On Topology and DNA
- A topological model of DNA replications

A. Mohamad and Tsukasa Yashiro

University of Nizwa, Sultan Qaboos University

Nizwa Seminar
University of Nizwa, Nizwa
Sultanate of Oman
25 February 2016

DNA Replication

Brief History of Knot Theory

In 19th century mathematical study on knots began. Gauss defined the linking number with integral:

$$\mathrm{lk}(K_1, K_2) = \frac{1}{4\pi} \int_{K_1} \int_{K_2} \frac{\mathbf{k}_1 - \mathbf{k}_2}{|\mathbf{k}_1 - \mathbf{k}_2|^3} \cdot (d\mathbf{k}_1 \times d\mathbf{k}_2).$$

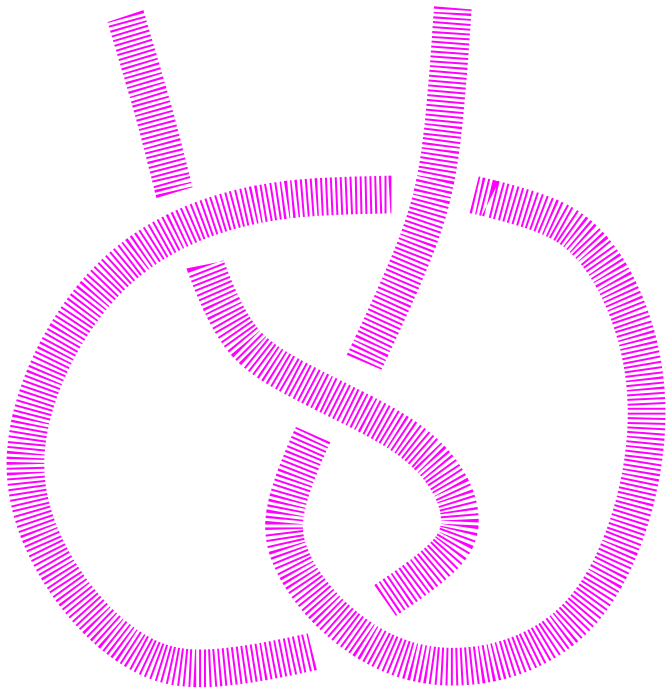
At the time most scientists believed the universe was filled with **ether** and in 1860s, Lord Kelvin proposed that atoms were knots and scientists wanted to study knots. However, as the atomic theory was accepted, scientists lost their interests in knots.

Brief History of Knot Theory

Dveloping the theory of knots was continued by mathematicians. In the early 20th century M. Dehn, J. W. Alexander and others studied knots using knot group. Alexander defined the **Alexander polynomial**.
From late of 1970's to the 1980s the development of the study was boosted. **Jones polynomial** was discovered by V. Jones.
1990s to the present, Knot Theory has been applied to other sciences such as statistical mechanics, biology, etc.

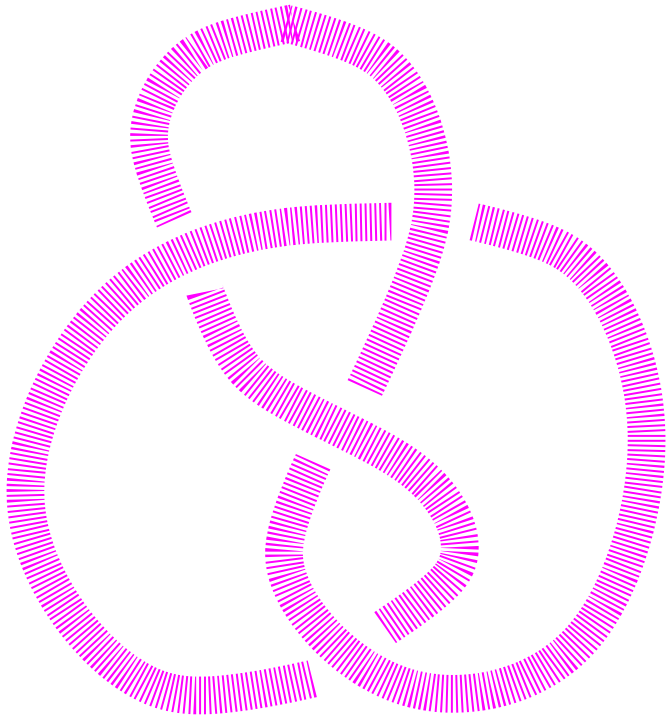
Introduction to Knots

A knot is a tangled string with connected ends. If a knot does not bound a disc in its complement, then it is called a **non-trivial knot**.



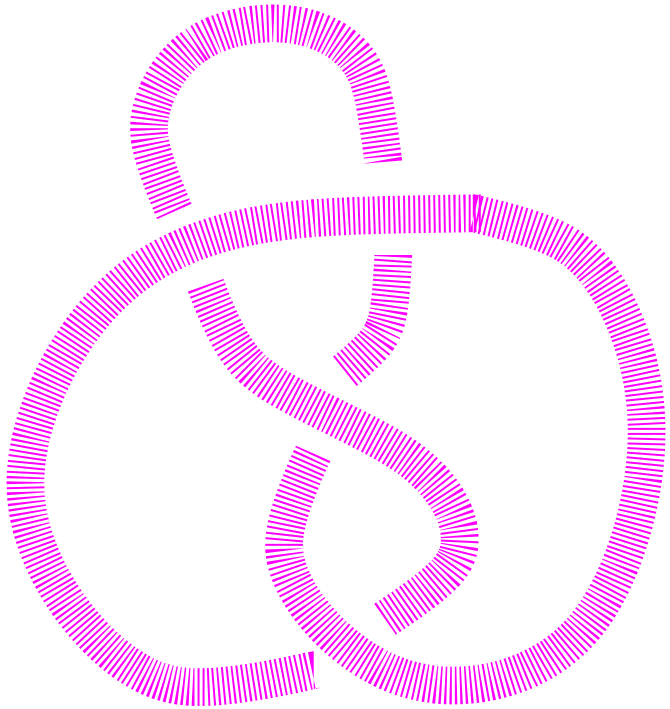
Introduction to Knots

A knot is a tangled string with connected ends. If a knot does not bound a disc in its complement, then it is called a **non-trivial knot**.



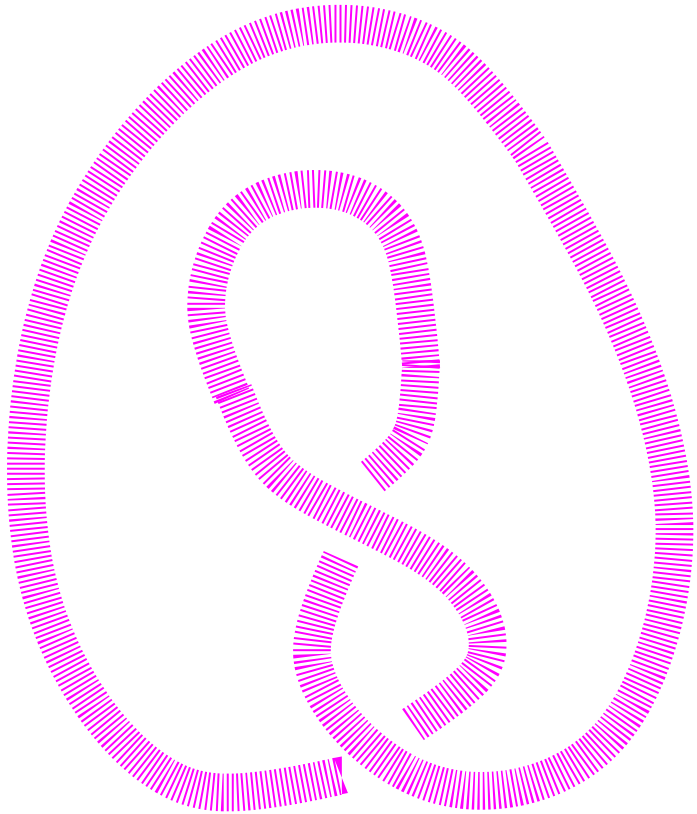
Introduction to knots

If there is a continuous move with the ambient space so that the knot is untied (trivial), then the knot is called a **trivial knot**.



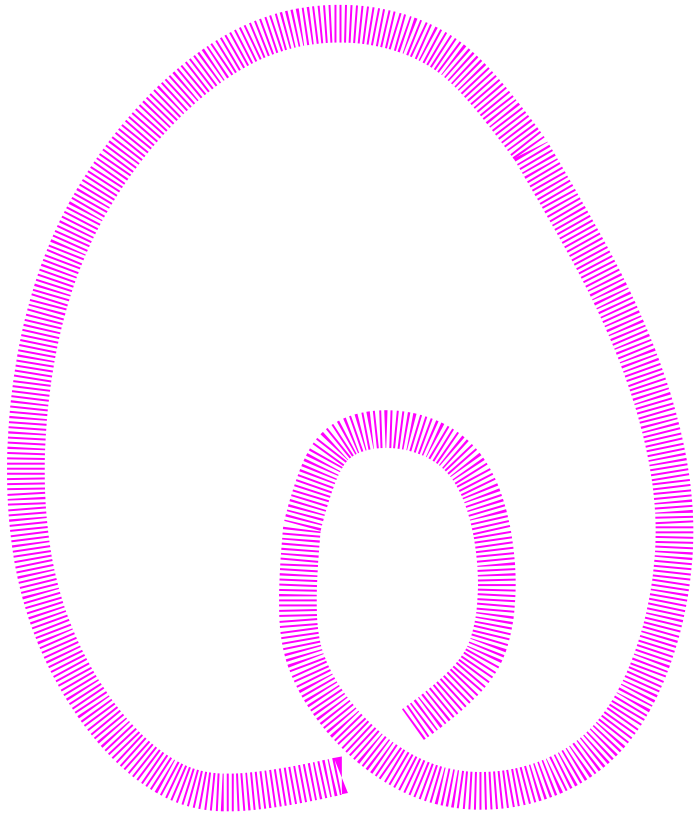
Introduction to knots

If there is a continuous move with the ambient space so that the knot is untied (trivial), then the knot is called a **trivial knot**.



