Application of $\mathcal{R}$-graphs to DNA modelling

A. Barbensi
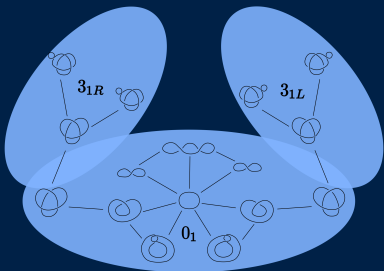*DPhil Student*
Supervisors: D.Buck, H.A.Harrington, M.Lackenby
Mathematical Institute
University of Oxford

CMO-BIRS Workshop, November 2017

Oxford
Mathematics

Goal

Finding implementable descriptions of the configurations of a knotted DNA molecule.

This problem is of remarkable importance from a biological point of view, and it is a hard and well studied problem in mathematics.

Here we present a newly developed mathematical tool that describes the configuration space of knots: the **Reidemeister Graph**.
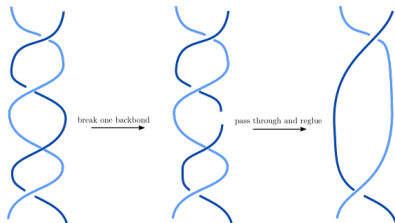
Modelling DNA

We represent the central axis of a circular (possibly knotted) molecule of DNA as a diagram, thought as the core of a ribbon whose edges are the two backbones of the DNA.

Picture taken from **DNA Topology**
kind courtesy of A. D. Bates and A. Maxwell
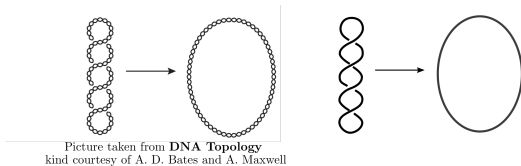
In our work we consider only the central axis of the DNA molecule, disregarding the double helix structure.

Modelling Type I Topoisomerases

Type I Topoisomerases are proteins that regulate supercoiling.

They bind to the DNA molecule, make a break in one of the DNA backbones and then pass the other backbone through before reattaching it.

Modelling Type I Topoisomerases
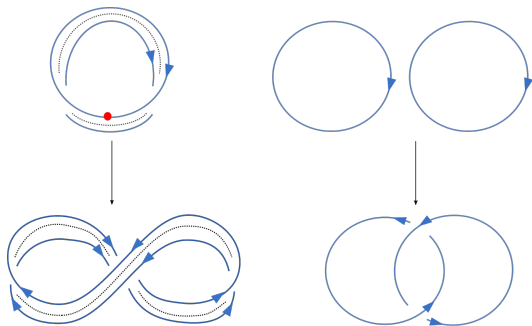
Each edge of the ribbon represents one strand of DNA. Cutting one strand at the red dot, introducing a 2π turn and reglueing results in a change by one the linking number of the edges of the ribbon.
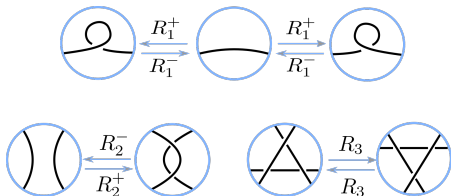
The effect of Topo I on the core of the ribbon
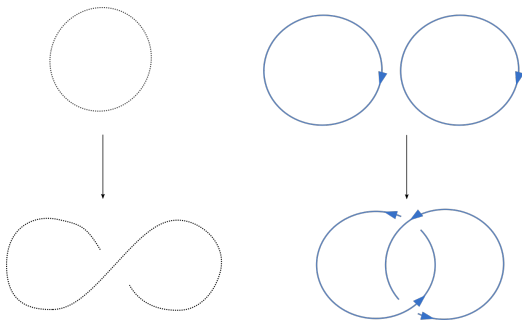
The effect on the central axis of the ribbon is the same as adding a supercoil (*a.k.a.* a **Reidemeister I** move).

Representing the Topo I action by local moves

We can represent local deformations (*i.e.* isotopies) and the action of Type I Topoisomerases by local moves.


Theorem [Reidemeister, 1927]

Two diagrams represent the same knot type if and only if they are related by a finite sequence of R_1, R_2 and R_3 moves.

