The pinning ideal of a multiloop

Christopher-Lloyd Simon and Ben Stucky

January 14, 2025

Abstract

A loop is a generic immersion of the circle in a surface, considered up to isotopy. A loop is taut when it is minimally intersecting in its homotopy class. A pinning set of a loop is a set of points P in the surface avoiding the loop, such that the loop is taut in the surface punctured at P. The collection of pinning sets forms a poset under inclusion, endowed with the cardinal function: its minimum defines the pinning number of the loop.

We show that the decision problem associated to computing the pinning number of a plane loop is NP-complete. First, we implement a polynomial algorithm to check if a point-set is pinning, adapting a method of Birman–Series for computing intersection numbers of curves in surfaces. Then we reduce the problem to a boolean formula whose solutions correspond to pinning sets: for this we study immersed discs in detail to improve a theorem of Hass–Scott characterising taut loops, and apply an algorithm of Blank and Frisch. Finally, we reduce the vertex-cover problem for graphs to the pinning problem for plane loops.

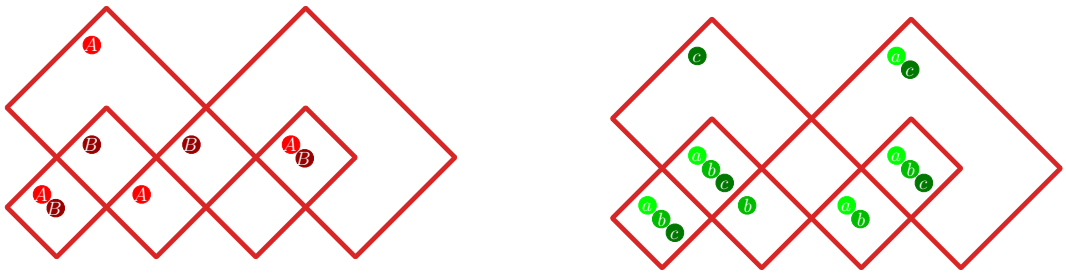


Figure 1: The loop 9^1_5 has 2 optimal pinning sets, and 3 other minimal pinning sets.

Acknowledgements

Both authors wish to thank Yago Antolín and Alan Reid for the invitation and financial support to attend the workshop on orderings and groups at ICMAT in Madrid during summer 2023, where their friendship and collaboration began.

We also recognise the help of Nathan Dunfield who gave us tips for adapting `SnapPy` to our purposes, as well as Gunnar Brinkmann and Brendan McKay who gave us tips and custom plugins for `plantri`. Many of the figures were created with the help of `SnapPy` [CDGW]. We thank Matt Clay for help with Figure 8.

The second author gratefully acknowledges the unwavering support of his partner Madeleine. He is also grateful to the faculty and staff of the department of Mathematical Sciences at Northern Illinois University for welcoming him as a visiting scholar during the Spring 2024 semester, where a large portion of this work was completed.

Last but not least, we are pleased to thank the referee for their thorough reading, pointing out several mistakes in our proofs and providing relevant critiques.

Plan of the paper

0 Introduction

Multiloops and their pinning ideals

Let $\mathbb{F}$ be a closed oriented smooth surface (we will often focus on the case of a sphere, which already contains most of the intricacies). For $P \subset \mathbb{F}$ we write $\mathbb{F}_P = \mathbb{F} \setminus P$. Denote by $\sqcup_1^s \mathbb{S}^1$ a disjoint union of $s \in \mathbb{N}$ oriented circles.

A *multiloop* with s strands is a generic immersion $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}_P$ considered up to orientation preserving diffeomorphisms of the source and the target. Here *generic* means that all multiple points are transverse double-points.

A multiloop γ yields an embedded 4-valent graph $\Gamma \hookrightarrow \mathbb{F}_P$ where $\Gamma = \mathrm{im}(\gamma)$. The *regions* of γ or Γ are the connected components of $\mathbb{F} \setminus \Gamma$, indexed by $R = \pi_0(\mathbb{F} \setminus \Gamma)$. The multiloop γ is *filling* $\mathbb{F}$ when $\mathbb{F} \setminus \Gamma$ is homotopic to a disjoint union of discs.

A *multicurve* is a homotopy class of multiloops. The set of multicurves is endowed with the *self-intersection* function $\mathrm{si}(\gamma)$ counting the minimal number of double-points among its generic representatives. A multiloop is *taut* when it has the minimal number of double-points in its homotopy class.

Definition 0.1 (pinning set)**.** *Consider a multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$ with regions R. We say that a set of regions $P \subset R$ is a* pinning set *if puncturing the surface $\mathbb{F}$ in those regions yields a taut multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}_P$. The pinning sets form the* pinning ideal $\mathcal{PI} \subset \mathcal{P}(R)$*: a sub-poset absorbing under union and containing R.*

A pinning set is called minimal *if it is minimal with respect to inclusion, and* optimal *if it has the minimum cardinal among all pinning sets. The* pinning number $\varpi(\gamma) \in \mathbb{N}$ *is the cardinal of its optimal pinning sets.*

Question 0.2 (goal)**.** *Given a multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$, how (efficiently) can we:*

- *Construct minimal or optimal pinning sets?*
- *Find which regions belong to every pinning set?*
- *Compute the pinning number $\varpi(\gamma)$?*

More generally, we wonder what can be the shape of the pinning ideal of a (multi)loop, and what are its statistics for certain random models.

We will address all of these questions, with a particular emphasis on the following.

Theorem 0.3 (MuLooPinNum)**.** *The MuLooPinNum problem defined by*

> *Instance: A filling multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$, and an integer $p \in \mathbb{N}$.*
> *Question: Does γ have pinning number $\varpi(\gamma) \leq p$?*

is NP-complete, and so is its restriction to loops in the sphere $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{S}^2$.

0.1 Combinatorial group theory: counting self-intersections

A filling multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$ is determined, modulo orientation of its strands, by its associated 4-valent filling graph $\Gamma \subset \mathbb{F}$, which can be encoded by a pair of permutations over its set of half-edges. We will use the size $|\gamma|$ of such an encoding as a measure of complexity: it is proportional to $\mathrm{Card}(R) + |\chi(\mathbb{F})|$.

Algorithm 0.4 (self-intersection)**.** *We will describe an algorithm which has:*

> *Input: A multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$ together with a subset of regions $P \subset R$.*
> *Output: The self-intersection number of the multicurve $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}_P$.*
> *Complexity: Running time $O(|\gamma|^2)$.*

The algorithm first computes a convenient presentation of the free group $\pi_1(\mathbb{F}_R)$. Its generators correspond to the edges of a spanning tree of the dual graph of $\Gamma \subset \mathbb{F}$. The homotopy classes of the loops $\gamma_i \subset \mathbb{F}_R$ correspond to conjugacy classes in that free group, whence to cyclically reduced words in the generators. Filling the regions $Q = R \setminus P$ yields explicit relations providing a presentation of $\pi_1(\mathbb{F}_P)$, and we deduce the conjugacy classes associated to $\gamma_i \subset \mathbb{F}_P$ by applying simple rewriting rules. Finally we adapt an algorithm of Birman–Series and Cohen–Lustig to compute the intersection number between homotopy classes.

Corollary 0.5 (Pinning and PinMin are P)**.** *Given a plane multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{S}^2$ and a subset of regions $P \subset R$, we can successively:*

> *Pinning: certify that P is pinning in time $O(|\gamma|^2)$, and if so*
> *PinMin: check if P is a minimal pinning set in time $O(\mathrm{Card}(P)\,|\gamma|^2)$, if not*
> *PinMinC: construct a minimal pinning set $M \subset P$ in time $O(\mathrm{Card}(P)\,|\gamma|^2)$.*

Remark 0.6 (higher genus)**.** *We will present the algorithm 0.4 for loops in the sphere but it can be adapted to any genus without essential modifications.*

The Corollary 0.5 holds identically in that context.

Corollary 0.7. *The MuLooPinNum problem is in NP.*

0.2 Geometric topology: computing immersed discs

We now focus on loops, namely multiloops with one strand. The [HS85, Theorem 4.2] characterises taut loops in terms of *singular monorbigons*, which are null-homotopic subloops delimited by one or two double-points.

For our purposes, we will need to improve [HS85, Theorem 4.2]. The statement relies on our Definition 2.8 of an immersed monorbigon for a loop γ, which are singular monorbigons that extend to immersions of a disc.

Theorem 0.8 (non-taut loops have immersed monorbigons)**.** *Consider a closed oriented surface $\mathbb{F}$ together with a set of points $P \subset \mathbb{F}$, and a loop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}_P$.*

The loop $\gamma \looparrowright \mathbb{F}_P$ is not taut if and only if it has an immersed monorbigon.

The subset of immersed monorbigons can be computed in polynomial time by an algorithm of Blank [Bla67, Poé95] and Frisch [Fri10]. For each immersed monorbigon, we may associate at most two subsets of regions called *mobidiscs*, whose collection is denoted $\mathrm{MoB}(\gamma) \subset \mathcal{P}(R)$ which can also be computed in polynomial time.

Algorithm 0.9 (pinning mobidiscs)**.** *Given a filling loop $\gamma\colon \mathbb{S}^1 \to \mathbb{F}$ with regions R, we will construct in time $O(\mathrm{Card}(R)^2\,\mathrm{Card}(V)^2)$ a collection $\mathrm{MoB}(\gamma) \subset \mathcal{P}(R)$ with the property that $P \in \mathcal{P}(R)$ is pinning if and only if it intersects every $D \in \mathrm{MoB}(\gamma)$.*

The algorithm 0.9 implies that the pinning sets of the loop γ correspond to the solutions of the monotone conjunctive normal form associated to $\mathrm{MoB}(\gamma) \subset \mathcal{P}(R)$, or equivalently to the vertex-covers of the hyper-graph whose vertices are indexed by the regions R and hyper-edges correspond to the *mobidiscs* $\mathrm{MoB}(\gamma) \subset \mathcal{P}(R)$.

Thus finding the minimum cardinal for these solutions may be solved efficiently with a particular type of SAT-solver.

0.3 From planar graph vertex cover to pinning loops

We know from Corollary 0.7 that the MuLooPinNum problem is in NP.

Algorithm 0.10 (reduction from PLANAR VERTEX COVER to LooPinNum)**.** *We describe a polynomial time algorithm which to a planar graph $G = (V, E)$ associates a plane loop $\gamma\colon \mathbb{S}^1 \to \mathbb{R}^2$ with $O(\mathrm{Card}(E)^2)$ double-points, such that the k-vertex-covers of G are in one to one correspondence with the $(k + 6\,\mathrm{Card}(E))$-pinning-sets of γ.*

Note that the LooPinNum problem contains as a sub-problem the problem of pinning loops in the sphere, so its complexity is at least as hard as the PLANAR VERTEX COVER problem, which is known to be NP-complete.

Corollary 0.11. *The LooPinNum problem is NP-hard, even in the sphere.*

To justify the reduction, we will derive from [HS85, Theorem 4.2] a Lemma 2.6 ensuring a sufficient condition that set of regions is pinning in terms of linking numbers of singular monorbigons with pins.

0.4 The pinning ideal of a multiloop

Consider a multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$ with regions R. The pinning ideal $\mathcal{PI} \subset \mathcal{P}(R)$ carries all the information related to the pinning problem for the multiloop. As an ideal, it is generated by the minimal pinning sets. It contains the *pinning semi-lattice*, which is the sub-poset obtained by all possible unions of minimal pinning sets.

Using Algorithm 0.4 and some functionalities from `plantri` [BM07], we computed the pinning ideals and semi-lattices of all irreducible indecomposible spherical multiloops with at most 12 regions (unoriented and up to reflection), and the statistics of certain numerical parameters. The results are available in the LooPindex [SS24b].

We make a few observations in section 3, either suggesting certain heuristics for approximation algorithms, or invalidating some naïve conjectures about the behaviour of pinning ideals. Let us record the main questions and observations.

Question 0.12 (structure of pinning ideals)**.** *Which ideals arise as pinning ideals of filling loops or multiloops?*

How do pinning ideals of filling multiloops behave under Reidemeister moves, flypes, and crossing resolutions?

How do pinning ideals of loops behave under the operations of spheric-sums and toric-sums introduced in [Sim23]?

The Figures 26 and 27 show that the pinning number can change under $R3$ moves and flypes as well as mutations, even for indecomposible loops in the sphere.

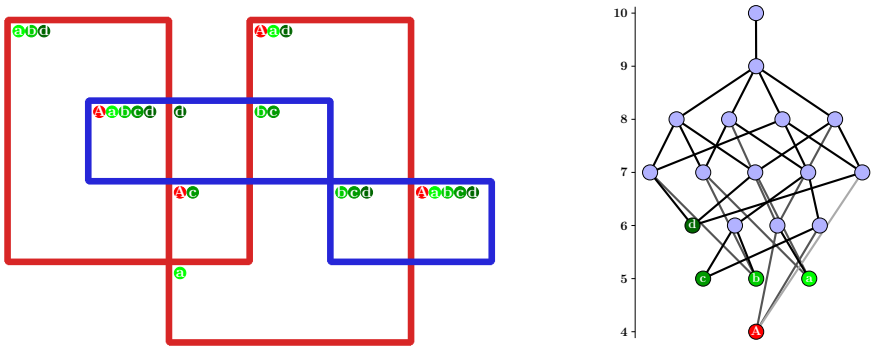


Figure 2: The multiloop 10^2_{16} and its pinning semi-lattice obtained by unions of minimal pinning sets (in red or green), together with the set of all regions.

0.5 Further directions of research and related works

Multiloops versus loops. The validity of Algorithm 0.9, reducing LooPinNum to the vertex cover problem for hypergraphs, relies on results which hold for loops but sometimes fail for multiloops with ≥ 2 strands.

Question 0.13 (loops versus multiloops)**.** *Does the MuLooPinNum problem belong to a strictly larger complexity class than LooPinNum? This would be a finer complexity distinction since the present work shows that both are NP-complete.*

By Algorithm 0.9, the pinning ideal of a loop γ with regions R is isomorphic to the solution ideal of the monotone conjunctive normal form associated to $\mathrm{MoB}(\gamma) \subset \mathcal{P}(R)$. The pinning ideal of multiloops is more mysterious.

Question 0.14 (counting strands)**.** *Can one find a topological notion generalising mobidiscs to multiloops, which would yield a basis for the pinning ideal ?*

Can we compute or estimate lower bounds for the number of strands of a multiloop from its pinning ideal?