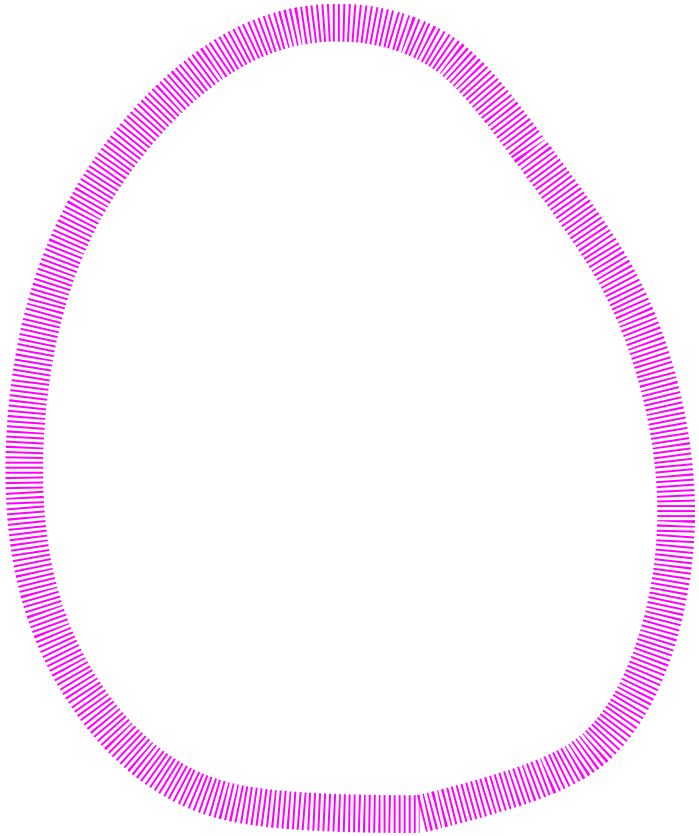
Introduction to knots

If there is a continuous move with the ambient space so that the knot is untied (trivial), then the knot is called a **trivial knot**.



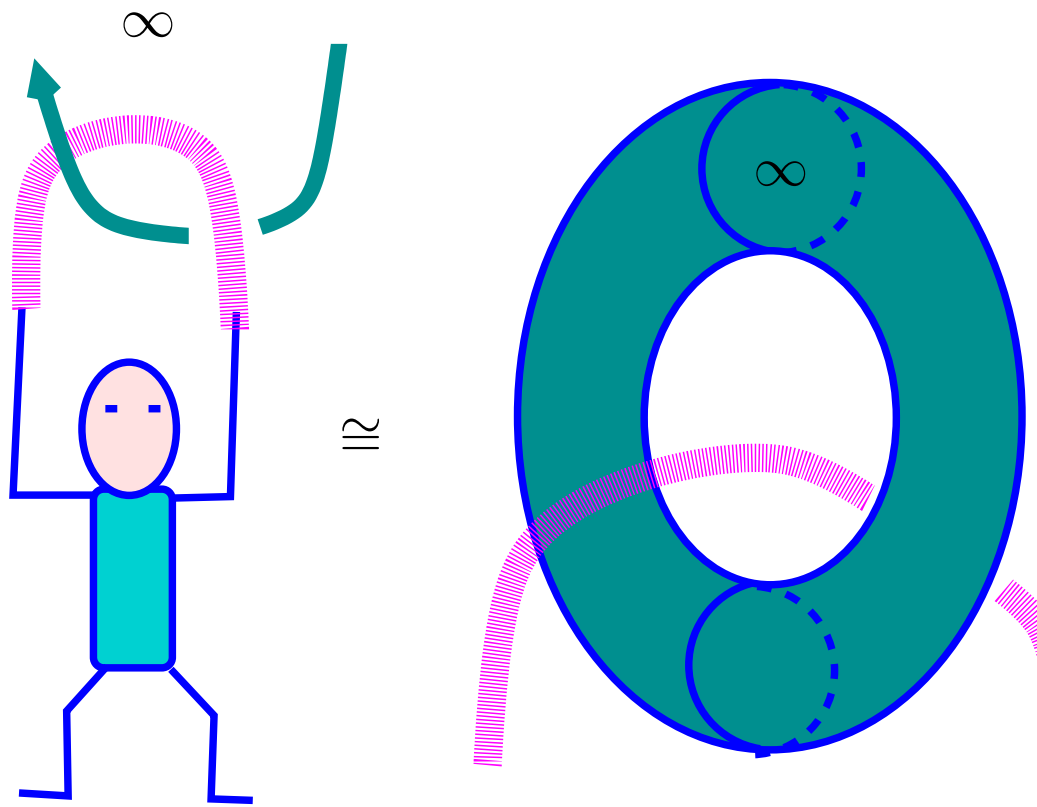
Introduction to knots

If there is a continuous move with the ambient space so that the knot is untied (trivial), then the knot is called a **trivial knot**.



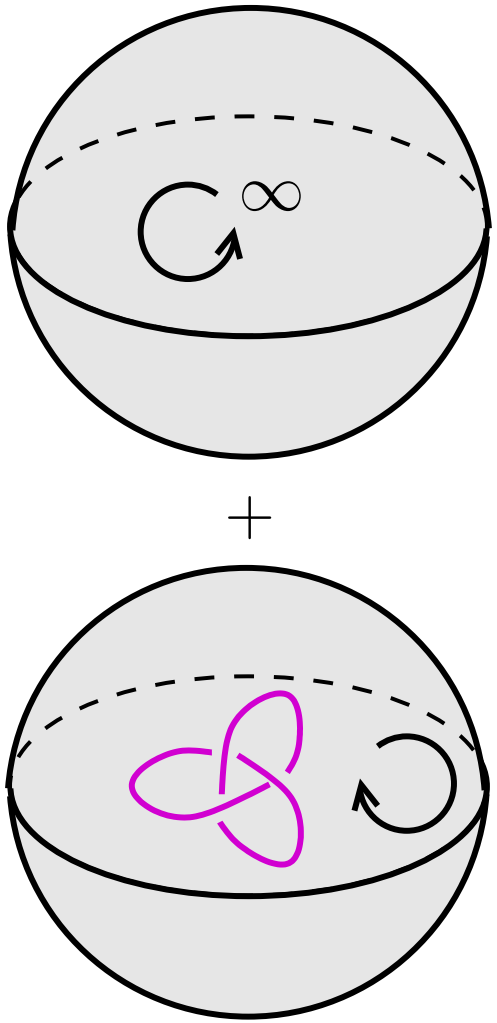
Introduction to knots

The ambient space of a trivial knot in $\mathbb{S}^3 \cong \mathbb{R}^3 \cup \{\infty\}$ is homeomorphic to a solid torus.



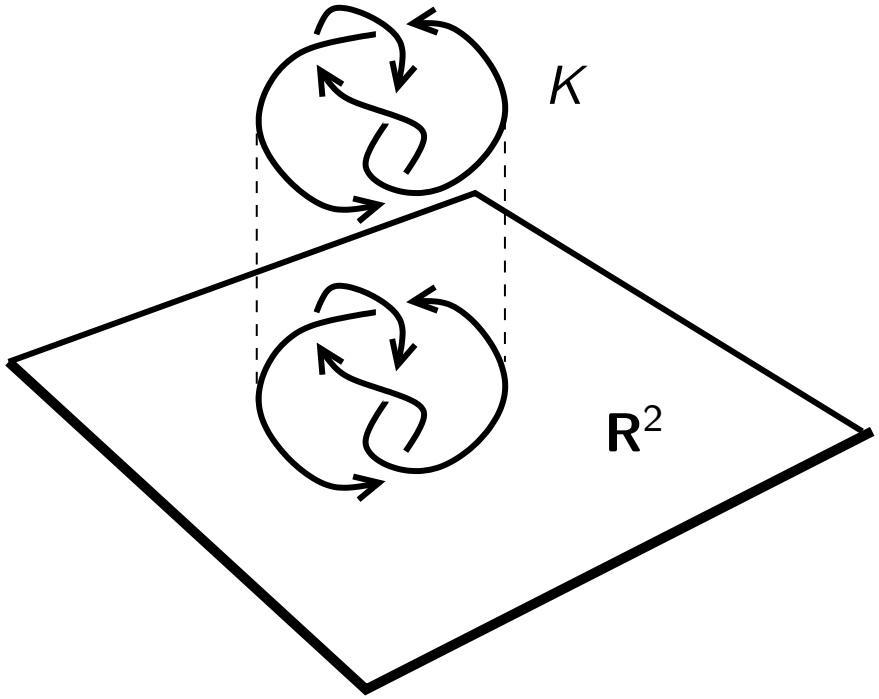
Knots

A **knot** is an embedded finite plygonal circle embedded in $\mathbb{R}^3$ (or $\mathbb{S}^3$). A **link** is a disjoint union of knots.



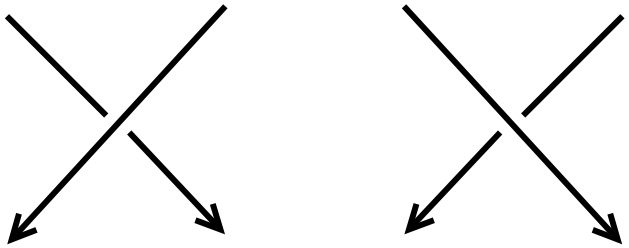
Knots

A **knot diagram** is a projected image of a knot into $\mathbb{R}^2$ with crossing information. We denote it by D_K.



Linking Number

Let D_L be a diagram of L. Add a signnature $\varepsilon(d)$ to each crossing of D as follows.



