# Application of R-graphs to DNA modelling

**A. Barbensi**[1], D. Celoria[1]  Supervisors: Prof. D. Buck[2], Dr. H. A. Harrington[1], Prof. M. Lackenby[1]

[1]University of Oxford, [2]University of Bath

UNIVERSITY OF OXFORD
Mathematical Institute

## Abstract

Stasiak and collaborators in [2] investigate the connectivity between **Knots Spaces** accessible to a **circular knotted double stranded DNA molecule**, under the action of **Type II Topoisomerases.** They describe the Spaces of Knots as a foam where neighbouring sub-spaces contain configurations of knot types differing by a single intersegmental passage.

Here we define a new mathematical tool to model the action of **Type I and II Topoisomerases** on a covalently closed circular DNA molecule: the **Reidemeister graph**. We then analyse the local properties of this object from a mathematical point of view.

Finally, we indicate how the Reidemeister graph can be used to infer information about the proteins' action.

## Problem

Our goal is to describe the set of possible configuration of **knots** (*i.e.* **covalently closed circular DNA molecules**) in a computable way, using a newly developed mathematical tool.

Finding implementable descriptions of possible configurations of a **knotted DNA molecule** has remarkable importance from a biological point of view, as highlighted by Stasiak and collaborators in [2], and in other works (see *e.g.* [9], [10]) and it is a hard and well studied problem in mathematics (see *e.g.* [1]).

## Approach

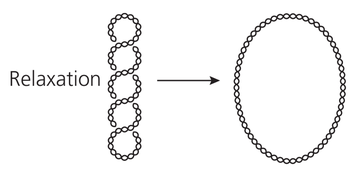
First approximation: encode the data regarding the configuration space of a single knot type K in a **graph**, the Reidemeister graph G(K).

This is done by working on **diagrams**, *i.e.* **drawings of knots**. We then study properties of the graph to infer information on the action of **Type I Topoisomerases**.

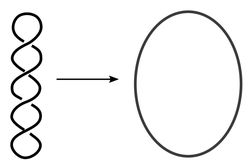
Finally, we consider a "bigger" graph, in which the vertices are diagrams of any knot type, and in which the action of **Type II Topoisomerases** can be analysed as well.

## Our model of the Space of Knots

We represent the central axes of a **circular** (possibly knotted) **molecule of DNA** as a diagram. In this setting, the action of **Type I Topoisomerases** can be modeled as shown in the Figure below.

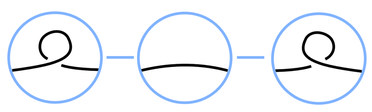



This Picture is taken from [8],
kind courtesy of A. D. Bates and A. Maxwell.

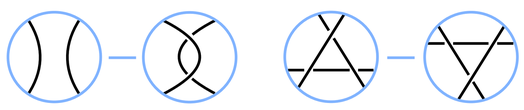


In our work [4] we consider only the central axes of the DNA molecule, disregarding the double helix structure.

We then represent deformations and the action of Type I Topoisomerases (assuming no chirality bias*) by local moves.



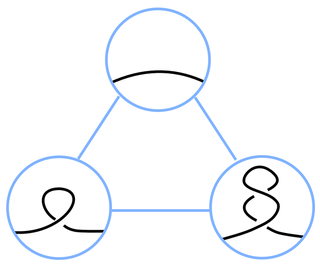
This local move represents the action of Type I Topoisomerases.
The two possibilities differ in the sign of the crossing, and lead to positive or negative coiling.
We will refer to this move as a Reidemeister I move.



These two other moves (called Reidemeister II and III, respectively) also represent deformations.
Together with the previous one they form a complete set of local moves relating
diagrams representing equivalent knots.

Our choice is consistent, since a famous Theorem [3] in topology (by K. Reidemeister, from whom the moves take the name) ensures that two diagrams represent the same knot type if and only if they are related by a finite sequence of these moves.