Question 0.15 (multi-simple-loops)**.** *We may restrict the MuLooPinNum problem to* multi-simple-loops*, namely whose double points only involve distinct strands.*

It can be approached using methods similar to those in subsection 2.1, relying on [FM12, Proposition 1.7]. Indeed, one may show that the problem is in P if we restrict to $s \leq 3$ strands, and NP-complete if we do not restrict $s \in \mathbb{N}$. Is there a critical value of $s \in \mathbb{N}$ for which it is NP-complete?

Approximate solutions and heuristics. When facing an NP-complete problem, one wonders whether it is possible to find approximate solutions in polynomial time.

Question 0.16 (approximability)**.** *Do there exist $\epsilon_1, \epsilon_2 > 0$ and a polynomial time algorithm which: given a multiloop γ and $k \in \mathbb{N}$, either confirms that $\varpi(\gamma) > k$ or else constructs $P \subset R$ with $\operatorname{Card} P \leq (1+\epsilon_1)k$ and $\operatorname{si}(\gamma, P) > (1-\epsilon_2).\operatorname{si}(\gamma, R)$?*

A weaker expectation would be an algorithm based on certain heuristics which computes an approximate solution with high probability (that is on most entries).

Question 0.17 (heuristics)**.** *Are there heuristics building on the idea of pinning regions of small degree and certain embedded monorbigons which lead to algorithms that find approximate solutions with high probability? We discuss some in section 3.*

Pinning statistics for random multiloops.

Question 0.18. *Consider a random model for unoriented multiloops in closed orientable surfaces (in fixed genus or not), by choosing a distribution on a set of 4-valent maps whose weights are inverse proportional to cardinal of the automorphism groups. What is the distribution of:*

the size of an optimal pinning set, that is the pinning number?
the number of optimal pinning sets, and of minimal pinning sets?

Geometric algorithms. Given a multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$ and a subset of regions $P \subset R$ such that $\chi(\mathbb{F}_P) < 0$, one could apply geometric algorithms to check whether P is pinning, by applying a curve shortening flow [HS94, DL19].

For instance, choose a negatively curved metric on $\mathbb{F}_P$ and apply a geometric curve-shortening flow algorithm relying on gradient descent: the multiloop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}_P$ is not taut if and only if it develops a self-tangency or a cusp under this flow.

Remark 0.19 (taut multiloops are shortest geodesics)**.** *It follows from [NC01] that a multiloop is taut if and only if it is isotopic to a union of shortest geodesics for some Riemannian metric. We emphasise that the term shortest is important here.*

However, the multiloops arising from geodesics in negatively curved or hyperbolic surfaces are much more subtle to describe (see for instance [HS99]).

Other variational problems. For a filling multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$, consider the increasing functionals on its lattice of regions $\mathcal{P}(R)$, which to a set $P \in R$ associate:

- its cardinal $\operatorname{Card}(P) = \sum_{P_i \in P} 1$, its *total-degree* $\operatorname{Deg}(P) = \sum_{P_i \in P} \deg(P_i)$;
- minus the self-intersection number $-\operatorname{si}(\gamma, P)$ of the multicurve $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}_P$.

Note that $\operatorname{Card}(P)$ and $\deg(P)$ are valuations in the sense of lattice theory.

The pinning problems addressed in this work concern the local and global minima in $\mathcal{P}(R)$ of the valuation $\operatorname{Card}(P)$ in restriction to the level-set $\{P \in R \mid \operatorname{si}(\gamma, P) = n\}$ where n is the number of double-points of the multiloop γ.

If we change $\operatorname{Card}(P)$ for $\operatorname{Deg}(P)$, then we may apply Algorithm 0.4 to deduce an analogous Corollary 0.5 showing that certain pinning problems are in P.

Question 0.20 (Changing the valuation)**.** *We believe that the* Deg*-weighted analogue* LooPinDeg *of the* LooPinNum *problem is also* NP*-complete.*

Can one adapt the Algorithm 0.10 to reduce the problem of finding a vertex cover for a plane graph to that of finding the minimum total region degree of a corresponding plane loop?

Question 0.21. *There are other interesting functionals to extremize, which from a game theoretic perspective can be thought of as varying the cost or gain function, and that may yield other interesting topological invariants of filling multiloops.*

What is the complexity of computing the minimal values of the monotonous functionals $\mathrm{Card}(P) - \mathrm{si}(\gamma, P)$ *or* $\deg(P) - \mathrm{si}(\gamma, P)$*? What about non-monotonous functionals of the form* $4\,\mathrm{Card}(P) - \deg(P) - 3\,\mathrm{si}(\gamma, P)$*?*

Which monotonous functions depending on $(\mathrm{Card}(P), \deg(P))$ *and* $-\mathrm{si}_P(\gamma)$ *have global minimal values that are in* P *to compute? What are the topological interpretations of these quantities?*

Pinning invariants for homeomophisms. Let $\psi \in \mathrm{Homeo}_0(\mathbb{F})$ be a homeomorphism which is isotopic to the identity. For a multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$, we define

$$\varpi(\psi, \gamma) = \min\{\varpi(\gamma_0 \cup \psi(\gamma_0)) \mid \text{generic immersions } \gamma_0 \text{ in the isotopy class } \gamma\}$$

This quantity measures the minimum cardinal of a puncture-set $P \subset \mathbb{F}$ yielding a topological model for the action of a class $[\psi] \in \mathrm{Map}(\mathbb{F}_P)$ on $\gamma \subset \mathbb{F}_P$. It may be compared with $\varpi(\gamma) + \varpi(\psi(\gamma)) = 2\varpi(\gamma)$ Now, for a certain class of multiloops $\mathcal{C}$ (say all multiloops, or those without double-points), we define

$$\varpi(\psi, \mathcal{C}) = \min\{\varpi(\psi, \gamma) \mid \gamma \in \mathcal{C}\}.$$

The case where $\mathcal{C}$ are loops with no double points appears closely related to the work [BHW22] about quasi-morphisms on the groups $\mathrm{Homeo}_0(\mathbb{F})$ and $\mathrm{Diff}_0(\mathbb{F})$. Indeed, their main tool is to relate the action of $\psi \in \mathrm{Diff}_0(\mathbb{F})$ on the fine curve graph of $\mathbb{F}$ to the actions of $[\psi] \in \mathrm{Map}(\mathbb{F}_P)$ on the surviving curve graphs of $\mathbb{F}_P$. We believe that the quantities $\varpi(\psi, \mathcal{C})$ may be of interest for such investigations.

Complexity of link invariants. The fact that MuLooPinNum and LooPinNum are NP-complete follows a general trend of complexity results regarding the computation of numerical invariants in topology.

Indeed, many numerical invariants in topology are NP-hard to decide, such as the crossing number of a knot [dMSS20] the unlinking number of a knot [KT21], and all so-called intermediate invariants of knots and links [dMRST21].

The problem of unknot recognition is noteworthy in that it belongs to NP and co-NP [HLP99]. Computing the genus for a knot in a general 3-manifold is NP-hard [AHT22] and for a knot in the sphere it is co-NP [Lac21].

As mentioned in [dMRST21], upper bounds on the complexity of deciding numerical invariants of knots and links are often elusive. In contrast, the present work shows that certain subtle measures of the complexity of multiloops are more tractable.

1 Group theory: counting self-intersections

In the first subsection we encode filling multiloops $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$ by 4-valent maps. We used this encoding extensively in the implementation of our algorithms, but one may gloss over it without much hindrance to understanding the rest of our work.

In the second subsection we use this encoding to derive a presentation of the fundamental group $\pi_1(\mathbb{F}_P)$ and the conjugacy classes associated to the strands of γ.

In the third subsection we explain an algorithm to compute the self-intersection number of the multicurve associated to γ in $\mathbb{F}_P$.

Altogether they imply Corollary 1.12, saying certain pinning problems are in P.

1.1 Encoding filling multiloops as 4-valent maps

Let us first explain the permutation encoding of filling graphs and filling multiloops. A general reference for this is [LZ04].

A *map* $\Gamma \subset \mathbb{F}$ is an embedded graph in our closed oriented surface whose complement is (homotopic to) a disjoint union of discs, considered up to orientation preserving diffeomorphisms.

Lemma 1.1. *A map $\Gamma \subset \mathbb{F}$ with n edges corresponds to a pair of permutations $\epsilon, \sigma \in \mathfrak{S}_{2n}$ considered up to conjugacy, such that:*

- *ϵ has orbits of size 2 corresponding to the edges*
- *σ has orbits corresponding to the vertices*
- *$\varphi = (\sigma\epsilon)^{-1}$ has orbits corresponding to the regions*

The degree *of a vertex or region is the cardinal of the corresponding orbit for σ or φ. The Euler characteristic of $\mathbb{F}$ is the alternated sum of the number of orbits of $\sigma, \epsilon, \varphi$. The dual map of $\Gamma \subset \mathbb{F}$ is obtained by exchanging σ and φ.*

Remark 1.2 (Computing regions)**.** *For a map $\epsilon, \sigma \in \mathfrak{S}_{2n}$ and a union of orbits of $\varphi = (\epsilon\sigma)^{-1}$, the corresponding union of closed regions and its interior have homotopy types which may be computed in polynomial time on n. For example, we often wish to determine their number of connected components and their genera.*

A filling multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$ is determined, modulo orientation of its strands, by its associated 4-valent map $\Gamma \subset \mathbb{F}$. Note that the orientation of γ corresponds to an orientation of the edges of Γ which go straight on at each vertex.

Lemma 1.3. *A 4-valent map $\Gamma \subset \mathbb{F}$ with n edges corresponds to a pair of permutations $\epsilon, \sigma \in \mathfrak{S}_{2n}$ considered up to conjugacy, such that:*

- *ϵ has orbits of size 2 corresponding to the edges*
- *σ has orbits of size 4 corresponding to the vertices*
- *$\varphi = (\sigma\epsilon)^{-1}$ has orbits corresponding to the regions*

Moreover, the orbits of $\delta = \epsilon\sigma^2$ define the oriented strands of Γ, and these are paired by ϵ to form its unoriented strands. If a multiloop γ yields the map Γ, the oriented strands of γ select one orbit of δ in each ϵ-pair.

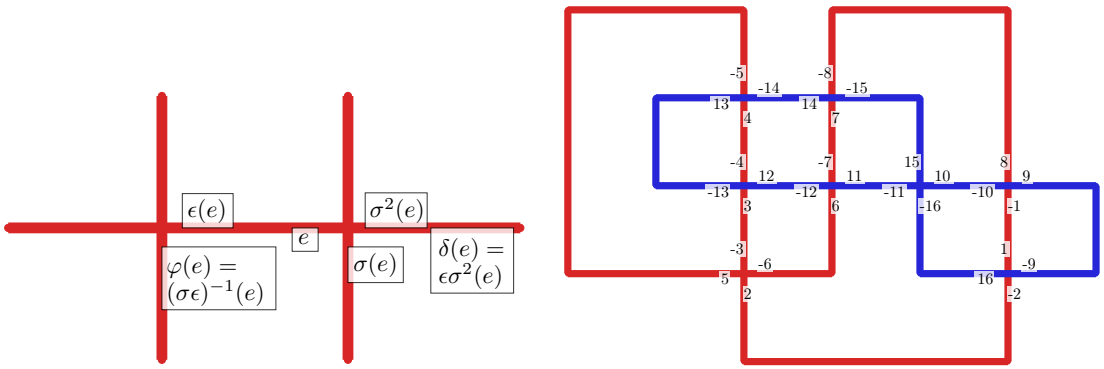


Figure 3: A map with labelled half-edges. The permutations $\epsilon, \sigma, \varphi, \delta$ are:

$$\begin{aligned}
\epsilon =&(-1,1)(-2,2)(-3,3)(-4,4)(-5,5)(-6,6)(-7,7)(-8,8)\\
&(-9,9)(-10,10)(-11,11)(-12,12)(-13,13)(-14,14)(-15,15)(-16,16),\\
\sigma =&(9,8,-10,-1)(5,2,-6,-3)(3,12,-4,-13)(13,4,-14,-5)\\
&(14,7,-15,-8)(10,15,-11,-16)(1,16,-2,-9)(6,11,-7,-12),\\
\varphi =&(-1,-9)(-2,5,-14,-8,9)(-3,-13,-5)(-4,13)(-6,-12,3)\\
&(-7,14,4,12)(-10,-16,1)(-11,6,2,16)(-15,10,8)(7,11,15),\\
\delta =&(1,2,3,4,5,6,7,8)(-8,-7,-6,-5,-4,-3,-2,-1)\\
&(9,10,11,12,13,14,15,16)(-16,-15,-14,-13,-12,-11,-10,-9).
\end{aligned}$$

The following observation will serve only to find all (multi)loops in the sphere whose regions all have degree ≥ 3, given their total number of regions.

Lemma 1.4. *For a 4-valent map $\Gamma \subset \mathbb{F}$ in a surface of Euler characteristic χ, the degree of the regions $R_i \in R$ satisfy*

$$\textstyle\sum_{R_i \in R}(\deg(R_i) - 4) = -4\chi.$$

In particular, a multiloop in the torus has average region-degree 4, and a multiloop in the sphere whose regions have degree ≥ 3 has at least 8 regions of degree 3.

Proof. Denote the vertices by V, the edges by E and the regions by R. The 4-valency of the graph implies that $2|E| = 4|V|$. The surface embedding yields $\sum_R \deg R_i = 2|E| = 4|V|$. The filling property computes the Euler characteristic $\chi = |V| - |E| + |R|$ so that $|R| = \chi + |V|$. Combine these, we find that for all $m \in \mathbb{N}$ we have $\sum_R(\deg(R_i) - m) = (4-m)|V| - m\chi$, whence $\sum_R(\deg R_i - 4) = -4\chi$. □

Remark 1.5 (Irreducible, indecomposible)**.** *Following [CZ16], a plane map is* irreducible *when it is not disconnected by removal of a vertex and* indecomposible *when it is not disconnected by cutting two distinct edges.*

These notions are the respective analogues of a link diagram having no nugatory crossings and being prime.

1.2 From multiloops to words in the free group

In this subsection we restrict, for simplicity of the exposition, to the case $\mathbb{F} = \mathbb{S}^2$.

Consider a multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$ together with a set of regions $P \subset R$. We address the question of checking whether P is a pinning-set, namely whether γ is a taut loop in the surface $\mathbb{F}_P = \mathbb{F} \setminus P$.