$+1$ $\quad\quad$ -1

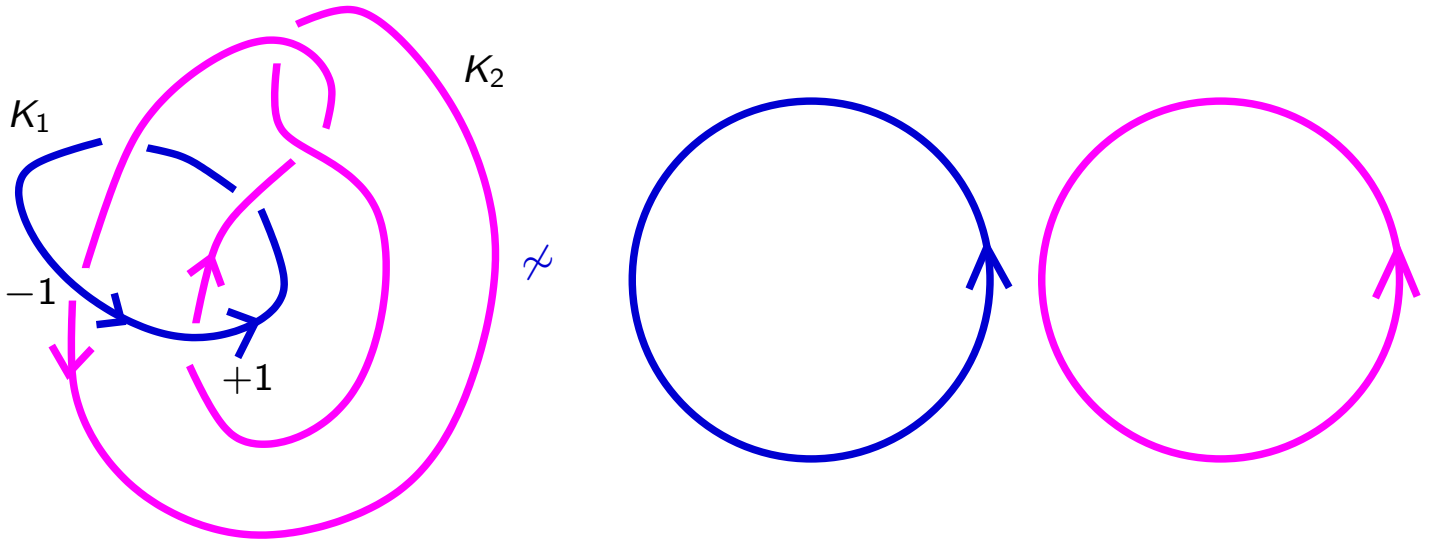
Let $L = K_1 \cup K_2$ be an oriented two-component link. Then we define a function $\mathrm{lk}(K_1, K_2)$ by

$$\frac{1}{2}\sum_{i=1}^{n}\varepsilon(d_i),$$

where $d_1, \ldots, d_n$ are crossings between D_{K_1} and D_{K_2}.

Splittable links

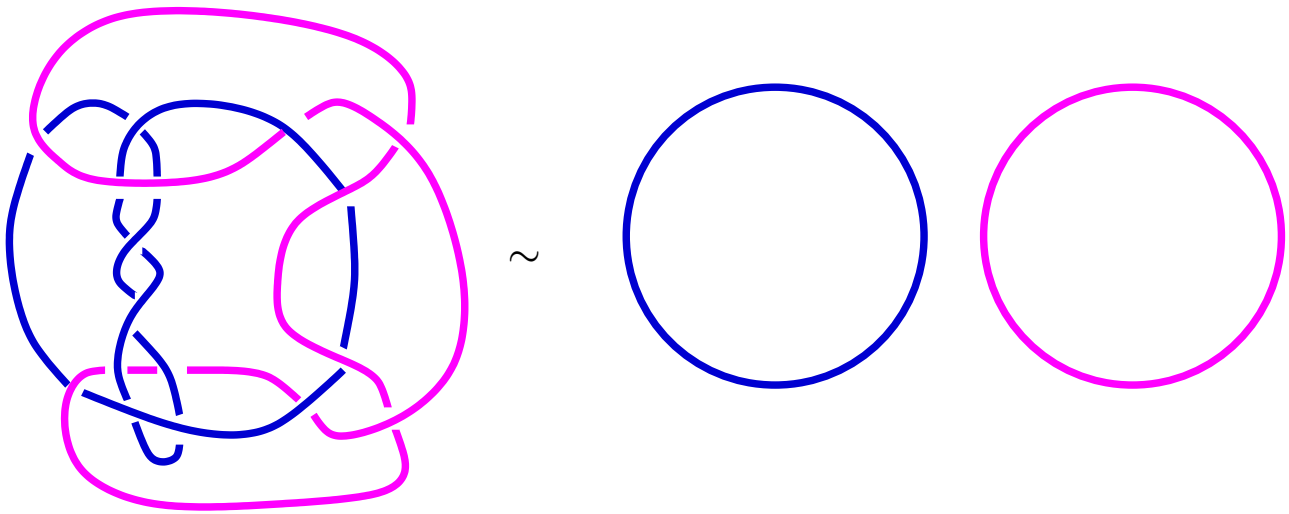
In the following $\mathrm{lk}(K_1, K_2) = +1 - 1 = 0$



The linking number does not tell us if it is splittable or not.

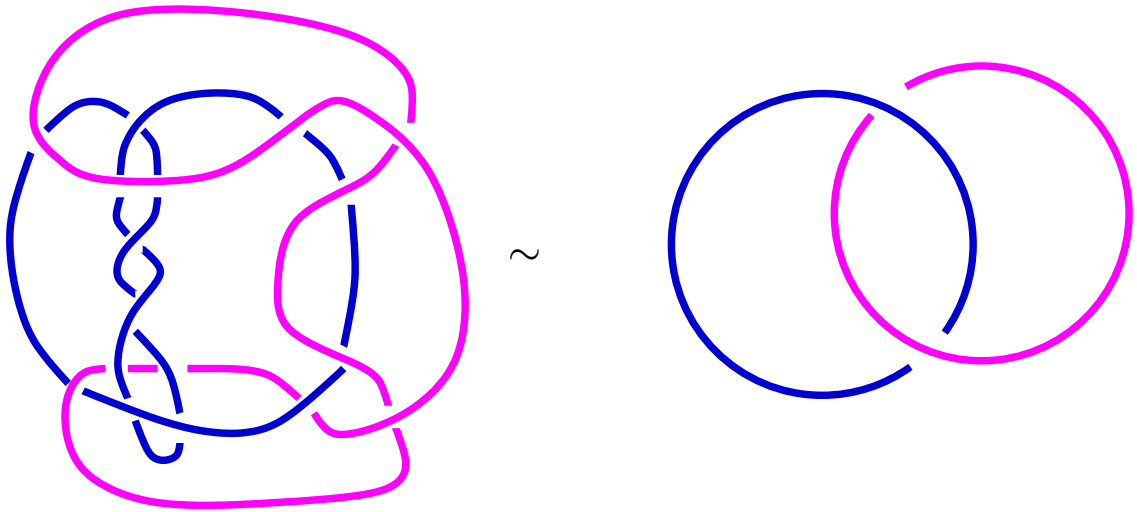
Splittable links

Some link diagrams look very complicated but it is splittable.



Splittable links

This is similar to the previous diagram but it is not splittable.



This suggests us that a small change in the diagram may change the splittability.

Writhe

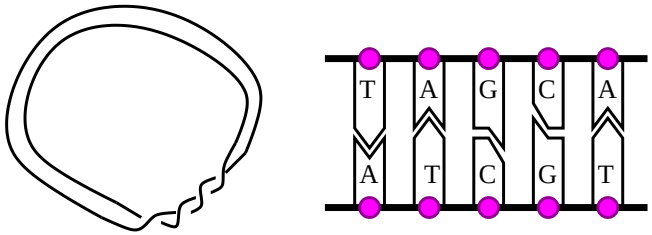
Let D_K be a diagram of a knot K. Then we define a function $\mathrm{Wr}(K)$ by

$$\sum_{i=1}^{n} \varepsilon(c_i),$$

where $c_1, \ldots, c_n$ are crossings of D_K.

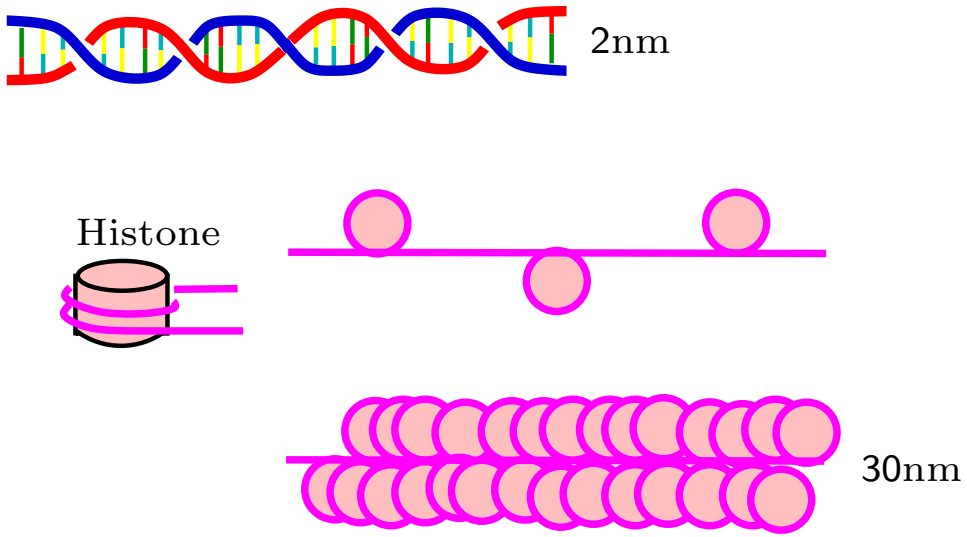
DNA Knots

Most cellular DNA (deoxyribonucleic acid) is double-stranded (duplex) and has a sturucture that has two linear backbones alternating sugar and phosphorus. Each sugar molecule is attached by one of four nucleotide bases: A= Adenine, T = Thymine, C = Cytosine, G = Guamine.