The Reidemeister Graph

Definition

Given a knot type K, the **Reidemeister graph** $\mathcal{G}(K)$ is the graph whose vertices are given by diagrams representing K. Edges connect two vertices whenever the corresponding diagrams are related by a single Reidemeister move.
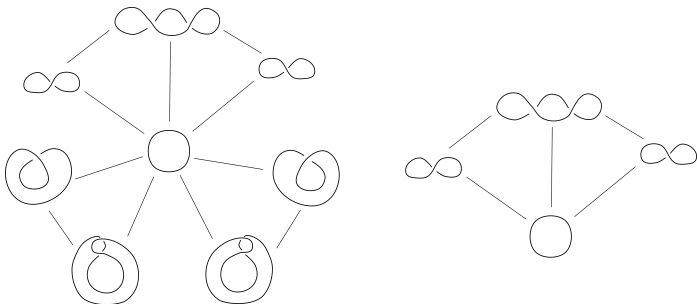
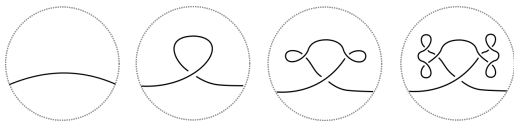**Some global properties of the Graph**

Observations

- $\mathcal{G}(K)$ is connected.
- $\mathcal{G}(K)$ is **infinite** (it presents also a fractal behaviour);
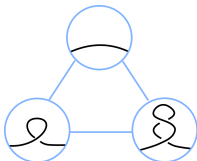

- $\mathcal{G}(K)$ is **locally finite** ($v(D) \leq 16cr(D) + n(D)$).

Theorem 1

We can endow the graph with the *path metric*, and consider short circular paths (**cycles**) on the graph.

Theorem [B.-Celoria]

The only cycles of length 3 are of the form described in the picture below.

The only length three cycles are composed of two successive Type I Topoisomerases-mediated actions of opposite sign, followed by a Reidemeister II move.

Sketch of the proof

- Using crossing number, Hass-Novik invariants [1], and Arnold's *St* invariant [2] we exclude all the possible combinations except $R_2^{\pm}$-$R_1^{\mp}$-$R_1^{\mp}$;

- we carefully consider all the possibilities involving $R_2^{\pm}$-$R_1^{\mp}$-$R_1^{\mp}$, and reduce to the triangle below.
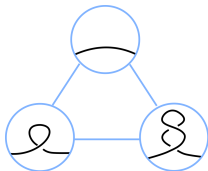
[1] Hass, Joel and Nowik, Tahl, *Invariants of knot diagrams*, Mathematische Annalen, vol 342, nr. 1, 2008.
[2] Arnold, Vladimir I., *Plane curves, their invariants, perestroikas and classifications*, Advances in Soviet Mathematics 21 (1994): 33-91.

Theorem 2

Theorem [B.-Celoria]

The Reidemeister graph (and more specifically some finite portions
of it) is a **complete** knot invariant up to mirroring:

$$\mathcal{G}(K) \cong \mathcal{G}(K') \Longleftrightarrow K = K' \text{ or } K = mK'$$
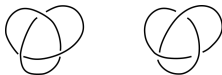
This is the first knot invariant whose completeness does not rely
on the completeness[*] of the knot complement $M_K = S^3 \setminus K$.

[*]Gordon, C. McA. and Luecke, J. *Knots are determined by their complements.* Journal of the American Mathematical Society 2.2 (1989): 371-415.

Sketch of the proof: a roadmap

Reconstruct the diagram corresponding to some specific vertex of the graph.

Sketch of the proof: a roadmap

Reconstruct the diagram corresponding to some specific vertex of the graph.

↑

Compute the projection (*i.e.* the diagrams disregarding the crossings, *a.k.a* a 4-valent graph).

Sketch of the proof: a roadmap

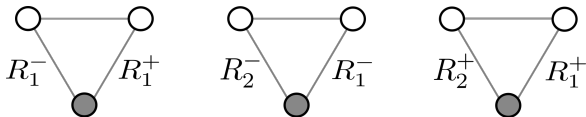Reconstruct the diagram corresponding to some specific vertex of the graph.

↑

Compute the projection (*i.e.* the diagrams disregarding the crossings, *a.k.a* a 4-valent graph).

↑

Compute the dual graph of the projection.

Sketch of the proof

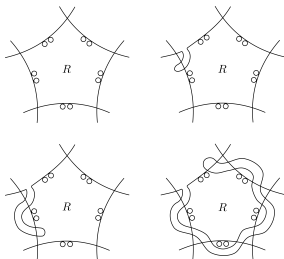- ▶ Using Theorem 1 we can label every edge and compute the number of crossings $cr(D)$ for every diagram D;

- ▶ we compute $P(D) = (P_1(D), \ldots, P_m(D))$ for every D with $P_1(D) = 0$, where $P_i(D) =$ number of regions with i edges in the complement of D;

Sketch of the proof

- we consider a big neighbourhood of a vertex v whose diagram satisfies $P_1(D) = 0$: its structure determines the projection of the diagram;

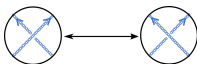- we prove that the neighbourhoods of two diagrams with the same projection differing in at least one crossing (but not in all) are distinct.