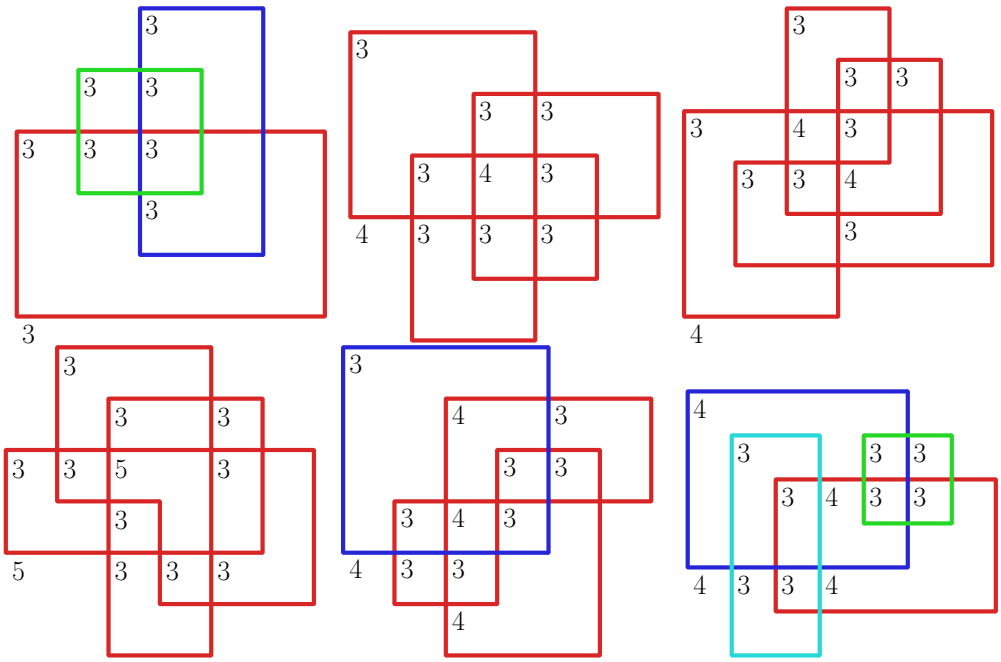


Figure 4: All irreducible indecomposible spherical multiloops with at most 12 regions and all region-degrees ≥ 3 (unoriented, up to reflection; see [OEI24, A078666]), with regions labelled by their degrees.

The multiloops γ admitting a pinning-set of cardinal ≤ 1 are those with no double-points. From now on we assume that $p - 1 = \mathrm{Card}(P) > 0$.

We choose a base point $\infty \in \mathbb{F} \setminus \Gamma$ in a region belonging to P, called the unbounded region. Hence the fundamental group $\pi_1(\mathbb{F}_P, \infty)$ is free of rank p.

Remark 1.6 ($\varpi(\gamma) = 2$)**.** *A loop γ with $\varpi(\gamma) = 2$ is a "snail", that is a generic perturbation of the n-th power of a circle which has $n - 1$ double-points. The multiloops γ with $\varpi(\gamma) = 2$ are the disjoint unions of $s_0 \geq 0$ circles and $s_1 > 0$ concentric snails.*

Recall that a filling multiloop $\gamma \colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$ is determined, modulo orientation of its strands, by its associated 4-valent plane graph $\Gamma \subset \mathbb{F}$. In its dual graph $\Gamma^* \subset \mathbb{F}$, we choose to include ∞ as a vertex which we call the root.

Fix a spanning tree T^* for the dual graph $\Gamma^* \subset \mathbb{F}$, orient its edges outwards from the root vertex ∞, and co-orient them using the orientation of the disc. We also order its vertices and its edges by a depth first search from ∞.

Step 1: $\pi_1(\mathbb{F}_R, \infty)$**.** We first compute an explicit presentation for the free group $\pi_1(\mathbb{F}_R, \infty)$ where $R = \pi_0(\mathbb{F} \setminus \Gamma)$ consists of all regions, from which we may easily express the conjugacy classes of the strands γ_j. The construction follows Figure 5.

Consider simple loops $\tau_1, \ldots, \tau_r$ which intersect only at their base point ∞, such that τ_j intersects T^* only at the edge with label r and in the direction prescribed by its co-orientation. The free group $\pi_1(\mathbb{F}_R, \infty)$ is freely generated by $\tau = \{\tau_1, \ldots, \tau_r\}$. The symmetric generating set $\tau^\pm = \{\tau_j^{\pm 1} \mid 1 \leq j \leq r\}$ inherits the cyclic order given by the emergence of paths from ∞. The complement $\mathbb{F} \setminus T^*$ is a polygon whose $2r$ edges are labelled by $\tau^\pm$ in accordance with its cyclic order. It lifts in the universal cover of $\mathbb{F}_R$ to a fundamental domain under the action of $\pi_1(\mathbb{F}_R, \infty)$, whose generators $\tau^\pm$ identify its edges two-by-two.

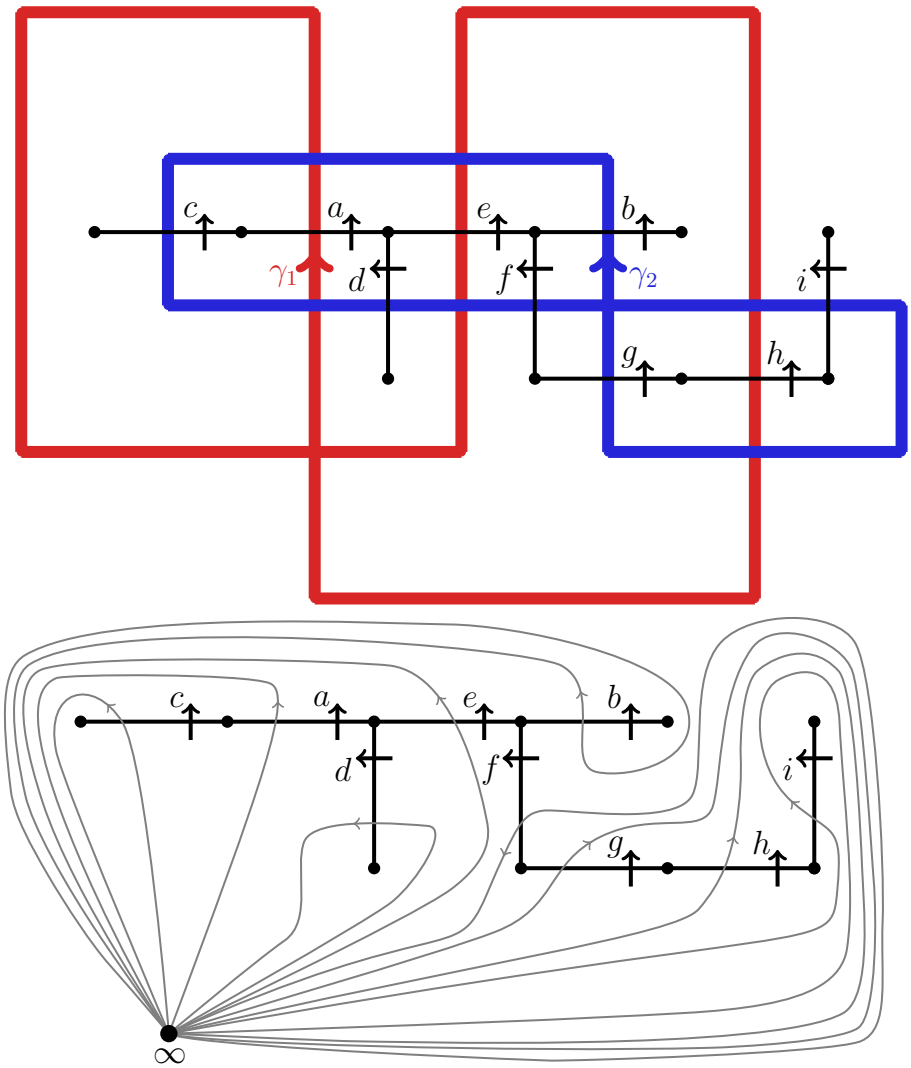


Figure 5: Top: Choosing a co-oriented labelled spanning tree T in the dual graph of γ. Bottom: We have that $\pi_1(\mathbb{S}^2_R, \infty) \cong F_9 = \langle a, b, c, d, e, f, g, h, i \rangle$ since the sphere punctured at the vertices of T retracts by deformation onto the dual bouquet. We also see that $\gamma_1 = aeH$, and $\gamma_2 = bCDFIg$ up to conjugacy (where capital letters denote inverses). Note the induced cyclic order on the symmetrized generating set: $hgFedDacCAEBbfGHIi$.

Note that each strand γ_j of γ corresponds to a conjugacy class in $\pi_1(\mathbb{F}_R, \infty)$, hence to a unique reduced cyclic word over $\tau^{\pm}$ which may be computed by recording the sequence of edges of T^* that it crosses, minding the co-orientation to determine the exponent of each generator.

Step 2: $\pi_1(\mathbb{F}_P, \infty)$. Now for a subset of regions $P \subset R$, we deduce a presentation of the free group $\pi_1(\mathbb{F}_P, \infty)$, and the new reduced cyclic words associated to the conjugacy classes of the strands γ_j. The construction follows Figure 6.

By the van Kampen Theorem, the group $\pi_1(\mathbb{F}_P, \infty)$ is the quotient of $\pi_1(\mathbb{F}_R, \infty)$ by the normal subgroup generated by small loops only surrounding the punctures to be filled, associated to the set of regions $Q = R \setminus P$.

A region $R_j \in R$ corresponds to a non-rooted vertex of T^*, which has one incoming edge and a certain number of outgoing edges. Let ρ_j be the word over $\tau^{\pm}$ obtained by reading the cutting sequence of T^* of a loop winding once around R_j, starting at the incoming edge. Thus ρ_j is a product of $\deg(R_j)$ generators in $\tau^{\pm}$ which are all distinct even up to inversion. We deduce the presentation:

$$\pi_1(\mathbb{F}_P, \infty) = \pi_1(\mathbb{F}_R, \infty) \mod \lhd \{\rho_j \mid R_j \in Q\} \rhd$$

in which every relator $\rho(R_j)$ yields a rewriting rule of the form $\tau_{j_1} \mapsto \tau_{j_d}^{\mp} \cdots \tau_{j_2}^{\mp}$ which we obtain by solving the equation $\rho(R_j) = 1$.

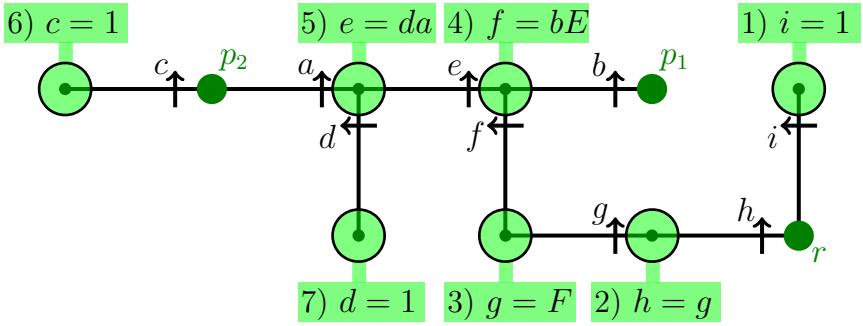


Figure 6: Filling punctures, adding relations, applying rewriting rules in DFS order. We consider the set of pins $P = \{r, p_1, p_2\}$ and rewrite from root r. We have that $\pi_1(\mathbb{S}^2_P, \infty) \cong \langle a, b, c, d, e, f, g, h, i \rangle \mod \lhd \{i = 1, h = g, g = F, f = bE, e = da, c = 1, d = 1\} \rhd$, and we rewrite γ_1 and γ_2 as follows: $\gamma_1 = aeH \mapsto_2 aeG \mapsto_3 aef \mapsto_4 aebE \mapsto_5 adabAD \mapsto_7 aabA \sim ab$. $\gamma_2 = bCDFIg \mapsto_{1,3} bCDFF \mapsto_4 bCDeBeB \mapsto_5 bCDdaBdaB = bCaBdaB \mapsto_{6,7} baBaB \sim aBa$.

Moreover, for an arbitrary word α over $\tau^{\pm}$, we may apply these rewriting rules in the order prescribed by a depth first search from ∞ to obtain a canonical representative of α in $\pi_1(\mathbb{F}_P, \infty)$ which uses none of the generators associated to regions in Q. This holds in particular for the words associated to the strands of γ.

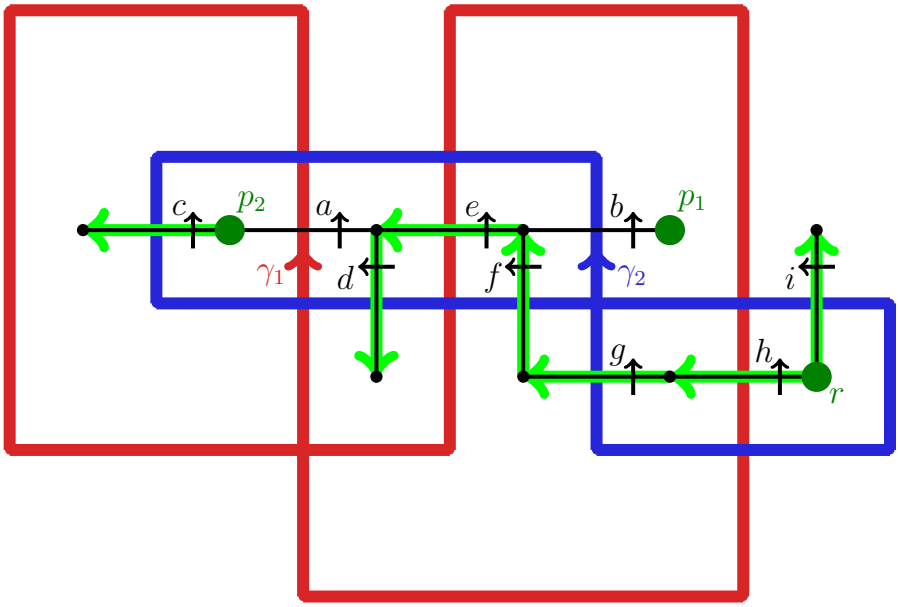


Figure 7: The fundamental domain of $\mathbb{S}^2_R$ collapses to that of $\mathbb{S}^2_P$. We slide strands of γ_1 and γ_2 "downstream" from the root r to avoid all edges "upstream" from elements of $R \setminus P$, then collapse these edges. The cyclic order on the remaining generators is inherited from the original: $hgFedDacCAEBbfGHIi \mapsto aABb$.

Observe that the remaining generators $\tau_P^{\pm} = \{\tau_j^{\pm} \mid R_j \in P\}$ of $\pi_1(\mathbb{F}_P, \infty)$ inherit a cyclic order from $\tau^{\pm}$, which is compatible with their action on a lift of $\mathbb{F}_P \setminus T^*$ to a polygonal fundamental domain in the universal cover of $\mathbb{F}_P$. See Figures 7 and 8.

Conclusion. Let us summarize and emphasise some details of the whole procedure.