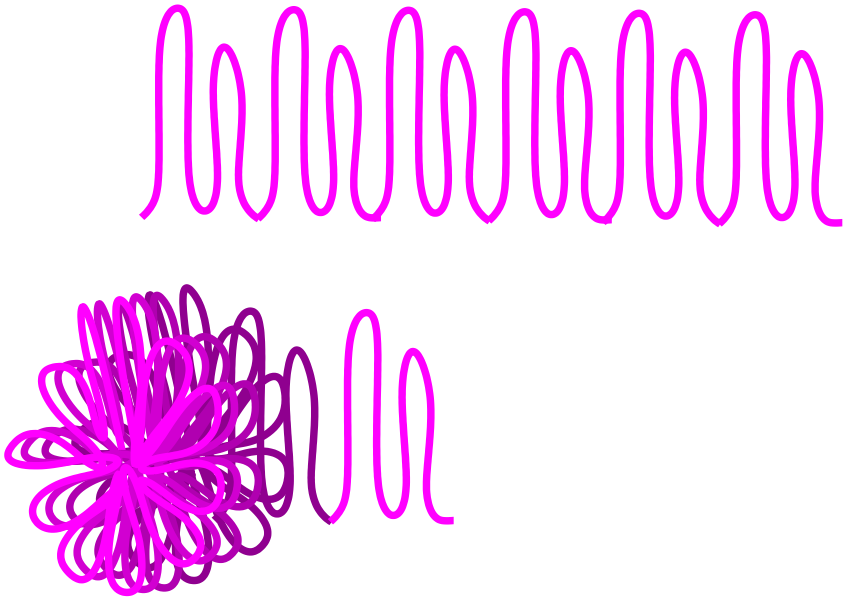
Chromosomes

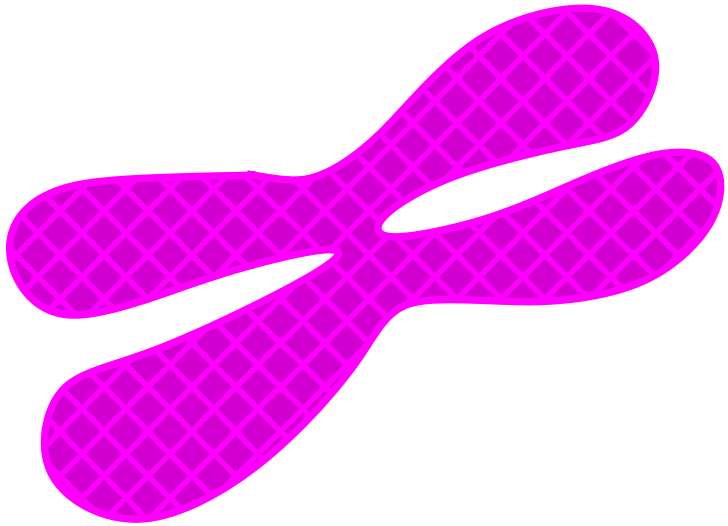
DNA is a double helix strands with the diameter 2 nm (1 nm = $1/10^9$ metre). It forms a winding structure around histones to make a beads structure. Also it forms a 30 nm fibre.



Chromosomes

The 30 nm fibre will form a loop structure and this structure gethers and forms a chromosome.





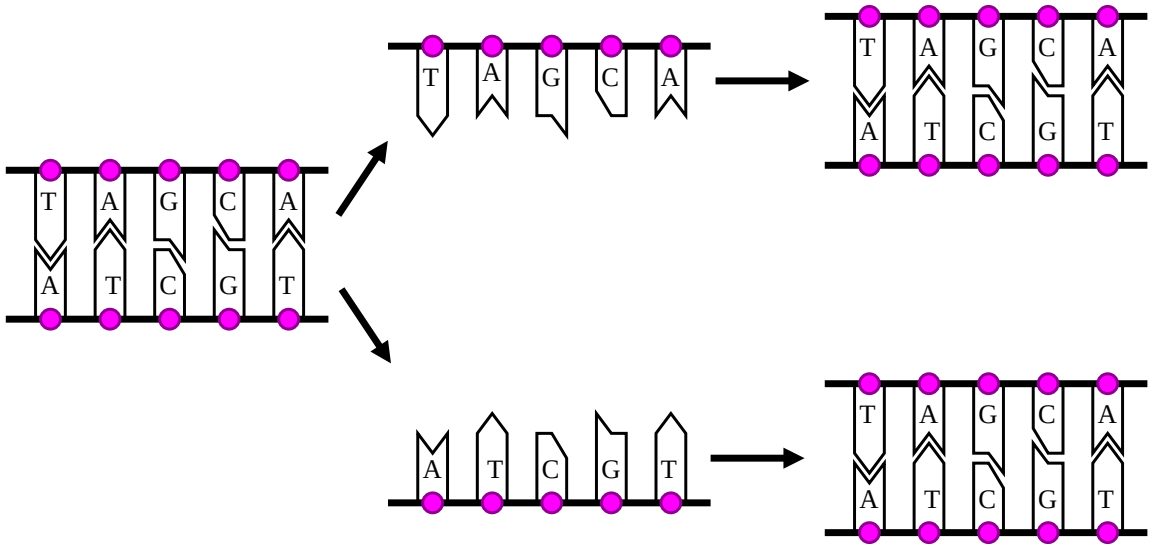
DNA Replication

DNA has the double helix structure. In eukaryotic cell, DNA is packed in the nuclear as tangled long strings. As we know a cell divides itself into two cells. Each cell has the same genetic information: the same DNA. the cell make a copy from DNA called **DNA replication**.

Semi Conservative Replication

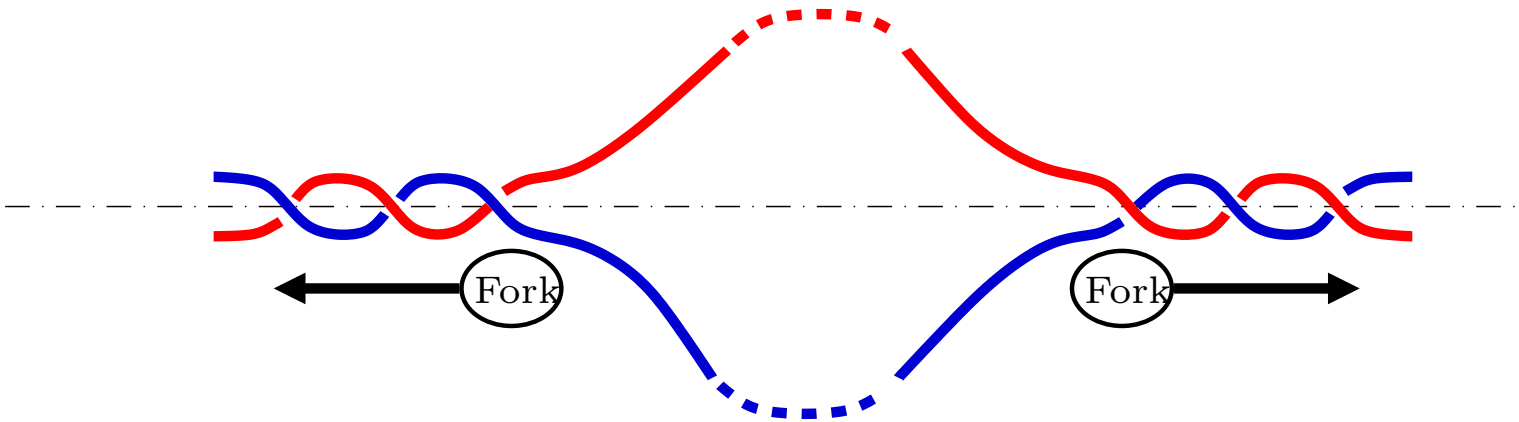
The double strands are connected by nucleotide bases: A= Adenine, T = Thymine, C = Cytosine, G = Guamine. T always joins A and G always joins C. Replication scheme is the following.

1. Locally, the double strands are separated into two chains.
2. Each chain has a sequence of bases A, T, C, G and this sequece casts its counterpart.



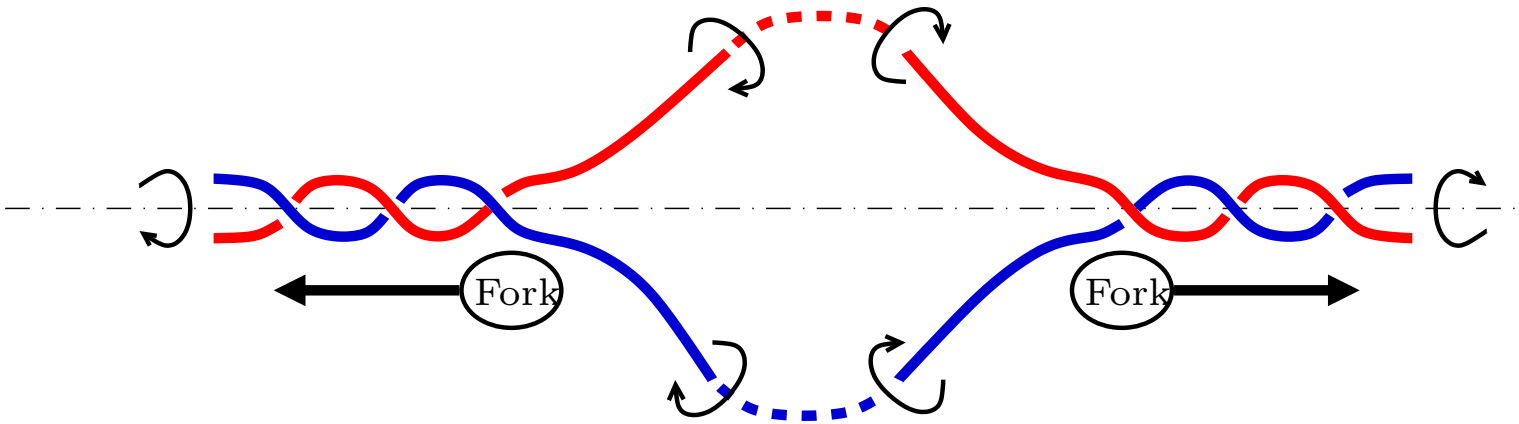
DNA Replication

The replication starts at a special site on the DNA starands called **Ori**. There are mainly two types of replications: Bidirectional and Unidirectional replications. Either case, at the *ori* DNA helix is relaxed and forms a site where the replication is done called **fork**.