In our work [4], for a given knot K, we consider the (infinite) graph whose vertices are given by diagrams representing K, and edges connect two diagrams related by any of the moves described above. We call the resulting locally finite graph obtained **the Reidemeister graph** of **K**.



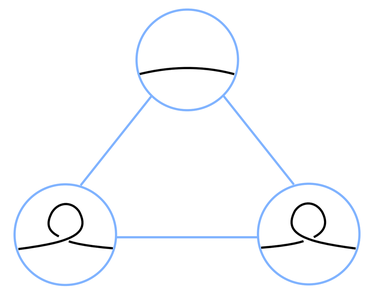
The action of Type I Topoisomerases plays a special role: the only
length three cycles are composed of two succesive Type I Topoisomerases-mediated
actions of opposite sign, followed by a Reidemeister II move.

We can consider **short circular paths (*cycles*)** on the graph. The only cycles of length 3 are of the form described in the picture. Thus, our Reidemeister Graph highlights the particular behaviour of Type I Topoisomerases.

*We discuss the possibility of using weighted edges in order to include chirality bias in the box "Applications and further work".

## A step further: include the action of Type II Topoisomerases

If we wish to include the action of **Type II Topoisomerases**, we need to consider all possible different knots at once, since the Topo II might change the topology of the (axes of the) DNA molecule.
We model the protein action as a crossing change as is standard (*e.g.* as done by Stasiak and collaborators in [2]). We then construct a larger graph, whose vertices are diagrams representing any knot type, and whose edges represent the Reidemeister moves plus the action of Type II Topoisomerases (assuming no chirality bias). We call this object **the blown-up Reidemeister graph G***.



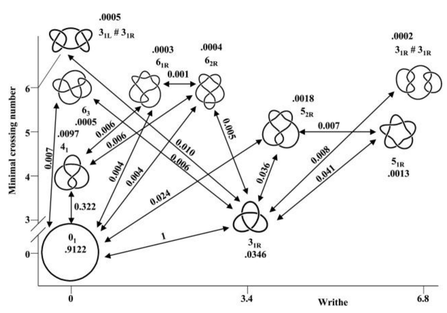
The action of Type II Topoisomerases fits into a length three cycle.
We proved in [4] that this configuration, together with the previous one are
the only admissible length three cycles.

We can then carry out a similar analysis as in the previous case, and look for **short cycles** in the graph. The action of Type II Topoisomerases is detected among the other local moves, and together with the Type I Topoisomerases, are the only moves that fit into length three cycles.

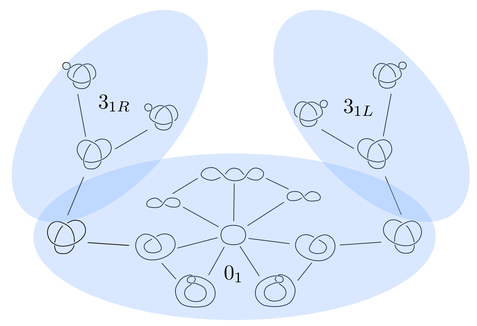
## Boundaries between Spaces of Knots

In their work "*Simulations of Action of DNA Topoisomerases to Investigate Boundaries and Shapes of Spaces of Knots*" Stasiak and collaborators obtain a schematic picture (shown in the left side of the Figure below) for the **connectivity** between spaces of knots. Labels in the edges represent the **probability** of passing form one knot type to another by random intersegmental passages, simulating the action of a Type II Topoisomerases without chirality bias.

With our construction we get a representation of the configuration spaces of knots "all at once", as an infinite graph. We can focus on small portions of this graph and look for information about exchanges between particular knot spaces.




This Picture is taken from [2],
kind courtesy of A. Stasiak.




In the center, the radius 1 ball centered
in the unkot diagram. On the sides, we
show paths leading to the Trefoil knot.