In short, given a multiloop $\gamma\colon \cup_1^s \mathbb{S}^1 \to \mathbb{R}^2$ and a subset of p regions $P \subset R$, we compute in time $O(|\gamma|)$ a presentation of the cyclically ordered group $\pi_1(\mathbb{F}_P)$ and a set of words representing the conjugacy classes of the strands γ_i.

The cyclically ordered presentation consists of the generating set $\tau_P = \{\tau_j\}_1^p$ for $\pi_1(\mathbb{F}_P)$ labelled by the bounded regions in P, which in our case is free from relations, together with a cyclic order on the symmetric set $(\tau_P)^\pm = \{\tau_j^{\pm 1}\}_1^p$.

Moreover for $P_1 \subset P_2 \subset R$, the natural quotient map $\pi_1(\mathbb{F}_{P_2}) \to \pi_1(\mathbb{F}_{P_1})$ is order preserving, and our presentations for these cyclically ordered groups (which we described by the relations ρ_j corresponding to elements in $Q = P_2 \setminus P_1$) are functorial in the sense that the cyclic order on $(\tau_{P_1})^\pm$ is inherited by that of $(\tau_{P_2})^\pm$.

Remark 1.7 (Winding numbers)**.** *For a loop $\alpha\colon \mathbb{S}^1 \looparrowright \mathbb{F}_P$ with $\infty \in P$ as above, every puncture $o \in P$ has a winding number (defined as the linking number between the 1-cycle α and the relative 0-chain $[o] - [\infty]$ in the oriented surface $\mathbb{F}_P \setminus \{\infty\}$).*

This yields a winding number function $w_P\colon \pi_1(\mathbb{F}_P) \to \mathbb{Z}^P$ which corresponds to the abelianisation $\mathcal{F}_p \to \mathbb{Z}^p$. Given our presentation of $\pi_1(\mathbb{F}_P)$ as above, this winding number function may be computed in time $\mathrm{Card}(P)^2$ as the matrix of winding numbers between the generators τ_j and the punctures P_i.

Remark 1.8 (higher genus)**.** *This algorithm adapts to filling multiloops $\gamma\colon \sqcup\mathbb{S}^1 \subset \mathbb{F}_P$ in closed surface of higher genus.*

For this one may choose a base point $\infty \in \mathbb{F}$ which is (close to) a puncture, and perform cuts along disjoint simple arcs between the punctures to represent $\mathbb{F}$ as the identification of the sides of a $2n$-gon whose vertices are removed. This leads to a presentation for $\pi_1(\mathbb{F} \setminus \{\infty\}) = \mathcal{F}_n$ by a free set of generators, together with a cyclic order on the corresponding symmetric generating set. The complexity is still $O(|\gamma|)$, keeping in mind that since γ is filling we have $|\gamma| = \mathrm{Card}(R) + |\chi(\mathbb{F})|$.

1.3 Computing self-intersection of homotopy classes

In this subsection we explain an adaptation of the algorithm proposed in [CL87], extending [BS84, BS87] and [Rei62], to compute the self-intersection number of a multicurve γ in $\mathbb{F}_P$. For clarity and completeness, we include a description of this algorithm in relation with our previous presentation so as to explain some details not covered in those references.

It takes as input a free group $\mathcal{F}_p$ of rank p over a set $\tau = \{\tau_j \mid 1 \le j \le p\}$ with a cyclic order on the symmetric set $\tau^\pm$, and a finite set γ of reduced words in these generators. It returns the self-intersection number:

$$\mathrm{si}(\gamma) = \textstyle\sum_k \mathrm{si}(\gamma_k) + \sum_{i \neq j} \mathrm{i}(\gamma_i, \gamma_j).$$

Remark 1.9. *For curves α, β in $\mathbb{F}_P$ and $a, b \in \mathbb{N}$ we have $\mathrm{i}(\alpha^a, \beta^b) = ab \cdot \mathrm{i}(\alpha, \beta)$. Note that the self-intersection number of a curve satisfies $\mathrm{si}(\alpha) = \frac{1}{2}\mathrm{i}(\alpha, \alpha)$.*

If α is primitive then for all $n \in \mathbb{N}^*$ we have $\mathrm{si}(\alpha^n) = n^2\,\mathrm{si}(\alpha) + (n-1)$. Conversely, if there exists $n \in \mathbb{N}_{>1}$ such that $\mathrm{si}(\alpha^n) = n^2\,\mathrm{si}(\alpha) + (n-1)$ then α is primitive (because otherwise $\alpha = \beta^m$ for some primitive β and $m \in \mathbb{N}_{>1}$, and applying the previous formula to $\alpha^n = \beta^{mn}$ would lead to $m \le 1$, a contradiction).

Description and justification of the algorithm

Combinatorial action. The elements of $\mathcal{F}_p$ correspond to words on the alphabet $\tau^\pm$ that are *reduced* in the sense that consecutive letters are not mutually inverse.

The Cayley graph of $(\mathcal{F}_p, \tau^\pm)$ is the infinite $2p$-regular tree $\mathcal{T}_p$ with a base vertex 1, whose edges around each vertex are labelled by $\tau^\pm$. We identify the boundary $\partial\mathcal{T}_p$ with the subset of one sided sequences $(\tau^\pm)^{\mathbb{N}}$ which are reduced. The cyclic order on $\tau^\pm$ extends to a cyclic order $\mathrm{cord}(x, y, z)$ on the boundary $\partial\mathcal{T}_p$.

The group $\mathcal{F}_p$ acts by left-translation on its Cayley graph $\mathcal{T}_p$ by left translation, thus on its boundary. A non-trivial element $\gamma \in \mathcal{F}_p \setminus \{1\}$ yields, by periodisation of the words γ, γ^{-1}, two points $\gamma^+, \gamma^- \in \partial\mathcal{F}_p$, which correspond to the attractive and repulsive fixed points for the action of γ on $\partial\mathcal{F}_p$.

Geometric action. Consider a collection of $2p$ disjoint geodesics in the hyperbolic plane $\mathbb{HP}$ with distinct endpoints (they appear very close in the figure), labelled in cyclic order by $\tau^\pm$ (as we would obtain by lifting the dual tree $T^* \subset \mathbb{F}_P$ from Figure 5 to the universal cover $\mathbb{HP} \to \mathbb{F}_P$). Let τ_j act as the hyperbolic isometry sending the geodesic τ_j^{-1} to the geodesic τ_j^{+1} to obtain a free action of $\mathcal{F}_p$ on $\mathbb{HP}$, with fundamental domain an open polygon bounded by the geodesics $\tau_j^\pm$.

This faithfully represents $\mathcal{F}_p$ as a subgroup of $\mathrm{Isom}^+(\mathbb{HP})$, its Cayley graph $\mathcal{T}_p$ as the dual tree of the tessellation by translates of the fundamental polygon in $\mathbb{HP}$, and its boundary $\partial\mathcal{T}_p$ as Cantor subspace in the circle boundary $\partial\mathbb{HP}$.

An non-trivial $\gamma \in \mathcal{F}_p$ acts by hyperbolic translation on $\mathcal{T}_p \subset \mathbb{HP}$ along an axis (γ^-, γ^+) joining its repulsive and attractive fixed points in the boundary $\partial\mathcal{T}_p \subset \mathbb{HP}$.

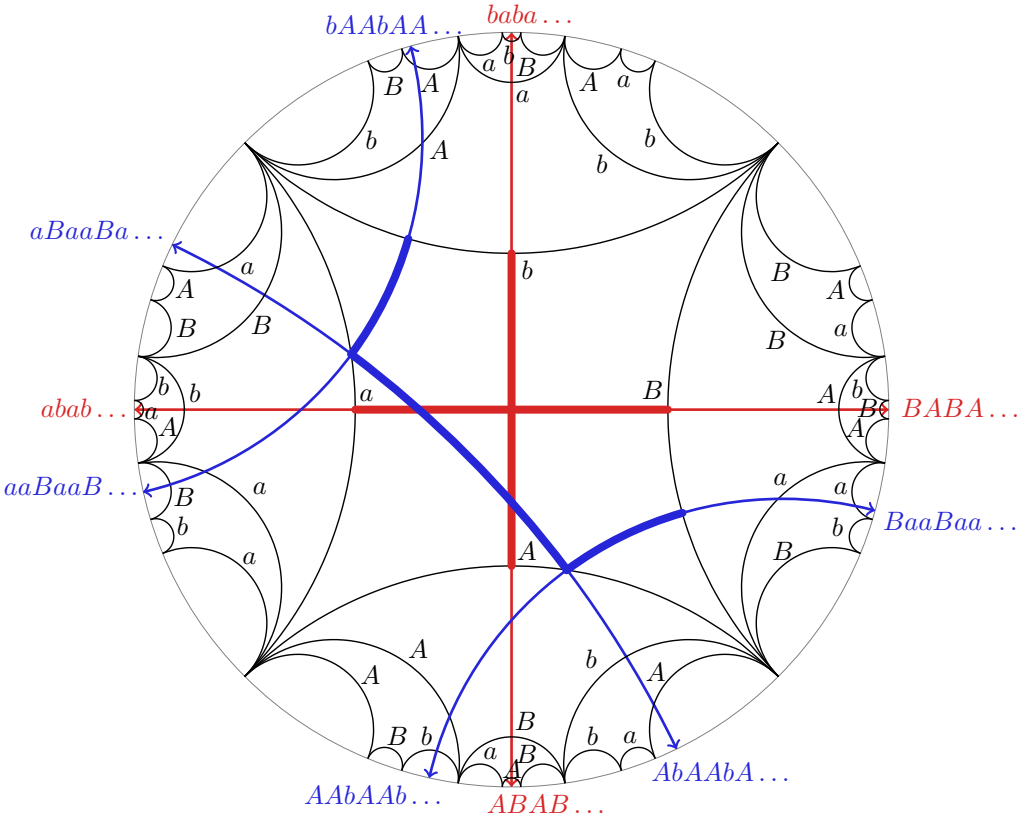


Figure 8: Illustration of Equation (cross-val) for the example in Figure 6, with lifts of $\gamma_1 = ab$ in red and lifts of $\gamma_2 = aBa$ in blue.

Cyclically reduced words. Recall that elements of $\mathcal{F}_p$ correspond to reduced words on the alphabet $\tau^\pm$. A word on $\tau_\pm$ is called *cyclically reduced* when all its cyclic permutations are reduced. Geometrically, the reduced word associated to a hyperbolic $\gamma \in \mathcal{F}_p$ is cyclically

reduced if and only if its hyperbolic axis intersects the fundamental region. Hence a (non-trivial) conjugacy class corresponds in $\mathcal{F}_p$ to the cyclic permutations of a cyclically reduced word on $\tau^{\pm}$, namely the representatives of its hyperbolic translations whose axis intersects the fundamental domain. We will denote by σ the cyclic shift of words which moves the first letter at the end.

ALGORITHM 1: For $\gamma \in \mathcal{F}_p$ represented as a word on $\tau^{\pm}$ of length $\mathrm{len}(\gamma)$, we can find a cyclically reduced representative for its conjugacy class in $O(\mathrm{len}(\gamma))$ steps.

Primitive roots. For an element $\gamma \in \mathcal{F}_p$, its primitivity exponent is the largest $n \in \mathbb{N}_{>0}$ for which there exists $\gamma_0 \in \mathcal{F}_p$ such that $\gamma = \gamma_0^n$: such a γ_0 is unique and called the primitive root of γ. A non-trivial conjugacy class in $\mathcal{F}_p$ is primitive when the cyclically reduced words that represent it are primitive words.

ALGORITHM 2: For a cyclically reduced word γ on $\tau^{\pm}$ of length $\mathrm{len}(\gamma)$, we can compute its primitive root and primitivity exponent in time $O(\mathrm{len}(\gamma))$.

Intersection number. For non-trivial $\alpha, \beta \in \mathcal{F}_p$ acting on $\mathcal{T}_p \subset \mathbb{HP}$ with axes (α^-, α^+) and (β^-, β^+), define their crossing number $\mathrm{cross}(\alpha, \beta) \in \{-1, 0, +1\}$ by:

$$\mathrm{cross}(\alpha, \beta) = \tfrac{1}{2} \left(\mathrm{cord}(\alpha^+, \beta^+, \alpha^-) - \mathrm{cord}(\alpha^+, \beta^-, \alpha^-) \right).$$

This can be computed from their corresponding words in time $O(\mathrm{len}(\alpha) + \mathrm{len}(\beta))$.

For non-trivial primitive $\alpha, \beta \in \mathcal{F}_p$, define their *order of contact* $\mathrm{val}(\alpha, \beta)$ as the number translates of the fundamental domain which they enter and leave from the same sides (the union of which is contractible by convexity of the hyperbolic metric). If α, β correspond to cyclically reduced primitive words, then denoting $w^{\pm}$ the longest common prefix of the infinite words $\alpha^{\pm}, \beta^{\pm} \in \partial \mathcal{F}_p$

$$\mathrm{val}(\alpha, \beta) = \mathrm{len}(w^+) + \mathrm{len}(w^-)$$

This can be computed from the finite words α, β in time $O(\mathrm{len}(\alpha) + \mathrm{len}(\beta))$. Indeed, we have $\mathrm{len}(w^{\pm}) \geq \mathrm{len}(\alpha) + \mathrm{len}(\beta)$ if and only if α, β have the same primitive roots, hence are equal by the primitivity assumption.

Proposition 1.10 (Intersection)**.** *For non-trivial primitive conjugacy classes in the cyclically ordered group $(\mathcal{F}_p, \tau^{\pm})$ represented by cyclically reduced words α, β on $\tau^{\pm}$, the intersection number of the corresponding curves is given by:*

$$\mathrm{i}(\alpha, \beta) = \sum_{i=1}^{\mathrm{len}(\alpha)} \sum_{j=1}^{\mathrm{len}(\beta)} \frac{|\mathrm{cross}(\sigma^i \alpha, \sigma^j \beta)|}{1 + \mathrm{val}(\sigma^i \alpha, \sigma^j \beta)} \qquad \text{(cross-val)}$$

counting the pairs of cyclically reduced representatives of the conjugacy classes up to simultaneous conjugacy, whose endpoints are linked.

Proof. The topological intersection number $\mathrm{i}(\alpha, \beta)$ counts the number of intersection points in the fundamental domain between representatives for the conjugacy classes.

It is given by summing over pairs of representatives whose axis intersect the fundamental domain (cyclically reduced words $\sigma^i \alpha, \sigma^j \beta$) which intersect (that is with $|\mathrm{cross}(\alpha, \beta)| = 1$), the inverse of the number of fundamental regions which they fellow travel (given by $\mathrm{val}(\alpha_i, \beta_j)$). This formula holds even when α and β are equal. □

ALGORITHM 3: For primitive $\alpha, \beta \in \mathcal{F}_p \setminus \{1\}$ represented as cyclically reduced words on the symmetric set of generators $\tau^{\pm}$, we can compute the topological intersection number $\mathrm{i}(\alpha, \beta)$ of the conjugacy classes in time $O(\mathrm{len}(\alpha)\,\mathrm{len}(\beta))$.