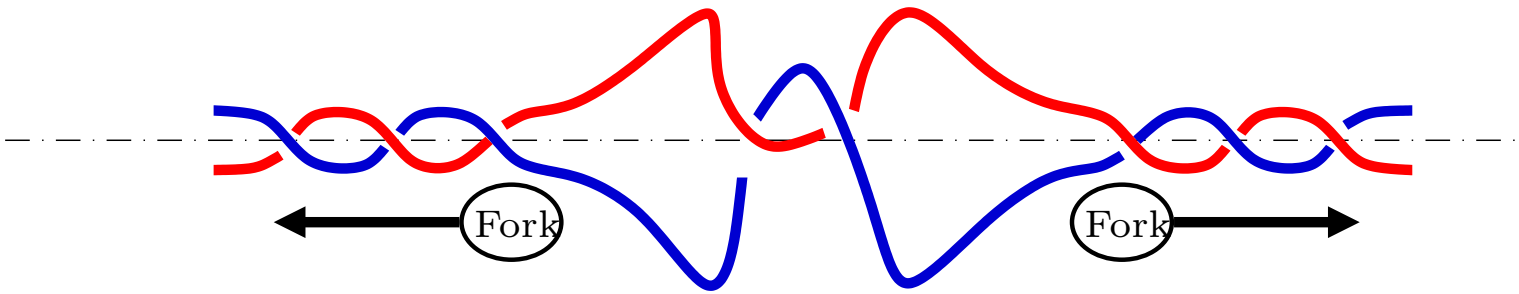
DNA Replication

We propose the follwoing mechanism for the fork moving. As the forks move away from the ori, the starand rotates as below.



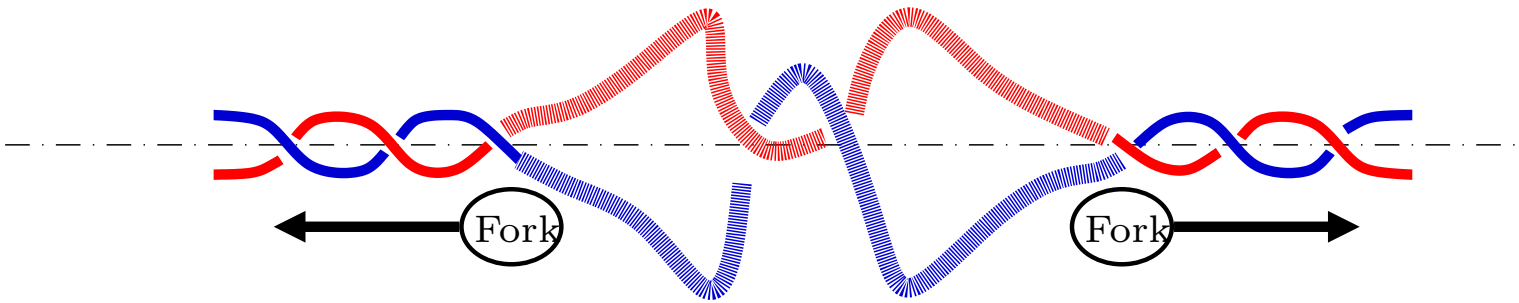
Topological Problems

When a replication is done in the nuclear of a cell, there is a topological problem. If the site **replication eye** (replication bubble) is twisted, then the resulting copies are linked (catenated): It is said that a catenated DNA is decatenated by action of enzymes (topoisomerase (Type II)).



Topological Problems

When a replication is done in the nuclear of a cell, there is a topological problem. If the site **replication eye** (replication bubble) is twisted, then the resulting copies are linked (catenated): It is said that a catenated DNA is decatenated by action of enzymes (topoisomerase (Type II)).



Topological Problems

The linking number between daughter DNAs is almost the same number of Tw.
The topological problems are:

- How many times topoisomerase (Type II) cut the daughter DNA strings to decatenate?
- After cut the DNA, how they match the cutting edges to recover the DNA?

Conjecture

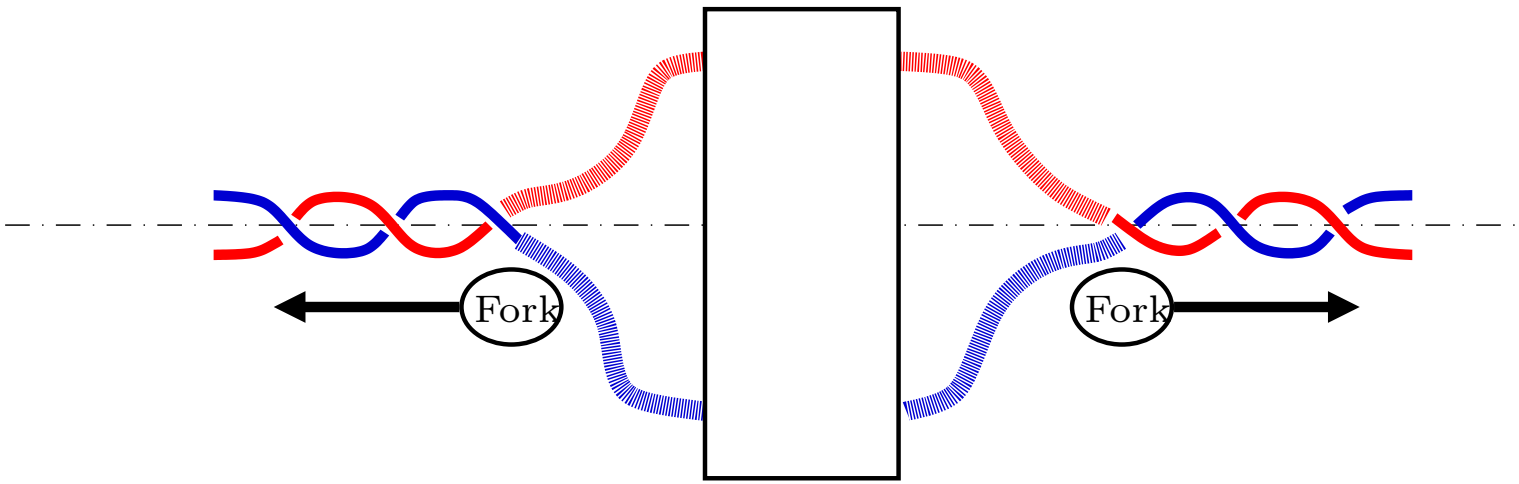
We conjectured the following:

$$\mathrm{Lk} = \mathrm{Tw} + \mathrm{Wr} \approx 0.$$

If the link is splittable, then the linking number must be zero. Of course the converse is not true in general.

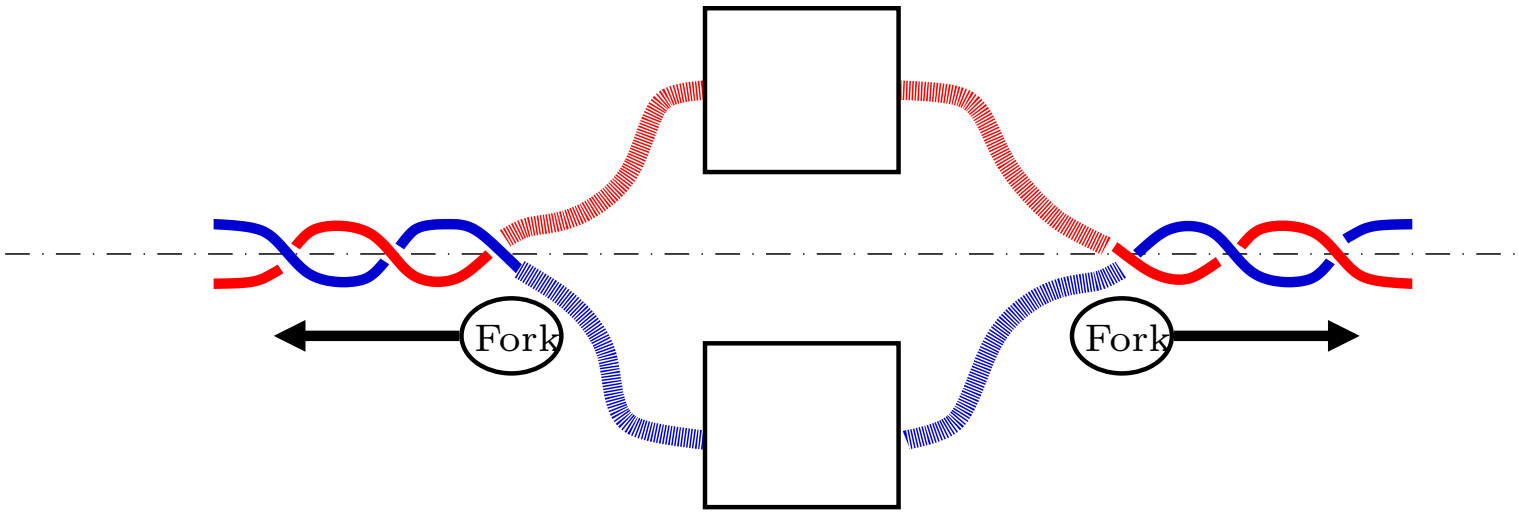
Possible Topological Model

We propose the model that has a mechanism to avoid creating the catenated DNA in the site of replication. The mechanism is placed in the box.



