k-1$.

There are two cases, depending on the parity of n. When n is even, the maximum number of crossings between two polygons in K_n would occur when each polygon had $\frac{n}{2}$ edges, all of which are interior edges, which would create n crossings. Then, the maximum linking number would be $\frac{n}{2} - 1 = \frac{n-2}{2}$.

If n is odd, then the maximum number of crossings between two polygons in K_n would be when one polygon has $\frac{n-1}{2}$ edges, all of which are interior, and the other has $\frac{n+1}{2}$ edges, all but one of which are interior. This would create $n-1$ crossings. Then the maximum linking number would be $\frac{n-1}{2} - 1 = \frac{n-3}{2} = \lfloor \frac{n-2}{2} \rfloor$.

Since the maximum linking number between two monotonic polygons in K_n is $\frac{n-2}{2}$ when n is even and $\lfloor \frac{n-2}{2} \rfloor$ when n is odd, we conclude that the maximum linking number between two monotonic polygons in K_n is $\lfloor \frac{n-2}{2} \rfloor$. □

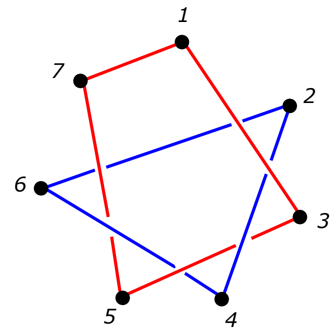

Figure 16: Two polygons in K_7 which have the maximum linking number of 2

8 Conclusion

While we have found the mean squared linking number for all polygons in book embeddings, finding the linking number distribution for all polygons has been a challenge. In future work, we hope to expand our results to include non-monotonic polygons, but we believe this may require a different approach. Additionally, our current distribution for monotonic polygons is not a closed-form expression, so an analysis of the limits of this equation may provide us with more information about how the distribution relates to other calculations.

Intuition as well as results from our systematic and random code suggest that the linking number distribution should approach a uniform distribution, but this is yet to be proved. For specifically monotonic polygons, Figure 14 motivates this aspect of our research. Investigating the limits of our calculation should encompass a significant part of our future work.

Acknowledgments

We would like to thank Dr. Kozai for advising us on this project. This research was funded by NSF Grant DMS-1852132.

References

[1] Javier Arsuaga, T Blackstone, Yuanan Diao, E Karadayi, and M Saito. Linking of uniform random polygons in confined spaces. Journal of Physics A: Mathematical and Theoretical, 40:1925, 02 2007.

[2] Robin Leigh Blankenship. Book embeddings of graphs. PhD thesis, Louisiana State University, 2003.

[3] J. H. Conway and C. McA. Gordon. Knots and links in spatial graphs. J. Graph Theory, 7(4):445–453, 1983.

[4] Chaim Even-Zohar, Joel Hass, Nati Linial, and Tahl Nowik. Invariants of random knots and links. Discrete & Computational Geometry, 56(2):274–314, Jun 2016.

[5] Erica Flapan. Knots, Molecules and the Universe: An Introduction to Topology. American Mathematical Society, United States of America, 2016.

[6] S. Harrell M. Beals, L. Gross. Dna and knot theory. TIEM, 1999, Online.

[7] R. Mishra and S. Bhushan. Knot theory in understanding proteins. J. Math. Biol., 65:1187–1213, 2012.

[8] Numberphile. How DNA unties its own knots nov. 2, 2015. YouTube Video.

[9] D. Rowland. Classification of book representations of k_6. Journal of Knot Theory and Its Ramifications, 26(12):1–26, 2017.

[10] De Witt Sumners. Lifting the curtain: using topology to probe the hidden action of enzymes. Notices of the AMS, 42:528–537, 1995.