Consequences and summaries of algorithms for pinning multiloops

We may now apply the algorithms from subsection 1.2 and 1.3 to the pinning problem.

Algorithm 1.11 (self-intersection)**.** *We have described an algorithm which has:*

> *Input: A plane multiloop* $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$ *together with a subset of regions* $P \subset R$*.*
> *Output: The self-intersection number of the multicurve* $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}_P$*.*
> *Complexity: Running time* $O(|\gamma|^2)$*.*

Corollary 1.12 (PINNING and PINMIN are P)**.** *Given a plane multiloop* γ *and a subset of regions* $P \subset R$*, we can successively:*

> PINNING*: certify that* P *is pinning in time* $O(|\gamma|^2)$*, and if so*
> PINMIN*: check if* P *is a minimal pinning set in time* $O(\mathrm{Card}(P)\,|\gamma|^2)$*, and if not*
> PINMINC*: construct a minimal pinning set* $M \subset P$ *in time* $O(\mathrm{Card}(P)\,|\gamma|^2)$*.*

Proof. Consider a multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$ and a set of regions $P \subset R$. We may compute the self-intersection number of the multicurve $\gamma \subset \mathbb{F}_P$ using the previous algorithm in time $O(|\gamma|^2)$, and compare it with the number of double-points of γ. This checks whether P is pinning, and if so confirms that $\varpi(\gamma) \leq \mathrm{Card}(P)$, proving the NP complexity of MULOOPINNUM.

Now suppose that we have checked that P is pinning. To a linear order of P we associate a minimal pinning subset computed as follows. We construct a decreasing chain of pinning sets, starting with P, by trying to remove them one by one.

We try removing the first region and check if the new set is pinning in time $O(|\gamma|^2)$: if it is we continue with that new set, otherwise we move to the next region. By doing so, we never need to try removing a pin twice: either it was already removed, or we know it cannot be removed. Thus in the worst case we need to call at most $\mathrm{Card}(P)$ times to our $O(|\gamma|^2)$ checking algorithm.

Note that this construction defines a surjective function from the set of linear orders of R to the set of minimal pinning sets of γ. □

Corollary 1.13 (MULOOPINNUM is NP)**.** *The* MULOOPINNUM *problem, defined by*

> *Instance: A filling multiloop* $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$*, and an integer* $p \in \mathbb{N}$*.*
> *Question: Does* γ *have pinning number* $\varpi(\gamma) \leq p$*?*

belongs to the class NP.

2 Geometric topology: computing immersed discs

This section focuses on loops, namely multiloops with one strand, maybe non-filling.

In the first subsection we define various notions of monogons and bigons, provide a sufficient condition for a subset of regions to be pinning which will be useful to justify the Algorithm 2.33, and prove the main Theorem 2.13.

The second subsection applies Theorem 2.13 to put the pinning sets of a loop in correspondence with the vertex covers of a hypergraph or equivalently the satisfying assignments of a positive conjunctive normal form. Hence the pinning number and an optimal pinning set of a filling loop can be found efficiently with certain SAT-solvers.

The third subsection will reduce the vertex cover problem for planar graphs to our LoopInNum problem, proving it is NP-hard (for any genus).

2.1 Taut loops, their monorbigons and mobidiscs

Fix an oriented surface $\mathbb{F}$, and recall that for a subset $P \subset \mathbb{F}$ we denote by $\mathbb{F}_P = \mathbb{F} \setminus P$. Consider a loop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}_P$ which may be filling or not; denote by $R = \mathbb{F} \setminus \gamma$ its set of regions and by $V \subset \mathbb{F}_P$ its set of double points.

Definition 2.1 (singular monorbigons)**.** *Consider a loop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}_P$.*

A singular monogon *is a non-trivial closed interval $I \subset \mathbb{S}^1$ such that $\gamma(\partial I) = \{x\}$ for some $x \in V$ and $\gamma(I)$ is null-homotopic. This singular monogon is* embedded *when γ is injective on the interior of I.*

A singular bigon *is a disjoint union of non-trivial closed intervals $I \sqcup J \subset \mathbb{S}^1$ such that $\gamma(\partial I) = \{x, y\} = \gamma(\partial J)$ for some $x, y \in V$ and $\gamma(I \sqcup J) \subset \mathbb{F}$ is null-homotopic. This singular bigon is* embedded *when γ is injective on the interior of $I \sqcup J$.*

A singular monorbigon *K refers to a singular monogon I or singular bigon $I \sqcup J$. The restriction $\gamma(K)$ is a subloop of γ and we call $\gamma(\partial K)$ its* marked points*.*

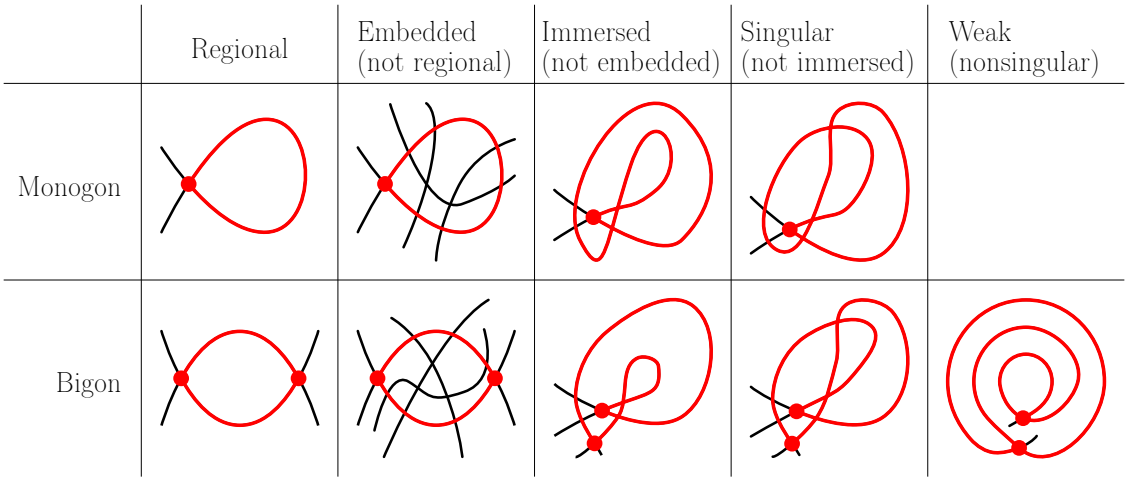


Figure 9: Examples and non-examples of different kinds of monorbigons.

Remark 2.2 (homology)**.** *For a loop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}_P$, a singular monorbigon K yields a subloop $\alpha = \gamma(K)$ which is homotopically trivial in $\mathbb{F}_P$, hence trivial in homology.*

In particular α is trivial in $H_1(\mathbb{F}_P; \mathbb{Z}/2)$ so it is two-sided. Moreover α is trivial in $H_1(\mathbb{F}_P; \mathbb{Z})$, so for every pair of points $p_0, p_1 \in P$ we have that $\mathrm{lk}(\alpha; [p_0] - [p_1]) = 0$.

Algorithm 2.3 (computing singular monorbigons)**.** *For a filling loop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}_P$, we can list all its singular monorbigons in time $O(\mathrm{Card}(V)^2 \times \mathrm{Card}(R))$.*

Proof. First list the $2\,\mathrm{Card}(V)$ intervals I and $< 2\,\mathrm{Card}(V)^2$ interval unions $I \sqcup J$ satisfying the conditions in definition 2.1, except the homotopical triviality.

Then check if the subloop $\gamma(I)$ or $\gamma(I \sqcup J)$ is homotopically trivial by computing the associated word in the fundamental group $\pi_1(\mathbb{F}_P)$ as in subsection 1.2. This applies when $P \neq \emptyset$, but when $P = \emptyset$ we perform this step for each choice of $P = \{p\} \in R$ to find all immersed monorbigons: indeed a subloop α of γ is null-homotopic in $\mathbb{F}$ if and only α if it is null-homotopic in some $\mathbb{F} \setminus \{p\}$. □

Our interest in the concept of embedded and singular monorbigons lies in the following Theorems which were proven by Hass-Scott in [HS85].

Theorem 2.4 (taut loops have no singular monorbigons)**.** *Fix a loop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}_P$. Recall that $\mathrm{si}(\gamma)$ denotes the self-intersection number of its homotopy class in $\mathbb{F}_P$.*

(2.7) If γ has > 0 double-points but $\mathrm{si}(\gamma) = 0$, then γ has an embedded monorbigon.
(4.2) If γ has $> \mathrm{si}(\gamma)$ double-points, then it has a singular monorbigon.

Definition 2.5 (Linking function)**.** *Consider a loop $\alpha\colon \mathbb{S}^1 \looparrowright \mathbb{F}_P$ with regions R, and assume that it yields a trivial homology class in $H_1(\mathbb{F}_P;\mathbb{Z}) = H_1(\mathbb{F};\mathbb{Z})$.*

For any two regions $o, \infty \in R$, the linking number $\mathrm{lk}(\alpha, [o] - [\infty]) \in \mathbb{Z}$ is defined as the signed intersection number between α and any arc β from ∞ to o.

The linking function *$\mathrm{lk}(\alpha, \cdot)\colon R \times R \to \mathbb{Z}$ defined by $(o, \infty) \mapsto \mathrm{lk}(\alpha, [o] - [\infty])$ changes sign under change of orientation of the 1-cycle α or the 0-chain $[o] - [\infty]$, in particular its absolute value and vanishing locus are well defined independently of these orientations.*

Lemma 2.6 (singular monorbigons link trivially)**.** *Consider a loop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}$ and a set of regions $P \subset R$. If P is not pinning γ, then there exists a singular monorbigon α of γ such that for every pair of regions $o, \infty \in P$ we have $\mathrm{lk}(\alpha, [o] - [\infty]) = 0$.*

Proof. If P is not pinning γ then the loop $\gamma_P\colon \mathbb{S}^1 \to \mathbb{F}_P \setminus \{\infty\}$ is not taut, so by 2.4 it has a singular monorbigon α. In particular α is trivial in the relative homology group $H_1(\mathbb{F}, P;\mathbb{Z})$, hence for all $o, \infty \in P$ we have $\mathrm{lk}(\alpha, [o] - [\infty]) = 0$. □

Remark 2.7 (converse is false)**.** *There are loops $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}$ admitting pinning sets $P \subset R$ which contradict the converse to the previous lemma: for every monorbigon α of γ and every pair of pinned regions $o, \infty \in P$, we have $\mathrm{lk}(\alpha, [o] - [\infty]) = 0$.*

One example is obtained as follows. The thrice punctured sphere $\mathbb{S}^2 \setminus \{0, 1, \infty\}$ with fundamental group $\langle \tau_0, \tau_1 \rangle$ has a unique hyperbolic metric: consider the geodesic representative γ for the commutator $[\tau_0, \tau_1]$ as depicted in Figure 10. As commutators have trivial abelianisation, for all $p, q \in \{0, 1, \infty\}$ we get $\mathrm{lk}(\gamma, [p] - [q]) = 0$.

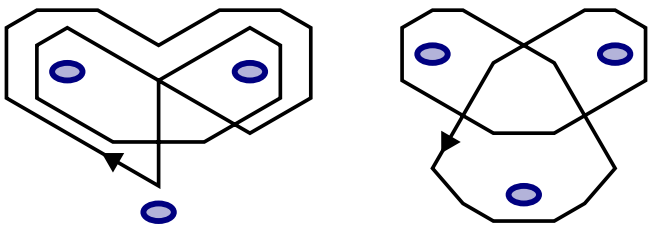


Figure 10: The trefoil in $\mathbb{S}^2 \setminus \{0, 1, \infty\}$ represents a commutator in $\pi_1(\mathbb{S}^2 \setminus \{0, 1, \infty\})$.

To characterise the pinning sets of loops in terms of the linking numbers of their singular monorbigons, we will thus restrict to the class of immersed monorbigons...

Definition 2.8 (immersed monorbigons and mobidiscs)**.** *Consider a loop* $\gamma\colon \mathbb{S}^1 \to \mathbb{F}_P$ *with a singular monorbigon* $K \subset \mathbb{S}^1$*.*

We say that K *is an* immersed *monorbigon when there exists an immersion* $\iota\colon \mathbb{D} \looparrowright \mathbb{F}_P$ *such that the restriction* $\gamma\colon K \looparrowright \mathbb{F}_P$ *factors through* $\iota\colon \partial\mathbb{D} \looparrowright \mathbb{F}_P$*.*

A mobidisc *consists of the regions of* γ *in the image* $\iota(\mathbb{D})$ *of such an immersion. We denote by* $\mathrm{MoB}(\gamma) \subset \mathcal{P}(R)$ *the set of all mobidiscs.*

Remark 2.9 (precisions)**.** *According to the definition of immersed monorbigons, the immersion* $\iota\colon \partial\mathbb{D} \looparrowright \gamma(K)$ *is surjective, and its only double-points are those of* $\gamma(K)$*.*

In particular $\iota\colon \partial\mathbb{D} \looparrowright \gamma(K)$ *cannot overlap itself (even with opposite directions), so the immersion suggested on the right of Figure 11 does not satisfy the conditions in Definition 2.8.*

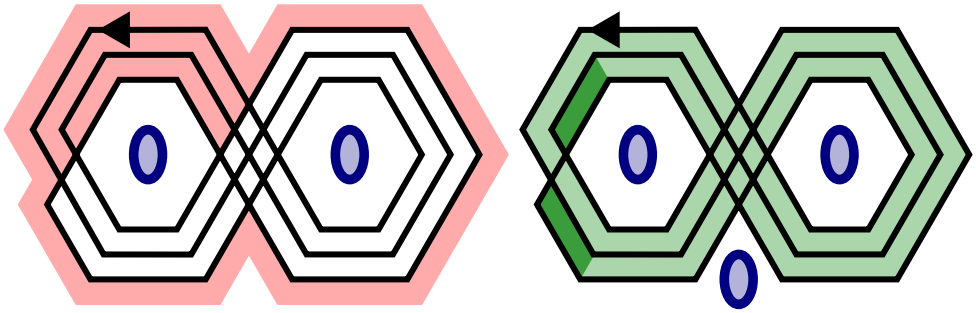


Figure 11: Left: The loop is not taut, since it has a singular bigon (which is immersed, and even embedded). Right: Placing an additional puncture at infinity makes the loop taut. Indeed, the immersed disc covering the green portions does not satisfy the conditions in Definition 2.8.