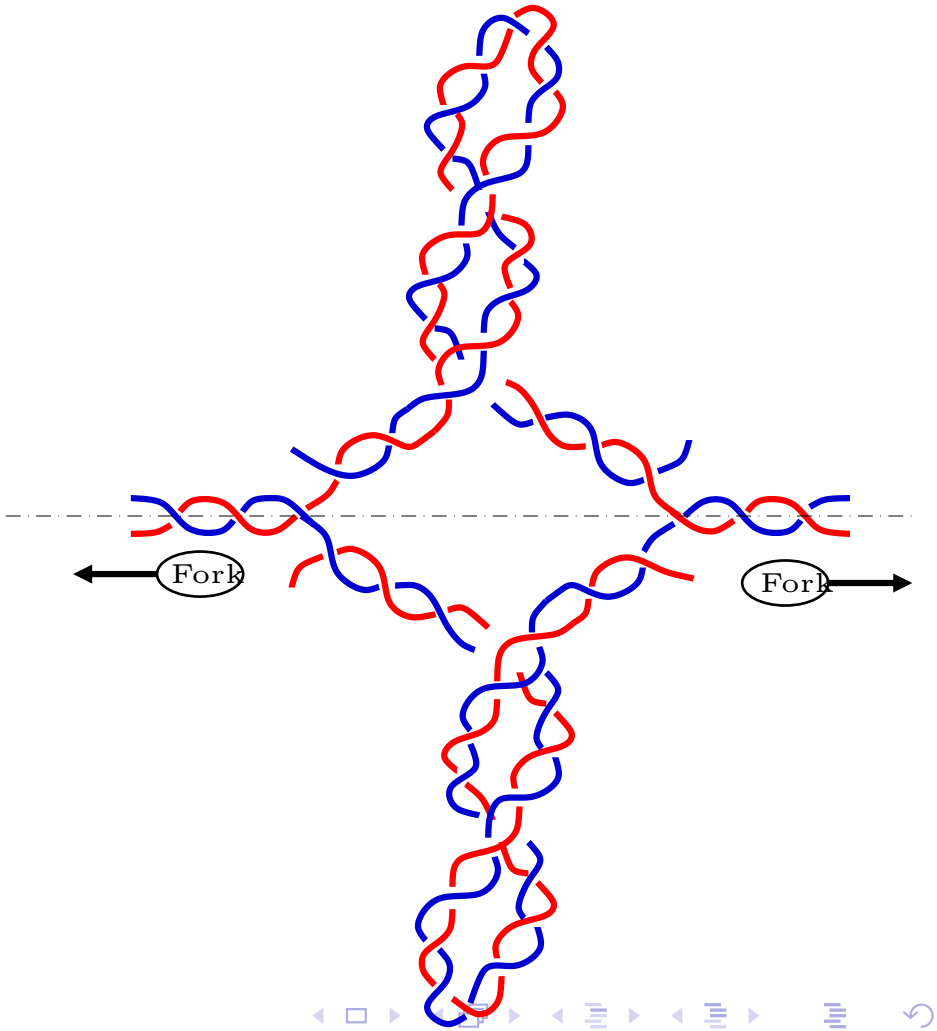
Possible Topological Model

In order to separate the doughter DNA, there may be two mechamisms on each daughter DNA. Each box may have a widing mechanism to form a chromosome.



Possible Topological Model

This is a possible topological model. As replication goes, supercoiling of replicated DNA will be created.



Estimate of the Writhe

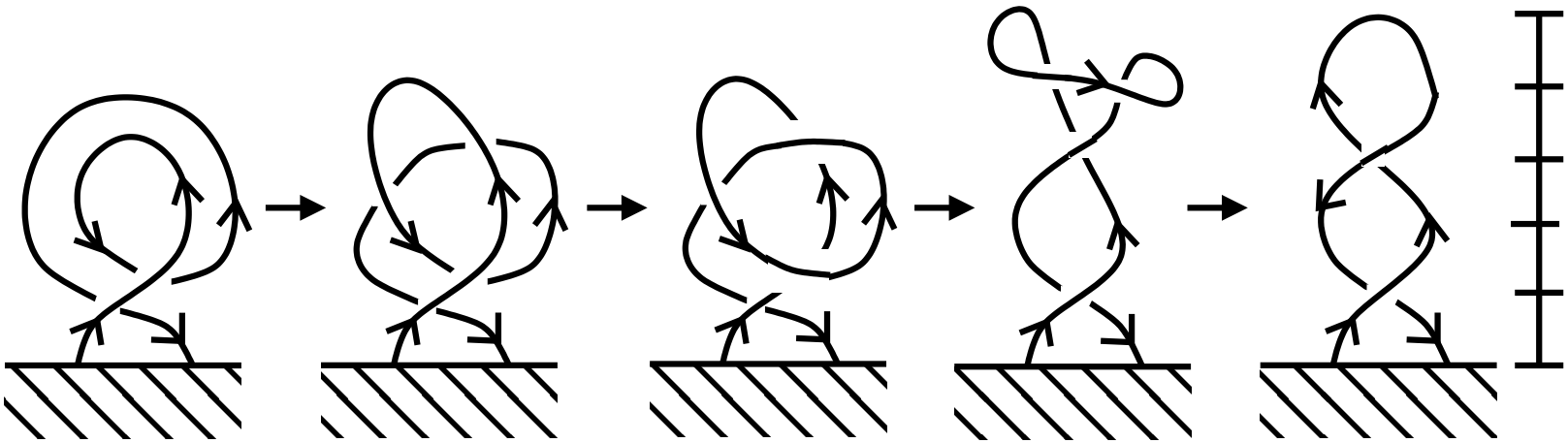
We now estimate the writhe of the DNA string around the histone. We may assume that a histone appears every 197 base pairs.

$$\frac{197}{10.5} \approx 18.76 \quad \text{(full twists)}.$$

$$\frac{Wr}{Tw} \approx \frac{2}{18.76} \approx 0.107$$

This means the number of supercoilings formed by histones is about 10.7%.

Estimate of the Writhe



As we can see the figure, the full twist occupies about 80% of the total loop. Therefore $197 \times 0.8 \approx 158$ bp. We may use this size as a unit to estimate $|\mathrm{Wr}|$.

Estimate of the Writhe

For this unit $|\mathrm{Wr}| = 2$ needs 158 bp. The size of the DNA is 3.1×10^9 bp. The number Tw for the human DNA is estimated as

$$\frac{3.1 \times 10^9}{10.5} \approx 2.95 \times 10^8$$

Estimate of the Writhe

On the other hand, the number of writhe is the sum of the number of writhe from histones and supercoilings. As we have seen, the writhe from histones is about 10% of the total twisting Tw. Therefore, the writhe from supercoilings is:

$$\frac{3.1 \times 0.9 \times 10^9}{158} \times 2 \approx 3.53 \times 10^7$$

Estimate of the Writhe

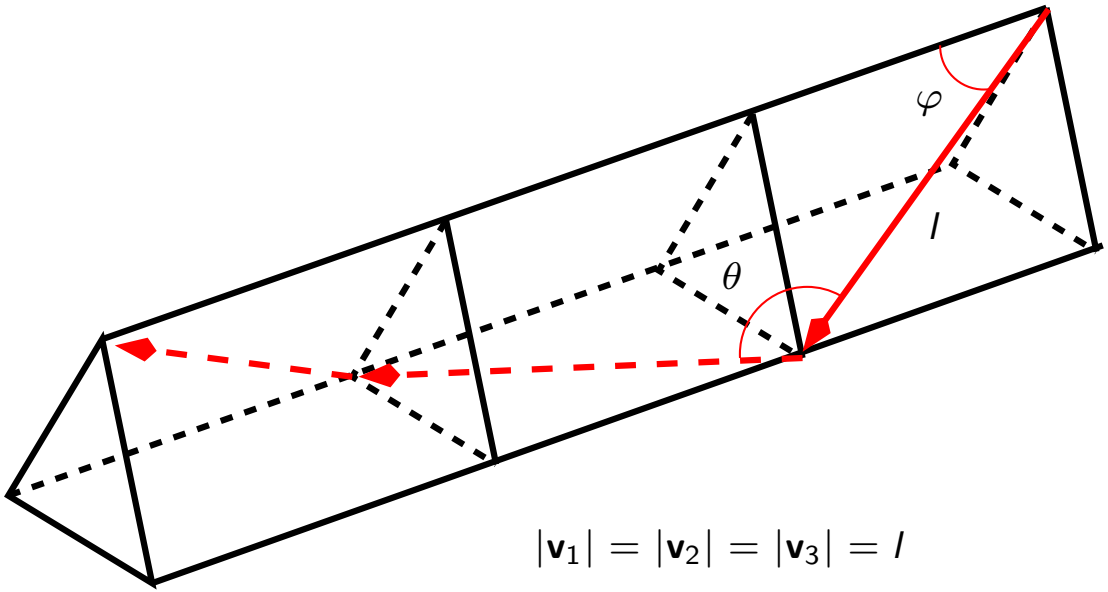
The ratio is:

$$\frac{3.53 \times 10^7}{2.95 \times 10^8} \approx 0.12.$$

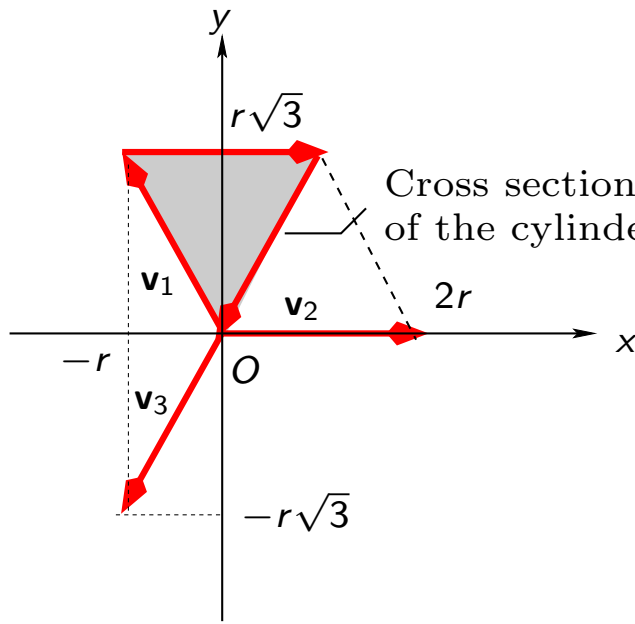
This shows that about $10 + 12 = 22\%$ of the writhe is produced by the histones and supercoilings. Therefore, 78 % of the twisting Tw must be cancelled to obtain $\mathrm{Lk} \approx 0$.