While Stasiak and collaborator's results are based on **simulations**, we provide a **mathematical framework** which allows us to **a priori obstruct** or **exclude** particular sequences of reactions, or the formation of specific knot types.
Moreover, with their simulation they obtain the probability of passing from one knot to the other, without considering the distinct configurations involved. Thus, by considering **specific configurations** (diagrams) we achieve a more granular understanding and a complementary approach to their model.

Passage from a trefoil
to an unknot.

This could possibly lead to better understanding the mechanics of the action of Type I and II Topoisomerases, since with our model we are able to to investigate not only **when** the change in the knot type happens, but also **where** (*i.e* between which configurations).

## Applications and further work

**Summarising**, for a circular **double stranded DNA molecule** we model the action of Type I and II Topoisomerases as a path in the blown-up graph, and use  mathematical tools to obstruct configurations or specific reactions.

Even if the graph is infinite, every vertex has finite valence. Moreover, if we consider  **finite-length knots** the resulting graphs are finite. These facts ensure that our approach is **implementable**.
The most feasible way to approach this point of view is through **grid diagrams** [5], [6]. In this case **a knot is encoded in a pair of permutations**. This fact obviously facilitates a **computational** approach.

We could refine our approach by adding **weights** to the edges of our graph and thus include a **chirality bias**.

Another approach could be to focus **only** on **Type II Topoisomerases** and the Reidemeister II and III moves. From a mathematical point of view, this means using diagrams representing **framed knots**.

In [7] the authors model **site-specific recombinations** as band surgeries. In terms of local moves between diagrams this translates to nullifications. We can define another graph, adding diagrams representing **catenanes** and edges representing nullification moves, to investigate short paths and configurations, or to obstruct their existence.

This local move represents the action of
site-specific recombination.

## References

[1] Bar-Natan, D. "On the Vassiliev knot invariants." Topology 34.2 (1995): 423-472.APA

[2] Flammini, A. and Maritan A. and Stasiak, A. "Simulations of action of DNA topoisomerases to investigate boundaries and shapes of spaces of knots." Biophysical Journal 87.5 (2004): 2968-2975.

[3] Reidemeister, K. "Elementare Begrundung der Knotentheorie." Abh. Math. Sem. Univ. Hamburg , 5 (1927) pp. 24–32

[4] Barbensi, A. and Celoria, D. "The Reidemeister graph is a complete knot invariant." In preparation, preprint available upon request.

[5] Cromwell, P. R. "Embedding knots and links in an open book I: Basic properties." Topology and its Applications 64.1 (1995): 37-58.

[6] Witte, S. and Brasher, R. and Vazquez, M. "Randomly Sampling Grid Diagrams of Knots."

[7] Buck, D. and Ishihara, K. "Coherent band pathways between knots and links." Journal of Knot Theory and Its Ramifications 24.02 (2015)

[8] Bates, A. D., and Maxwell, A. "DNA topology."  Oxford University Press, USA, 2005.

[9] Hua, X. and Nguyen, D. and Raghavan, B.and  Arsuaga, J. and Vazquez, M. "Random state transitions of knots: a first step towards modeling unknotting by type II topoisomerases." Topology and its Applications 154.7 (2007): 1381-1397.

[10] Burnier, Y. and Weber, C. and Flammini, A. and Stasiak, A. (2007). Local selection rules that can determine specific pathways of DNA unknotting by type II DNA topoisomerases. Nucleic acids research, 35(15), 5223-5231.

## Acknowledgements

AB would like to thank her supervisors for their support and useful suggestions, and Nicola De Mitri for graphic adjustments.

The authors would like to thank Marco Marengon for his remarks and valuable help.

AB is supported by the EPSRC grant “Algebraic and topological approaches for genomic data in molecular biology” EP/R005125/1.

DC has received support from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 674978).

DB gratefully acknowledges support from the Leverhulme Trust (Grant RP2013-K-017).