Lemma 2.10 (uniqueness of mobidiscs)**.** *For a loop* $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}_P$ *with* $\infty \in P \neq \emptyset$*, every immersed monorbigon* K *gives rise to exactly one mobidisc: it consists of the regions* $o \in R \setminus P$ *such that* $|\mathrm{lk}(\gamma(K), [o] - [\infty])| \neq 0$*.*

Proof. One may deduce this from the more general statements in [Pap96], but here is a simple explanation.

Consider a region o and an immersed disc $\iota\colon \mathbb{D} \mapsto \mathbb{F}_P$ such that the immersed monorbigon $\gamma\colon K \looparrowright \mathbb{F}_P$ factors through $\iota\colon \partial\mathbb{D} \looparrowright \mathbb{F}_P$.

If $\mathrm{lk}(\gamma(K), [o] - [\infty]) \neq 0$ then o must belong to $\mathrm{im}(\iota)$ since $\gamma(K)$ is null-homotopic inside $\mathrm{im}(\iota)$ (this holds even for a singular monorbigon bounding a singular disc).

If o belongs to $\mathrm{im}(\gamma)$ and $\mathrm{lk}(\gamma(K), [o] - [\infty]) = 0$, then there must exist two points in the preimage $\iota^{-1}(o)$ at which the Jacobian of ι has opposite determinants. Hence the Jacobian determinant of ι must vanish along any arc of $\mathbb{D}$ joining these two points, contradicting the immersion property. □

Remark 2.11 (non-uniqueness of extensions)**.** *Note however that an immersed monorbigon* K *of* γ *can extend to several non-isotopic immersions* $\iota\colon \mathbb{D} \looparrowright \mathbb{F}$ *which define the same mobidisc. A famous example given by Milnor is depicted in Figure 12.*

Given a loop in a surface, the description and computation of all its extensions to immersions of a disc had been raised by Hopf and Thom [Tho59]. It was first solved by Blank when $\mathbb{F} = \mathbb{R}^2$ *[Bla67, Poé95], improved by Shor and Van Wyck [SV92], and generalised to other surfaces in [Fri10]. See also [Pap96] for another viewpoint.*

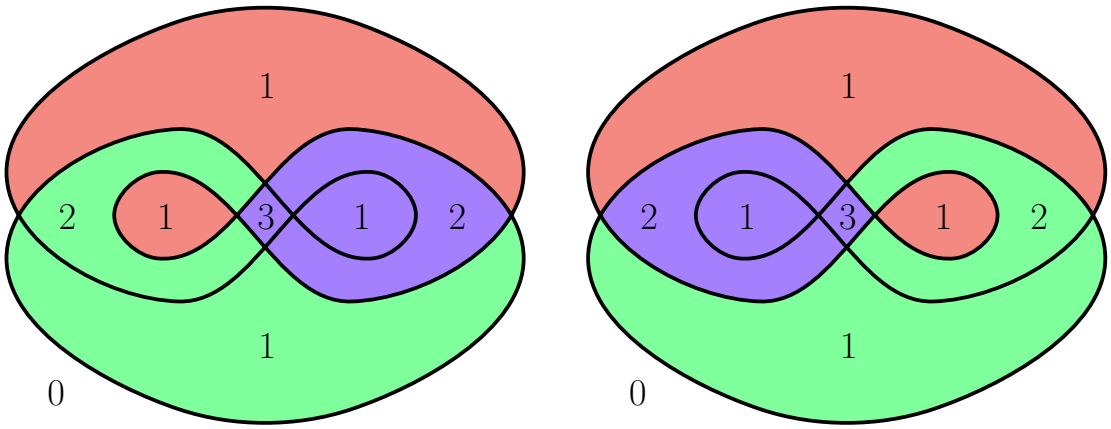


Figure 12: The "Milnor doodle" bounds an immersed disk in two different ways. When viewed from above, the violet portions are the closest whereas the red portions are the farthest. The numbers in each region correspond to the local degree of the immersion - note that they agree in both cases.

Definition 2.12 (extroverted monorbigon)**.** *For a loop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}_P$, a monorbigon K will be called* extroverted *when it extends to an* extroverted *immersion $\iota\colon \mathbb{D} \looparrowright \mathbb{F}_P$, namely such that each marked point $z \in \partial\mathbb{D}$ has a neighbourhood $U_z \subset \mathbb{D}$ satisfying $U_z \cap \iota^{-1}(\gamma(\mathbb{S}^1)) \subset \partial\mathbb{D}$, meaning that U_z avoids the preimage $\iota^{-1}(\gamma(\mathbb{S}^1))$ which lies in the interior of $\mathbb{D}$. An immersed monorbigon which is not extroverted will be called* introverted*.*

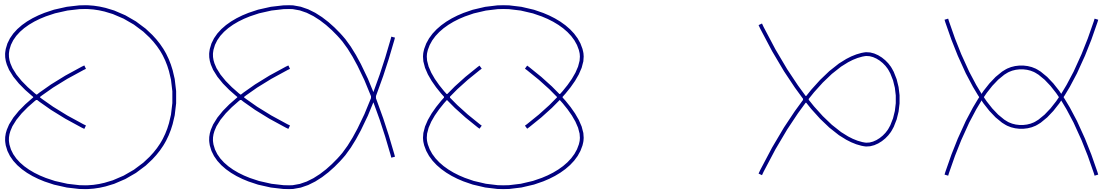


Figure 13: Left: Introverted monogon and two introverted bigons. Right: Extroverted monogon and bigon.

The main result in this subsection is an improvement of [HS85, Theorem 4.2].

Theorem 2.13 (non-taut loops have extroverted monorbigons)**.** *Consider a closed oriented surface $\mathbb{F}$ together with a set of points $P \subset \mathbb{F}$, and a loop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}_P$.*

The loop $\gamma \looparrowright \mathbb{F}_P$ is not taut if and only if it has an extroverted monorbigon.

The proof will make use of another theorem by Hass–Scott [HS94].

Theorem 2.14 (curve shortening flow)**.** *If loop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}_P$ is homotopic to a taut loop $\gamma_0\colon \mathbb{S}^1 \looparrowright \mathbb{F}_P$, then there is a sequence of isotopies and shortening-Reidemeister moves $R1, R2, R3$ (which never increase the number of double-points) leading from γ to γ_0.*

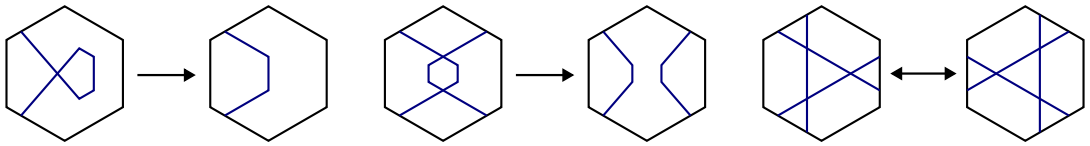


Figure 14: The shortening Reidemeister moves of type $R1$, $R2$, $R3$.

The following key lemma shows that along a curve shortening flow we can (back)track a sequence of extroverted monorbigons for the loop.

Lemma 2.15. *Consider a loop β and a shortening-Reidemeister move to a loop α. If α has an extroverted monorbigon then β has an extroverted monorbigon.*

Proof. Consider a shortening-Reidemeister move from a loop β to a loop α. If that move is of type $R1$ or $R2$ then β has an embedded monorbigon, so the implication is obviously true. Now assume that it is a move of type $R3$.

Choose an extroverted monorbigon K_a of α , together with an extrovert disc $D_a\colon \mathbb{D} \hookrightarrow \mathbb{F}_P$ such that $\alpha(\partial D_a)$ coincides with the restriction of α to K_a.

We list all possible local configurations for the $R3$ move. For each configuration, we show how to find an extroverted immersion $D_b\colon \mathbb{D} \hookrightarrow \mathbb{F}_P$ for β.

The collection of local configurations is partitioned into types and listed systematically in Figures 15 and 17 according to the following conventions. Those involving $i \in \{0, 1, 2\}$ marked points of $\alpha(\partial K_a)$ and $j \in \{0, 1, \ldots, 3\}$ local strands of $\alpha(K_a)$ will be called of type $R3(i, j)$. For each type, we run through all compatible bicolourings for the 9 segments in the triangle neighbourhood; red portions represent portions of the monorbigon whereas thin blue portions represent represent portions of its complement. Each case breaks into further subcases according to the choice of (co)orientations for the j local strands of $D_a(\partial\,\mathbb{D}) = \alpha(K_a)$ yielding the variation in local degree of the immersion between adjacent regions (we do not precise all of these subcases in the figures except for those which require some explanation). When a picture has an arrow, it means that only the source can be a configuration of α.

Note that if the move restricts to a regular homotopy of ∂D_a (that is a homotopy in the space of immersions, given either by an isotopy or a Reidemeister move of type $R2$ or $R3$), then one may continue the immersion D_a to obtain D_b. This deals with all cases involving $i = 0$ and $i = 1$ marked points in Figure 15 (we illustrate subcases for $R3(1, 2)B$ in Figure 16 for completeness), as well as the case $R3(2, 1)A$ involving $i = 2$ marked points in Figure 17. Cases $R3(2, 1)B$ and $R3(2, 1)C$ in Figure 17 are ruled out because an immersion involving the marked points must be introverted.

We are left with the moves of type $R3(2, 2)A$ and $R3(2, 2)B$ in Figure 17. In the $R3(2, 2)A$ case, the immersed disk may cobound $\alpha(K_a)$ in four distinct ways (two of which are mirror images of each); as illustrated at left in Figure 18. We rule out three cases as introverted, and in the remaining case we note that the outgoing strands labelled x_1 and x_2 must be consecutive (since we are dealing with a bigon), as must those labelled y_1 and y_2. But this is impossible since it yields a multiloop with more than one strand. Similarly, in the $R3(2, 2)B$ case, the immersed disk may cobound to $\alpha(K_a)$ in the four distinct ways illustrated at the right in Figure 18, three of which are ruled out as introverted. We are left with the remaining case $R3(2, 2)B$ which is rightmost in Figure 18: this requires a different argument which will occupy the remainder of the proof.

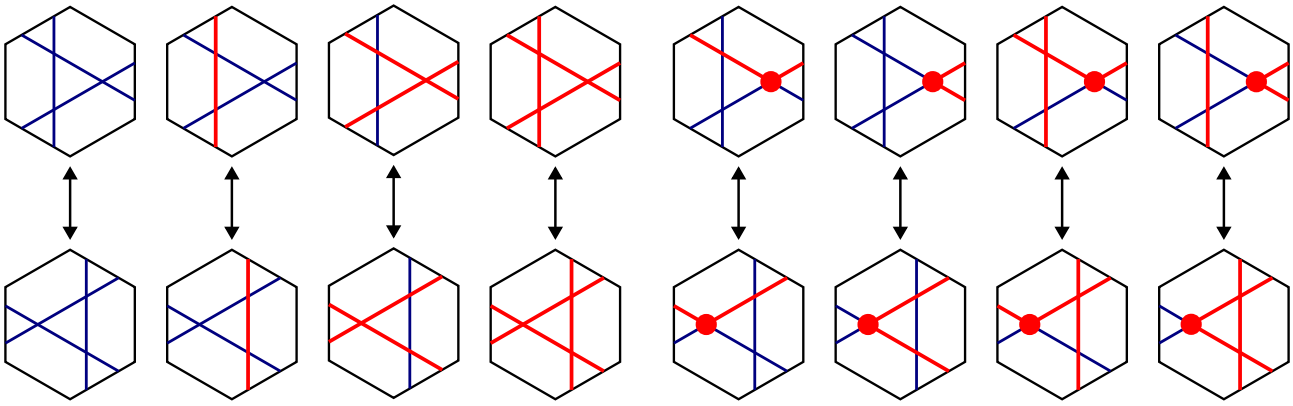


Figure 15: Configurations of type $R3(0,j)$ and $R3(1,j)$. Left to right:
$R3(0,0)$, $R3(0,1)$, $R3(0,2)$, $R3(0,3)$ and $R3(1,1)A$, $R3(1,1)B$, $R3(1,2)A$, $R3(1,2)B$.

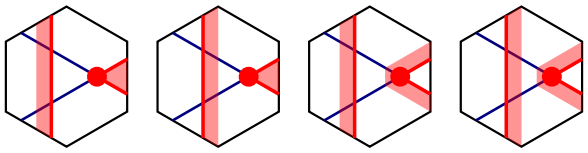


Figure 16: Possible co-orientations of immersed disks for α prior to a top-to-bottom move of type $R3(1,2)B$. The immersed monorbigons for the leftmost two are trivially preserved under the move, while those for the rightmost two are introverted.

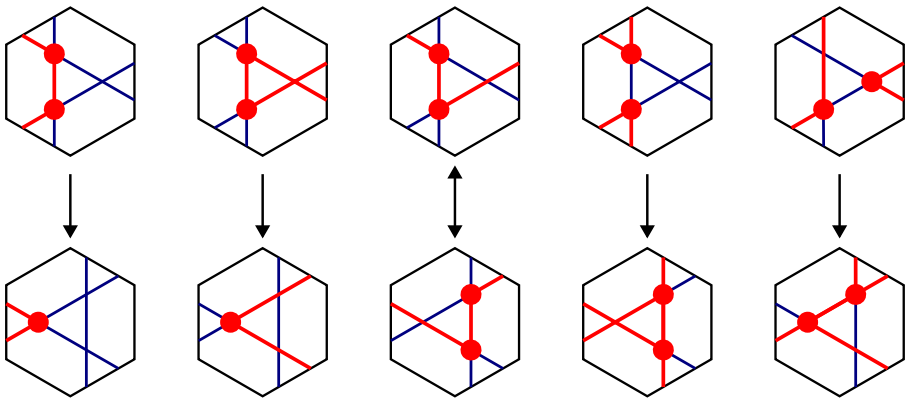


Figure 17: Configurations of type $R3(2,j)$, with arrows pointing from α to β.
Left to right: $R3(2,1)A$, $R3(2,1)B$, $R3(2,1)C$, $R3(2,2)A$, $R3(2,2)B$.

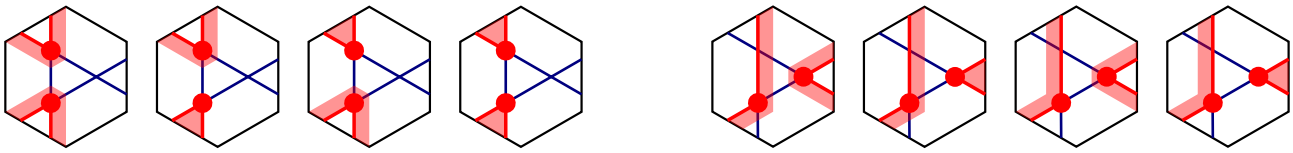


Figure 18: Left: Possible co-orientations of immersed disks for α prior to a move of type $R3(2,2)A$. The first three are introverted. In the fourth, such an immersion would imply that the vertical segment has its endpoints connected outside the local region by a strand of the loop, contradicting the fact that the loop has a single strand.
Right: Possible co-orientations of immersed disks for α prior to a move of type $R3(2,2)B$. The first three are introverted; the fourth subcase requires more work.