Hidden Writhe

We assume that there is an imaginary triangular cylindrical thin tube inside the double helix of DNA.



$$|\mathbf{v}_1| = |\mathbf{v}_2| = |\mathbf{v}_3| = l$$



The core curve consists of **core edges**. The core edges form angles φ and θ.

Hidden Writhe

Let l (bp) be the length of one edge along one face of the cylinder. Suppose that the length of the edge of the equilateral triangle is $2r$. In order to form one writhe, we need three vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$:

$$\begin{aligned} \mathbf{v}_1 &= (-r, r\sqrt{3}, \sqrt{l^2 - 4r^2}) \\ \mathbf{v}_2 &= (2r, 0, \sqrt{l^2 - 4r^2}) \\ \mathbf{v}_3 &= (-r, -r\sqrt{3}, \sqrt{l^2 - 4r^2}) \\ \mathbf{n} &= (0, 0, 1) \end{aligned}$$

Hidden Writhe

The angles θ are interpreted as angles between two consecutive vectors and they are equal:

$$\cos\theta = \frac{-\mathbf{v}_1 \cdot \mathbf{v}_2}{|\mathbf{v}_1||\mathbf{v}_2|} = \frac{-\mathbf{v}_1 \cdot \mathbf{v}_2}{l^2}$$

Also the angle φ between one of three vectors and the normal vector to the xy-plane is given by

$$\cos\varphi = \frac{\mathbf{v}_1 \cdot \mathbf{n}}{|\mathbf{v}_1||\mathbf{n}|} = \frac{\mathbf{v}_1 \cdot \mathbf{n}}{l}$$

Hidden Writhe

The total number of twists Tw is:

$$\mathrm{Tw} \approx 2.95 \times 10^8$$

From the observation, 78 % of twistings must be cancelled. Thus

$$2.95 \times 0.78 \times 10^8 \approx 2.30 \times 10^8$$

To find the length of DNA in bp to make one writhe, divide the total number of base pairs by this:

$$\frac{3.1 \times 10^9}{2.30 \times 10^8} \approx 13.5 \text{ bp}$$

Hidden Writhe

Therefore, for the triangular cylinder model,

$$l = \frac{13.5}{3} \approx 4.5 \text{ bp}$$

Examples

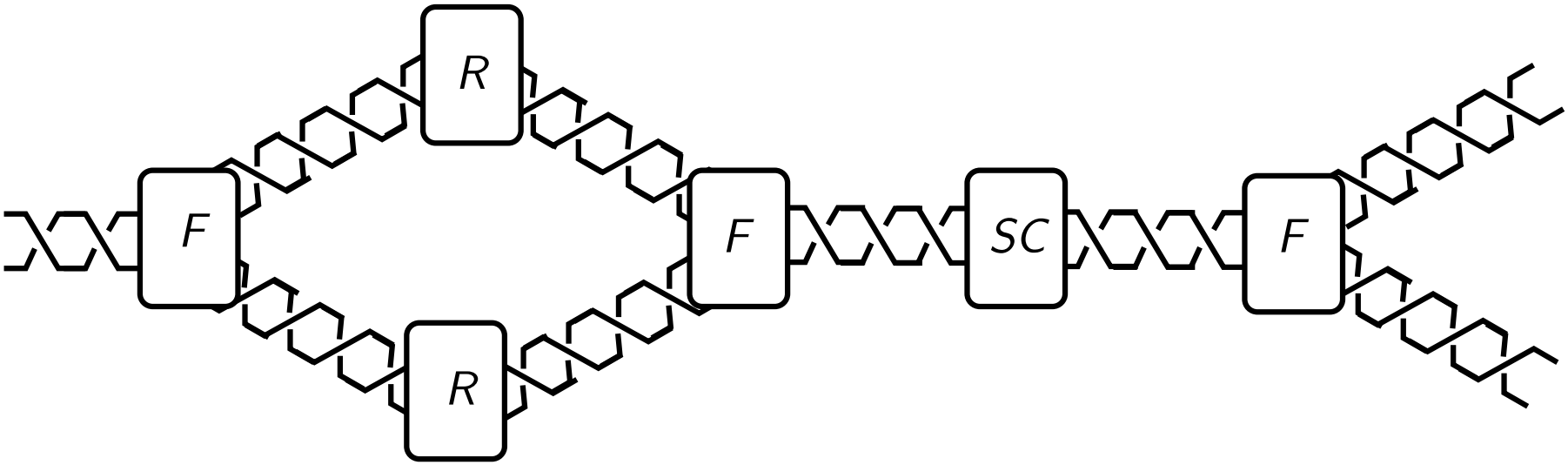
We caluculate the angles θ and φ for some r and l.

r (bp)	l (bp)	$\cos\theta$	θ	$\cos\varphi$	φ
0.25	4.5	-0.9815	169.00°	0.9997	1.41°
0.5	4.5	-0. 9259	157.81°	0. 9988	2.83°
1	4.5	-0. 7037	134.73°	0. 9951	5.67°

When $r = 0.25$ bp and $l = 4.5$ bp, $\theta \approx 169°$ and $\varphi \approx 1.4°$. If l is longer, then the angles will be less. This shows that to create writhe, we do not need to bend DNA widely.

Rotations and Supercoilings

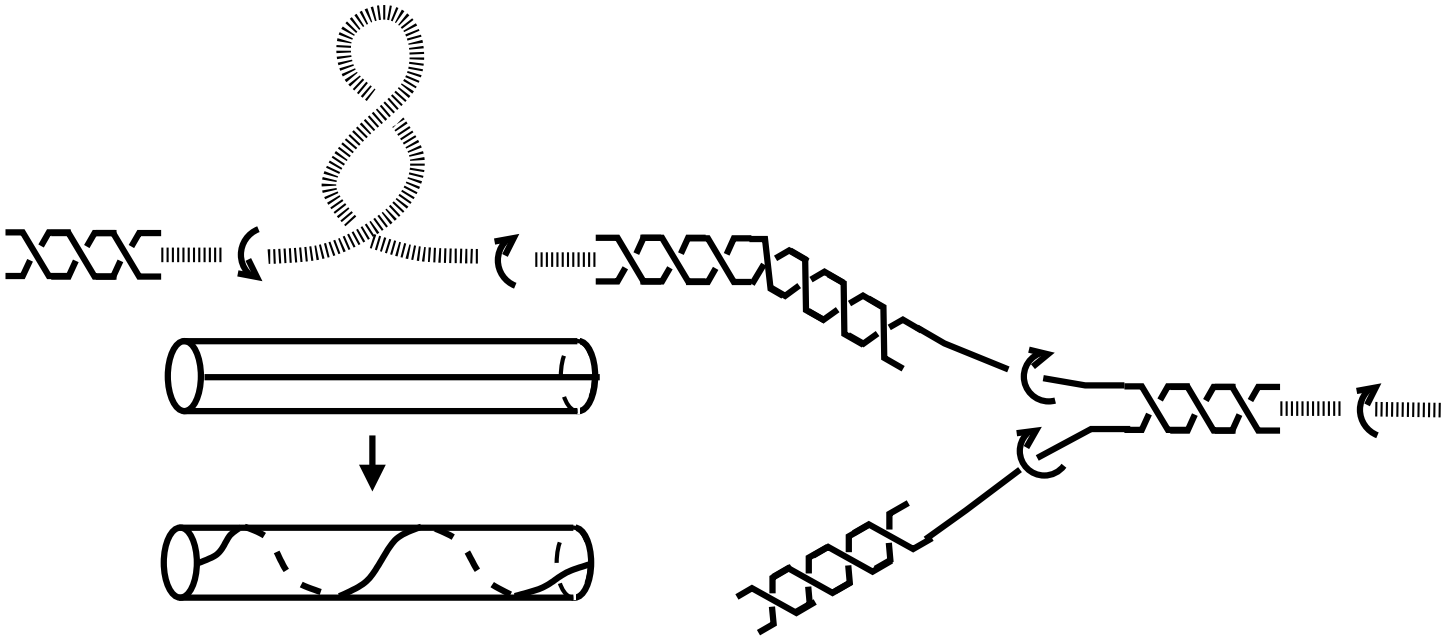
In the following, we restrict our arguments on replication bubbles with two forks. Our topological model of a replication eye consists of two forks and two parts in which negative supercoilings are produced. We call this winding mechanism a **reel**.



Rotations and Supercoilings

At the fork, the DNA strands are separated and this creates the rotation of each single strand around its core curve as well as it creates the rotation of the original DNA around its core curve. The strand between two replication sites, the rotation of the original DNA creates the positive supercoiling. Between two consecutive eyes a negative supercoiling will be produced.
In order to start the replication, the specified site of DNA is relaxed by the enzyme: **topoisomerase (Type II)**. Then two forks are produced in the replication eye.