A step further: include the action of Type II Topoisomerases

Type II Topoisomerases' main function is to change the knot or
link type of the molecule.
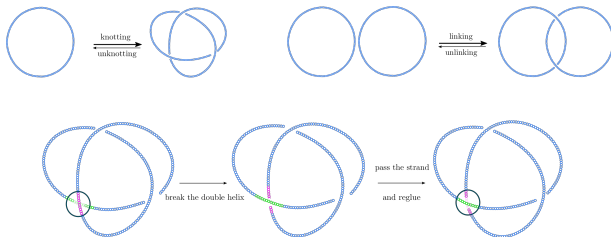
We model the action of Type II Topoisomerases as a crossing
change.

The blown-up Reidemeister graph.

Topo II might change the topology of the (axis of the) DNA molecule: we need to consider all knot types at once.

Definition
The blown-up Reidemeister graph $\mathcal{G}^*$ is the graph whose vertices are diagrams representing any knot type, and whose edges represent the Reidemeister moves plus the action of Type II Topoisomerases (*i.e.* crossing changes).

Theorem 3

- The action of Type II Topoisomerases is detected among the other local moves in $\mathcal{G}^*$;

- together with the Type I Topoisomerases, are the only moves that fit into triangles.

Theorem [B.-Celoria]

The only length three cycles in $\mathcal{G}^*$ are the previous one plus this new one.
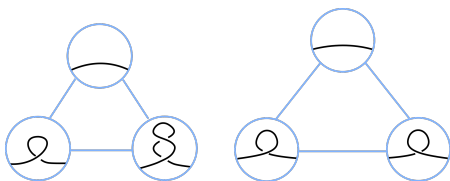
Previous models for the Space of Knots

Stasiak and collaborators [*] investigate the **connectivity** between spaces of knots:

- ▶ Start with a polymer configuration resembling the ideal configuration (*a.k.a* ropelength minimiser) (knots are considered up to 6 crossings);
- ▶ the configuration is evolved by applying random crank-shaft moves without intersegmental passages;
- ▶ the polymer is further evolved, allowing moves that result in one intersegmental passage;
- ▶ determine the resulting knot type.

This method can be applied to quantify biases manifested by Type II Topoisomerases.

*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.

A schematic Picture for the Spaces of Knots
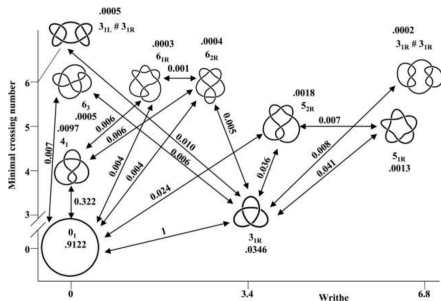
This Picture is taken from [*], courtesy of A. Stasiak.

Labels represent the probability of passing form one knot type to another.

*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.

Interfaces and shapes of individual knot spaces

A previous hypothesis was that relative probability of occurrence of various knots should decrease exponentially with the L/D of their corresponding ideal representations:

- ▶ their model contradicts this hypothesis: 5_2 occurence higher than 5_1;
- ▶ agrees on the hypothesis that the probability of remaining in the same knot type exhibits such decrease.

Our model for the Space of Knots

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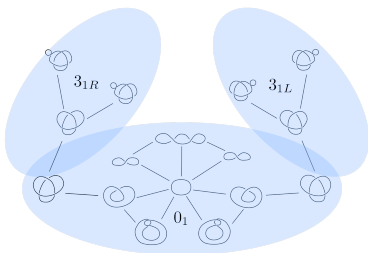
We can focus on small portions of this graph and look for information about exchanges between particular knot spaces.

A complementary approach

- Previous models are based on simulations: we provide a mathematical framework which allows us to *a priori* obstruct or exclude particular sequences of reactions;

- by considering specific configurations, we achieve a more granular understanding;

- by implementing our model we can compute exact values (*i.e.* not probabilities) for the connectivity between knots.

- we are however not able to infer information about relations between shapes of individual knot spaces and ideal configurations.

Summary

- For a circular double stranded DNA molecule we model the reactions due to Type I and II Topoisomerases as paths in the blown-up graph;

- we obtain a a representation of the configuration spaces as an infinite graph;

- we use mathematical tools to obstruct configurations or specific reactions.

Applications and further work

- add weights to the edges of our graph in order to include a chirality bias;
- add diagrams representing catenanes and edges representing nullification moves.

A local model for the action of site-specific recombination, following Buck and Ishihara*.

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

Thanks to:

- My supervisors for their support and suggestions, and Marco Marengon for his remarks and help.
- I am 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 Unions Horizon 2020 research and innovation programme (grant agreement No 674978).
- DB gratefully acknowledges support from the Leverhulme Trust (Grant RP2013-K-017).

And of course thanks to CMO and to the organisers for this great opportunity!