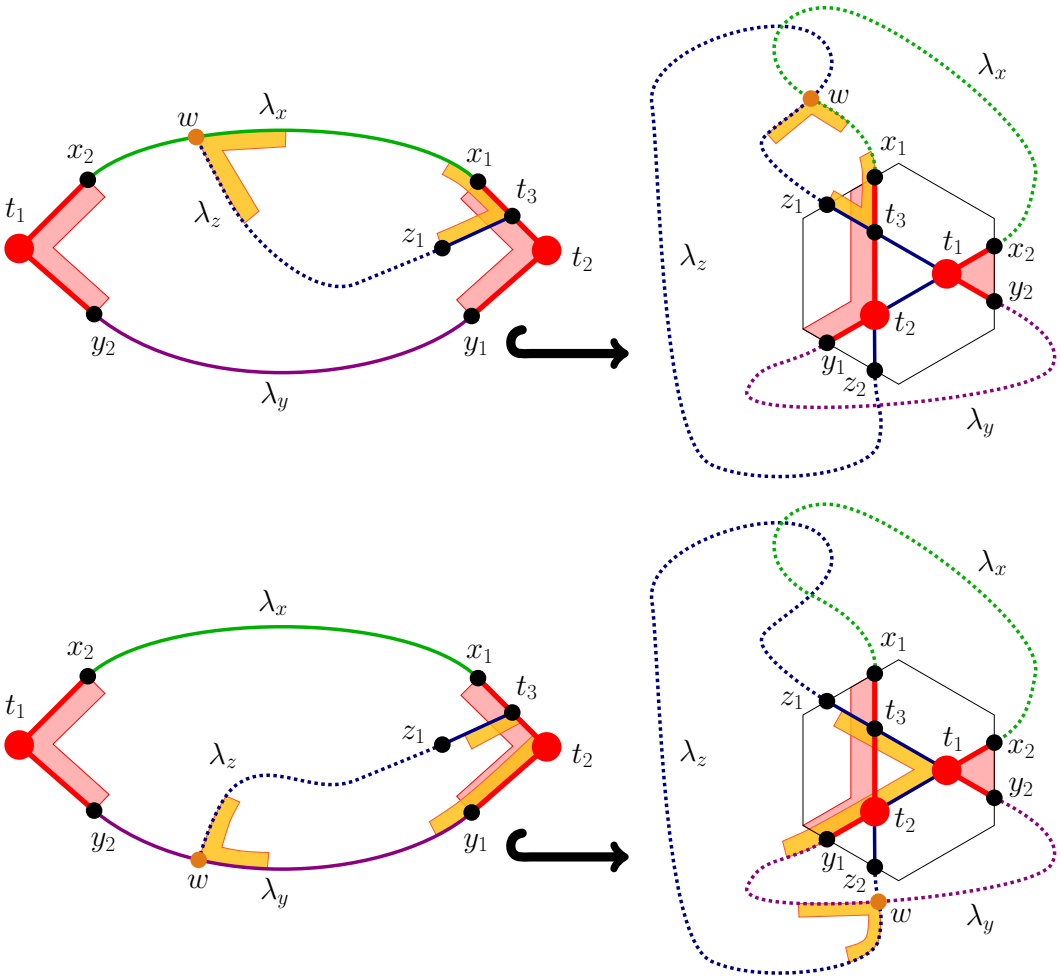


Figure 19: Reducing the remaining $R3(2,2)B$ subcase to the $R3(1,1)B$ case.

The subcase $R3(2,2)B$.

Consider an extroverted bigon $D_a\colon \mathbb{D} \to \mathbb{D} \subset \mathbb{F}$ whose boundary maps to $\alpha(K_a)$ as shown at the right in Figure 18. We will show that we can find a different extroverted bigon with the property that at most one of its marked points is among the vertices of the triangle move, thus reducing this case to the $R3(i,j)$ case for $i \in \{0,1\}$.

Note that the orientation of the immersed disc D_a implies that outside the local region, the loop γ connects the points x_1 and x_2 by a segment γ_x, the points y_1, y_2 by a segment γ_y, the points z_1, z_2 by a segment γ_z. We label the vertices of the triangle t_1, t_2, t_3, so that the immersed bigon K_a has marked points t_1 and t_2. Consider the preimage of D_a in $\mathbb{D}$ under D_a, and label points and arcs of $\mathbb{D}$ according to their images under D_a.

There is a unique lift of λ_z in $\mathbb{D}$ starting at the unique lift of t_3 on the boundary of $\mathbb{D}$: it is an immersed arc in $\mathbb{D}$ which must therefore terminate at some point $w \in \partial\,\mathbb{D}$, either on λ_x or λ_y as depicted in pictures in Figure 19.

We may assume that the lift of λ_z is embedded in $\mathbb{D}$, for otherwise we could apply [HS85][Lemma 1.1] to find an embedded monogon within $\mathbb{D}$ to which the restriction of D_a yields an extroverted monogon for α not involving the triangle move, thus reducing to the case $R3(0,0)$.

To complete the proof, we show that both configurations ($w \in \lambda_x$ and $w \in \lambda_y$) reduce to the case $R3(1,1)B$. If w lies on λ_x, then in $\mathbb{D}$ we have an embedded bigon between w and t_3 whose image under D_a is an extroverted bigon of type $R3(1,1)B$. If w lies on λ_y, then in $\mathbb{D}$ we have an embedded triangle with vertices w, t_2, t_3 whose union with the embedded triangle with vertices t_1, t_2, t_3 yields an extroverted bigon with marked points w and t_1 of type $R3(1,1)B$.

The proof is complete. □

Proof of Theorem 2.13. $\Longleftarrow$ If γ has an immersed monorbigon K, then $\gamma(K)$ is null-homotopic in a neighborhood of the union of regions of an associated mobidisc, implying that γ is homotopic to a loop with strictly less double-points.

$\Longrightarrow$ If γ is not taut then by Theorem 2.14 there is a sequence of shortening-Reidemeister moves from γ to a loop with less double-points, which begins with a sequence of $R3$ moves followed by a move of type $R2$ or $R1$. Just before performing the first move of type $R2$ or $R1$, the loop has a region of degree 2 or 1, hence an immersed bigon. The Lemma 2.15 shows that γ must also have an extroverted monorbigon. □

Corollary 2.16 (non-taut loops have immersed monorbigons)**.** *Consider a closed oriented surface $\mathbb{F}$ together with a set of points $P \subset \mathbb{F}$, and a loop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}_P$.*

The loop $\gamma \looparrowright \mathbb{F}_P$ is not taut if and only if it has an immersed monorbigon.

2.2 Reducing pin the loop to boolean formula

For this subsection, we fix a loop $\gamma\colon \mathbb{S}^1 \to \mathbb{F}_Q$ with regions R and double-points V.

Let us first reformulate Theorem 2.13 in terms of pinning sets.

Theorem 2.17 (pinning mobidiscs)**.** *For a loop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}_Q$ with regions R, a subset of regions $P \in \mathcal{P}(R)$ is pinning if and only if it intersects each mobidisc $D \in \mathrm{MoB}(\gamma)$.*

Algorithm 2.18 (computing mobidiscs)**.** *For a filling loop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}_Q$, we can construct its collection of mobidiscs $\mathrm{MoB}(\gamma) \subset \mathcal{P}(R)$ in time $O(\mathrm{Card}(V)^2\,\mathrm{Card}(R))$ when $Q \neq \emptyset$ and in time $O(\mathrm{Card}(V)^2\,\mathrm{Card}(R)^2)$ when $Q = \emptyset$.*

Proof. We first apply algorithm 2.3 to find all singular monorbigons.

For each singular monorbigon K, we check if $\gamma(K)$ bounds an immersed disc using [Fri10, Theorem 3.1.5 and Theorem 3.2.4 and Theorem 5.2.7]. This may be done in time $O(\mathrm{Card}(R))$ when the surface has a puncture, and in time $O(\mathrm{Card}(R)^2)$ when the surface is closed by trying all possible punctures as in algorithm 2.3.

Then, for each of these $O(\mathrm{Card}(V)^2)$ immersed monorbigons, we compute the corresponding mobidiscs, using winding numbers as in Lemma 2.10, again in time $O(\mathrm{Card}(R))$ if $Q \neq \emptyset$ or $O(\mathrm{Card}(R)^2)$ if $Q = \emptyset$. This yields the desired collection of $O(\mathrm{Card}(V)^2)$ mobidiscs $\mathrm{MoB}(\gamma) \subset \mathcal{P}(R)$ in the announced time. □

Let us now reformulate the LooPin problem as a satisfiability problem.

Definition 2.19 (formulae and satisfying assignments)**.** *Consider a set of variables $x_1, \dots, x_r$ which may take boolean values $\{1, 0\}$. A* formula *ϕ is an expression in the variables x involving negations $\neg$, disjunctions $\wedge$, conjunctions $\vee$ (and parentheses).*

Evaluating the variables x at $\xi \in \{1,0\}^r$ yields a value $\phi(\xi) \in \{1,0\}$ for the formula ϕ. Such an assignment ξ of the variables x is said to satisfy *ϕ when $\phi(\xi) = 1$.*

Two formulae ϕ, ϕ' on the set of variables x are called logically equivalent *when their evaluations on every assignment ξ of the variables x are equal $\phi(\xi) = \phi'(\xi)$.*

Every formula is equivalent to a conjunctive normal form. (One can also show that there is an essentially unique "minimal" conjunctive normal form.)

Definition 2.20 (conjunctive normal form)**.** *Consider boolean variables $x_1, \dots, x_r$.*

A clause *is a formula of the form $l_1 \vee \cdots \vee l_k$ where each l_j is a* literal*, namely a variable x_t or its negation $\neg x_i$. A* conjunctive normal form *(abbreviated CNF) is a formula which is conjunction $c_1 \wedge \cdots \wedge c_m$ of clauses c_j. A conjunctive normal form is called* positive *when none of the clauses feature negated variables.*

Definition 2.21 (solution ideal)**.** *Consider a positive conjunctive normal form ϕ over a set of boolean variables $x_1, \dots, x_r$. For an assignment ξ of x satisfying ϕ, we call $\{x_i \mid \xi_i = 1\} \in \mathcal{P}(x)$ a* solution *of ϕ.*

The solutions of ϕ form an ideal $\mathcal{S}(\phi) \subset \mathcal{P}(x)$: a sub-poset which is absorbing under union, and it contains the whole set of all elements $\{x_1, \dots, x_r\}$.

Remark 2.22 (hyper-graph vertex-cover)**.** *The solutions of a positive conjunctive normal form correspond to the vertex-covers of the hyper-graph whose vertices are indexed by the variables and hyper-edges correspond to the clauses.*

We will only need to consider conjunctive normal forms which are positive.

Definition 2.23 (mobidisc formula)**.** *To a loop $\gamma\colon \mathbb{S}^1 \to \mathbb{F}$ we associate its* mobidisc formula*: it is the positive conjunctive normal form on the set of boolean variables indexed by its set of regions whose clauses correspond to* $\mathrm{MoB}(\gamma) \subset \mathcal{P}(R)$*.*

Question 2.24. *Can we describe, among all positive conjunctive normal forms, the mobidisc formulae arising from filling loops, and from those in genus g?*

The next subsection will show that this class is complex even when $g = 0$: the task of finding (the cardinal of) an optimal solution is NP-complete.

We may now reformulate Theorem 2.17 as follows.

Theorem 2.25 (pinning a loop = satisfying its mobidisc formula)**.** *The pinning sets of a loop correspond to the solutions of its mobidisc formula. In particular, the pinning ideal of a loop is isomorphic to the solution ideal of its mobidisc formula.*

Hence the Algorithm 2.18 recovers the following corollary (which we already knew as a consequence of Corollary 1.12).

Corollary 2.26 (LooPinNum is NP)**.** *The LooPinNum problem, defined by*

> *Instance: A filling loop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{F}$, and an integer $p \in \mathbb{N}$.*
> *Question: Does γ have pinning number $\varpi(\gamma) \leq p$?*

belongs to the class NP.

Remark 2.27 (SAT-solver)**.** *Once the subset $\mathrm{MoB}(\Gamma) \subset \mathcal{P}(R)$ has been computed in time $\mathrm{Card}(V)^2\,\mathrm{Card}(R)$ as in the algorithm 2.18, one may send it to a SAT-solver designed to compute optimal solutions to positive conjuctive normal forms.*

In practice, the current SAT-solvers for these problems are surprisingly fast.

Example 2.28 (a mobidisc formula)**.** *The mobidisc formula of the loop 11^1_{97} in Figure 20 is equivalent, after taking the set of minimal clauses with respect to implication, to:*

$$(1 \vee 2) \wedge (2 \vee 3) \wedge (2 \vee 10) \wedge (4 \vee 6) \wedge (6 \vee 7) \wedge (6 \vee 8) \wedge$$
$$(1 \vee 4 \vee 5) \wedge (1 \vee 8 \vee 9) \wedge (3 \vee 5 \vee 7) \wedge (3 \vee 8 \vee 11) \wedge (4 \vee 10 \vee 11) \wedge (7 \vee 9 \vee 10) \wedge$$
$$(1 \vee 5 \vee 7 \vee 9) \wedge (3 \vee 4 \vee 5 \vee 11) \wedge (8 \vee 9 \vee 10 \vee 11)$$

Its minimal satisfying assignments correspond to optimal and minimal pinning sets, shown in uppercase and lowercase letters.

The mobidiscs giving rise to the clauses $(2 \vee 10)$ and $(8 \vee 9 \vee 10 \vee 11)$ are the bigons shaded in orange and blue.