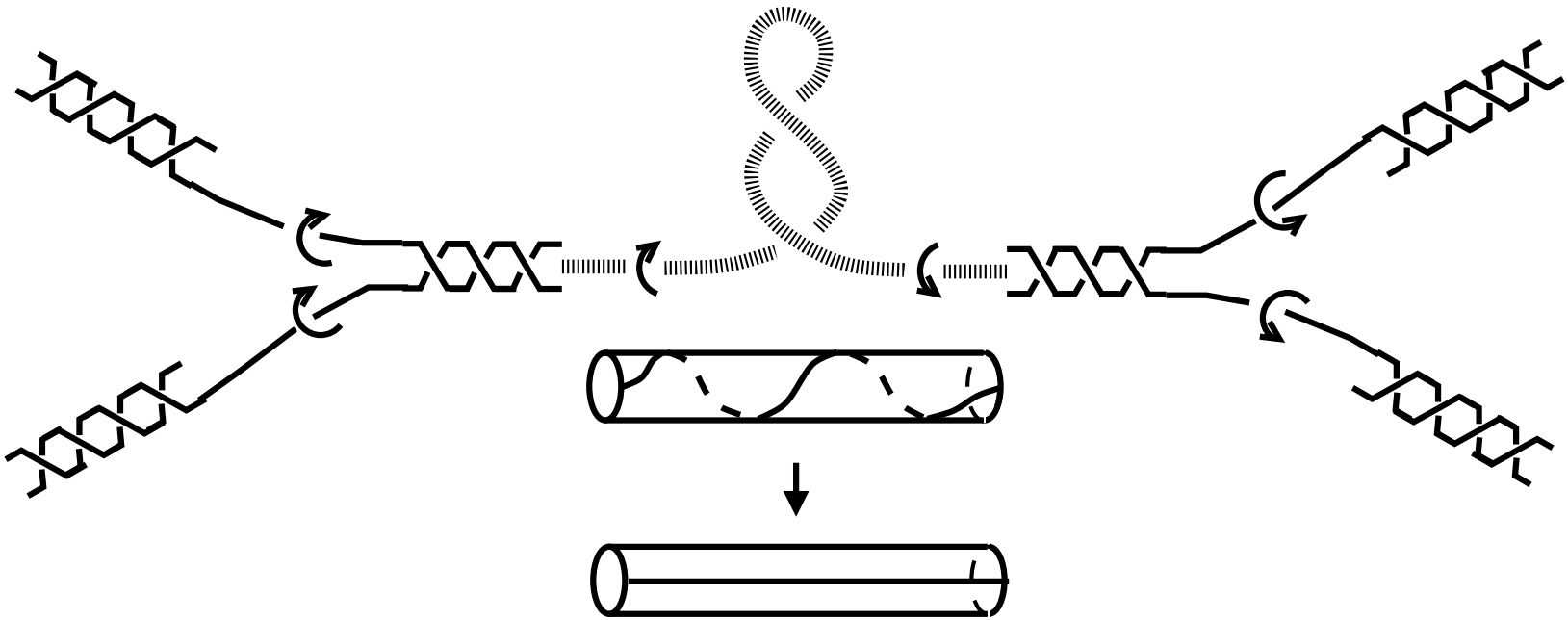
Rotations and Supercoilings



Rotations and Supercoilings

As the replication proceeds, the fork will move away from the site. In order to move the fork, the double helix strands must be relaxed. To do this the single strands and the original DNA must be rotated. Inside eye, the rotation will create the hidden writhe in the daughter DNAs and also it will introduce negative supercoilings of daughter DNAs. This is done in the reel part of our model. Between two consecutive eyes the rotation will cancel the hidden writhe and also it will introduce positive supercoiling.

Rotations and Supercoilings



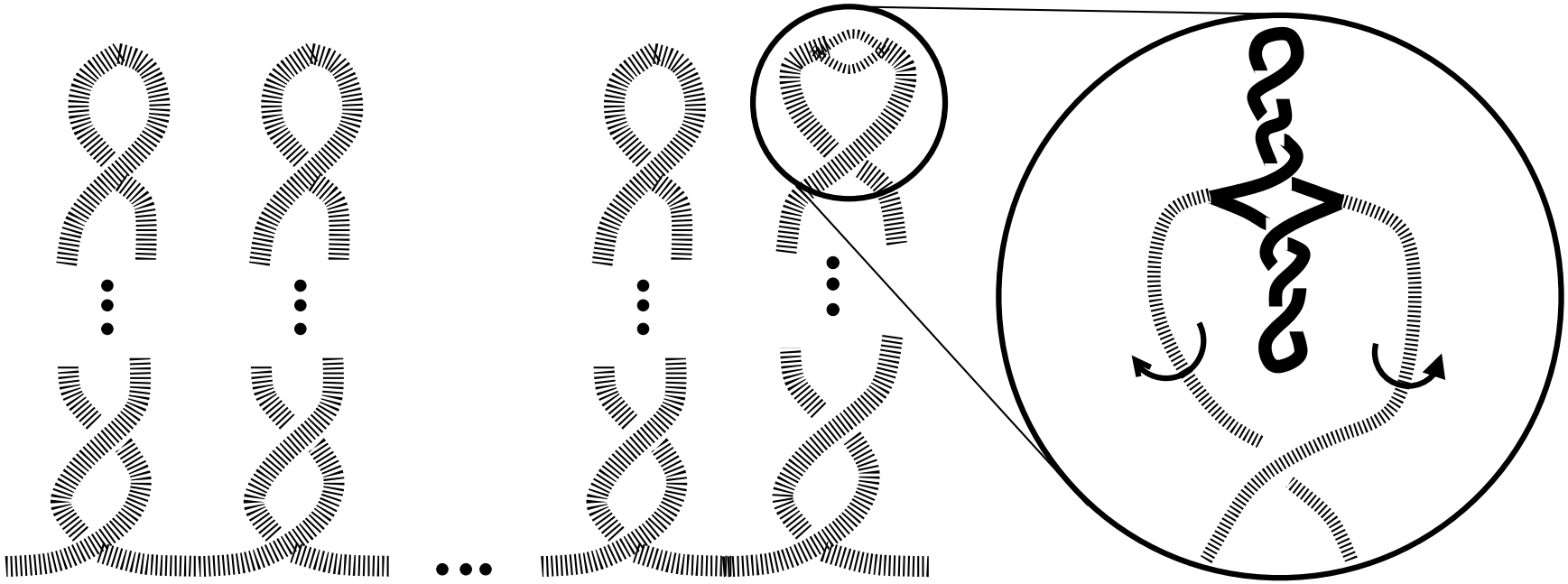
Topological Model

Our model must satisfy the following conditions.

1. The process in the model is compatible with the process in the replication process at the forks.
2. The copied DNA preserves the same shape and the configuration of the original DNA.

In order to satisfy these conditions, we assume that there exists a topological domain of DNA depicted in the figure:

Topological Model



Note that in the figure we omit to draw histones.

Topological Model

We suppose that the topological domain has the following properties:

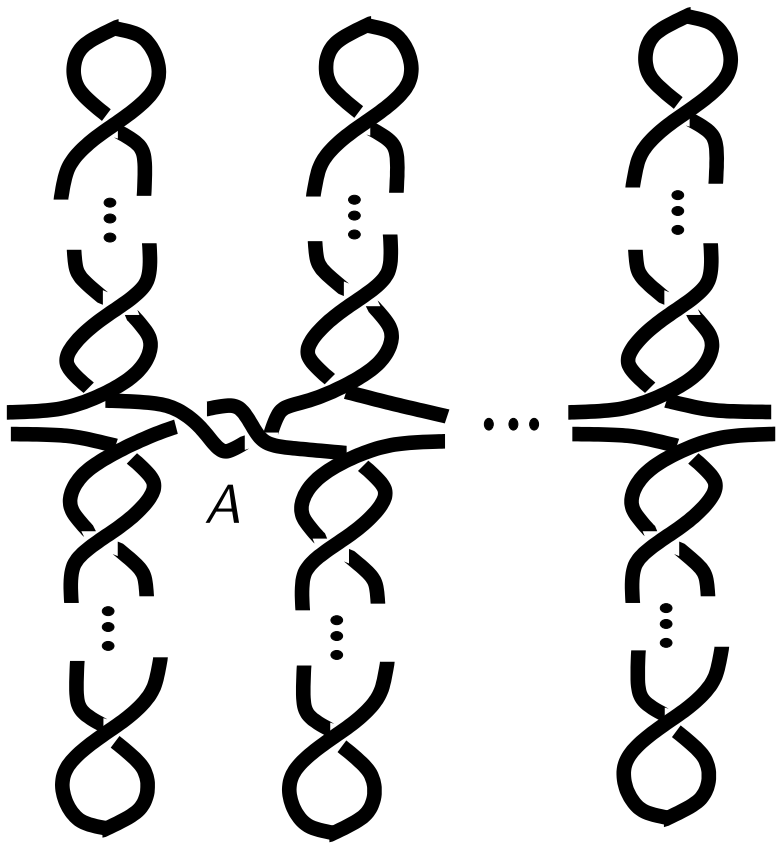
1. it consists of negative supercoilings and histones. We call each coiling a **tower**, and
2. each ori lies on the top of the tower so that the number of towers is the number of oris in the domain.

Topological Model

At the end of the process we obtain the duplicated DNA attached at the meeting points of forks denoted by A in the figure.

Topological Model

At A the replicated DNA may be linked and they need the topoisomerase (Type II) to split them.



Conclusion

The proposed topological model consists of negative supercoiling towers and oris at the top of each tower. When the replication starts from these oris, the daughter DNAs are produced and rotated to create the hidden writhe and then negative supercoilings and histones will be created. Between eyes, the rotation cancel the hidden writhe and creates positive supercoilings but they are cancelled by the negative supercoilings of the topological domain. Therefore, the model satisfies the equation.

$$\mathrm{Lk} = \mathrm{Tw} + \mathrm{Wr} \approx 0$$

Conclusion

Also this model creates the identical pair of topological domains linked at the base of towers. The number of the linking part is therefore, less than the number of towers. This means the number of times of cut and paste operations by topoisomerase (Type II) in the replication process can be reduced.

Thank you!