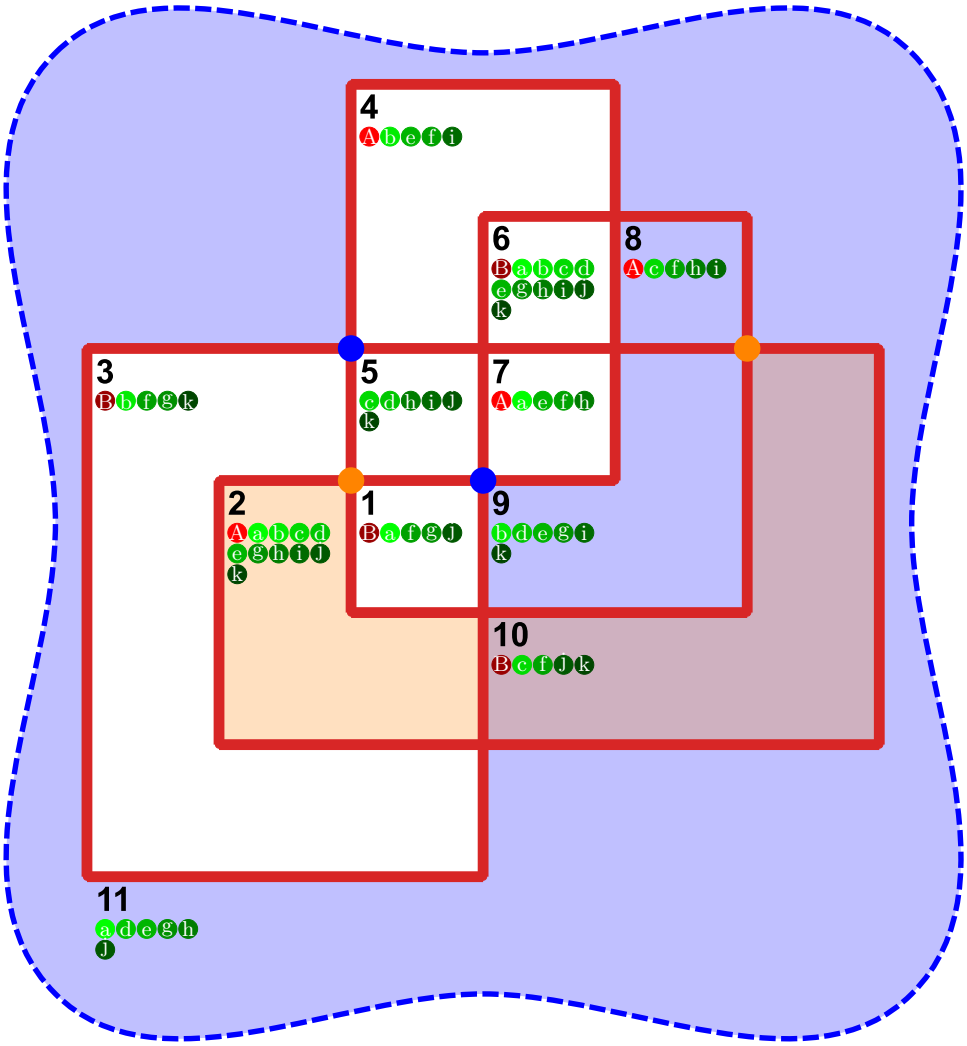


Figure 20: The loop 11^1_{97}, its minimal pinning sets, and some highlighted mobidiscs.

Remark 2.29 (just for loops)**.** *The Algorithm 2.18 relies on Theorems 2.13 and 2.17 which only hold for loops, and do not have obvious generalisations to multiloops.*

Indeed, [HS85, figure 0.1] exhibits a multiloop with 2 *strands inside an annulus which is not taut, but that does not bound any singular bigon involving both strands.*

After adding a puncture as in Figure 21, we obtain a multiloop in a thrice-punctured sphere which is not taut, but now has no singular monorbigons at all.

This is why this subsection restricts to loops, and we wonder if the MuLooPinNum *problem belongs to a stricter larger sub-complexity class than* LooPinNum*.*

Conjecture 2.30 (characterising non-taut multiloops)**.** *The proof of Lemma 2.15, especially the cases* $R3(2,2)$*, lead us to the following conjecture. It would rely on appropriately defining a* unicorn annulus *as the set of regions inside an immersed annulus with one* $\frac{3\pi}{2}$ *interior corner generalizing the embedded shape which is shaded in Figure 21.*

Given a multiloop, define a formula in conjunctive normal form whose clauses correspond to regions of its mobidiscs and *unicorn annuli: we conjecture that the solutions of the formula correspond to the pinning sets of the multiloop.*

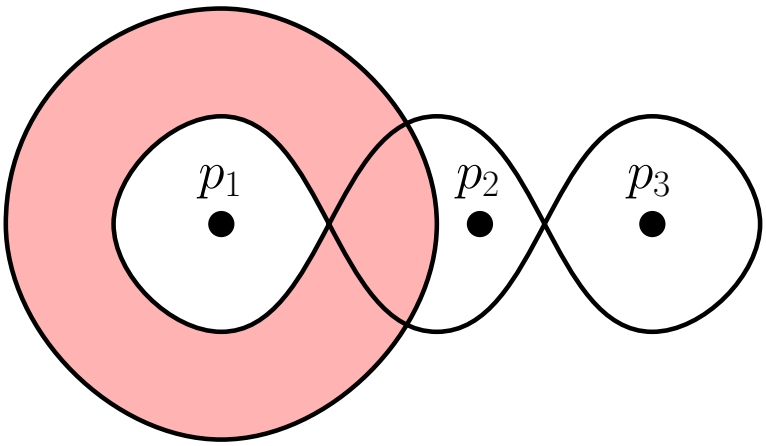


Figure 21: A multiloop $\mathbb{S}^1 \sqcup \mathbb{S}^1 \looparrowright \mathbb{S}^2 \setminus \{p_1, p_2, p_3\}$ which is not taut, but having no singular monorbigons. There is a sequence of shortening Reidemeister moves swallowing the *unicorn annulus*.

2.3 Reducing planar vertex cover to pin-the-loop

We now show that the LooPinNum problem is NP-hard by reduction from the following decision problem, which is known to be NP-hard by [GJS76][Theorem 2.7].

Definition 2.31 (Planar vertex cover)**.** *The PLANAR VERTEX COVER problem has:*

> *Instance: A planar graph $G = (V, E)$ and a positive integer $k \leq \mathrm{Card}(V)$.*
> *Question: Is there $U \subset V$ with $\mathrm{Card}(U) \leq k$ such that for all $\{v_1, v_2\} \in E$, at least one of v_1 and v_2 belongs to U?*

Remark 2.32 (Plane connected)**.** *For our purposes, the PLANAR VERTEX COVER problem has the same complexity as its restriction to connected plane graphs.*

On the one hand, the vertex covers of a graph correspond to the disjoint union of the vertex covers of each connected components.

On the other hand, there are linear time algorithms [HT74] which given a graph, decide whether it is planar and when so compute a plane embedding (encoded as a cyclic order of the edges around each vertex).

Algorithm 2.33 (PLANAR VERTEX COVER reduces to LooPinNum)**.** *The PLANAR VERTEX COVER problem admits a polynomial reduction to the LooPinNum problem.*

More precisely, there exists a polynomial time algorithm which to a planar graph $G = (V, E)$ associates a plane loop $\gamma \colon \mathbb{S}^1 \to \mathbb{D}$ with $O(\mathrm{Card}(E)^2)$ double-points, such that the k-vertex-covers of G are in one to one correspondence with the $(6\,\mathrm{Card}(E) + k)$-pinning-sets of γ.

Proof outline. Given an instance $(G = (V, E), k \in \mathbb{N})$ of PLANAR VERTEX COVER, we will construct an instance $(\gamma, 6\,\mathrm{Card}(E) + k)$ of LooPinNum, and show that they yield equivalent problems. As mentioned in Remark 2.32 we may assume that G is connected and endowed with a plane embedding in $\mathbb{B}(0, 1)$.

Following the example in Figure 24, we will associate to every edge $e \in E$ an edge-gadget g_e, and connect them all together through the boundary circle $\partial\,\mathbb{B}(0, 1)$ to form a loop γ. The technical part of the construction is to separate the gadgets adequately so as to ensure to ensure the genericity of γ and the equivalence of the two problems. In particular, we will have a correspondence (by the (deformation-retract) property) between the vertices of G and the maximal nonempty intersections of certain edge-bigons of γ. Finally we will prove the equivalence of the two problems. One direction will make extensive use of the Lemma 2.6 on linking numbers which we derived from the characterization 2.4 of taut loops.

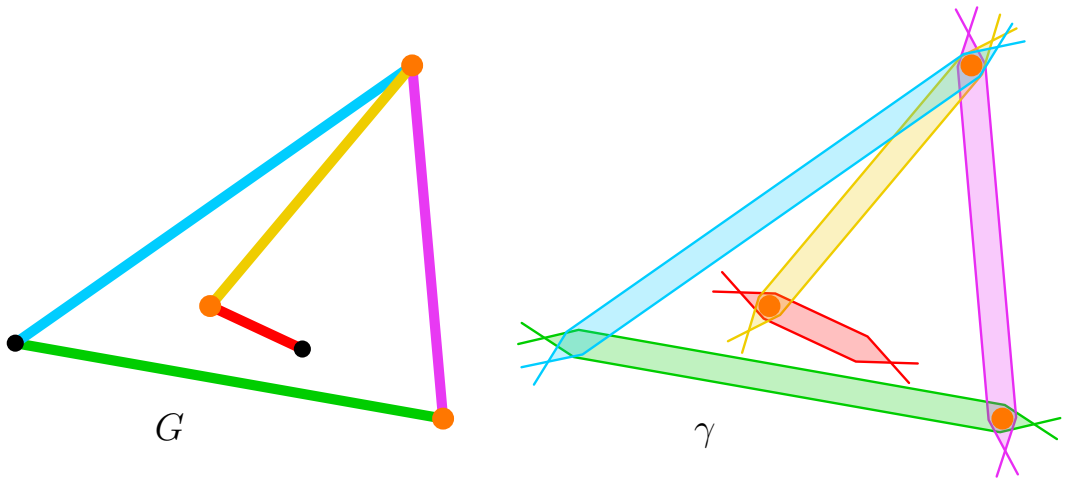


Figure 22: Rough idea of the reduction. To a plane graph G is associated a loop γ with a bigon for each edge. A vertex cover yields a choice of pins in the local pictures. The technical part of the construction is to join these bigons to form a single loop with the correct pinning lattice.

It is worth mentioning that the metric and combinatorial convolutions of the gadgets will ensure both the validity of the reduction and its justification. □

Construction of the loop. Let us now perform the construction in detail.

We work in the Euclidean plane $\mathbb{R}^2$ with distance function d. For a subset $X \subset \mathbb{R}^2$ and $\epsilon \in \mathbb{R}_+$ we denote by $\mathbb{B}(X, \epsilon)$ the ϵ-neighbourhood of X.

Definition 2.34 (Edge gadget)**.** *Fix an oriented segment $e \subset \mathbb{B}(0,1)$ and denote by L_e be the oriented line containing e.*

For a small $\epsilon > 0$, the edge gadget $g_e = g_e(\epsilon)$ is the pair of piecewise-smooth curves depicted in Figure 23. It is piecewise-linear except for portions of arcs in $\partial\,\mathbb{B}(0,1)$.

We assume that ϵ is small enough so that $\partial\,\mathbb{B}(0,1) \setminus \mathbb{B}(L_e, 3\epsilon\tan(\frac{\epsilon}{2}))$ consists of two disjoint arcs which together with $g_e(\epsilon)$ form a single piecewise-smooth loop.

The shaded bigon M_e is contained in $\overline{\mathbb{B}(e,\epsilon)}$, and we assume that ϵ is small enough so that $\overline{\mathbb{B}(e,3\epsilon)} \subset \mathbb{B}(0,1)$.

To such a gadget is associated the six forced pins $\{p_1, p_2, p_3, p_4, q_1, q_2\}$*, all of which necessarily exist for the piecewise-smooth curve to be taut after surgery.*

The double-points $\{u_i\}_{i=1}^4 \cup \{v_i\}_{i=1}^2 \cup \{w_i\}_{i=1}^6$ will be referred to in the proof justifying the validity of the reduction.

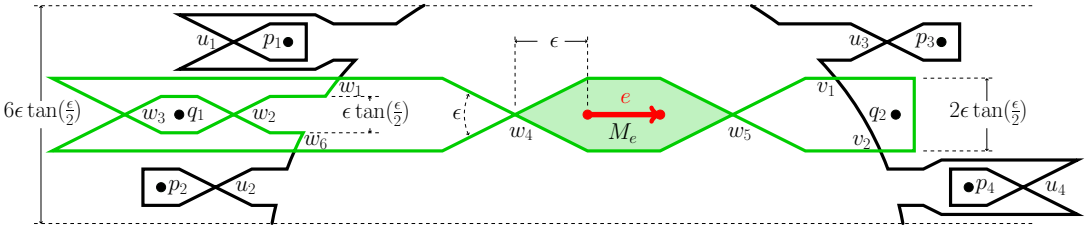


Figure 23: The edge gadget in metric details (coloured core, black bark).

Now to a connected plane graph $G \hookrightarrow \mathbb{B}(0,1)$ we associate a loop $\gamma\colon \mathbb{S}^1 \looparrowright \mathbb{R}^2$, by positioning edge-gadgets meticulously and assembling them. One may check that every step of the construction can be performed in polynomial time.

Embed graph: According to [Wag36, F48, Ste51], we may construct in polynomial time a plane embedding of $G \hookrightarrow \mathbb{B}(0,1)$ with straight edges.

Edge configuration: We may assume (after a generic perturbation of the vertices) that no edges of G are parallel. Each edge e of G spans a line L_e in $\mathbb{R}^2$, and we rescale the embedding so that all intersections $L_e \pitchfork L_{e'}$ lie inside $\mathbb{B}(0,1)$.

Fix labels: Fix a total order on the vertices V and orient each edge $e \in E$ from its smallest vertex $e^- \in V$ to its largest vertex $e^+ \in V$. Each oriented edge $e \in E$ generates an oriented affine line intersecting $\mathbb{B}(0,1)$ along an oriented chord $c_e = (c_e^-, c_e^+)$. By the edge configuration, we have $2\operatorname{Card}(E)$ distinct points $c_e^\pm \in \partial\mathbb{B}(0,1)$ on the unit circle.

Ensure deformation retracts: Choose $\epsilon > 0$ such that $\mathbb{B}(G,\epsilon) \subset \mathbb{B}(0,1)$, and $\forall e_1, e_2 \in E$ with $e_1 \neq e_2$ we have $\mathbb{B}(c_{e_1}, 3\epsilon\tan(\frac{\epsilon}{2})) \cap \mathbb{B}(c_{e_2}, 3\epsilon\tan(\frac{\epsilon}{2})) \subset \mathbb{B}(0,1)$ (possible by the edge configuration), and for all $\{e_i\}_{i=1}^k \subset E$:

$$\bigcap_{i=1}^k \overline{\mathbb{B}(e_i,\epsilon)} \quad \text{retracts by deformation onto} \quad \bigcap_{i=1}^k e_i.$$

Embed gadgets: For each $e \in E$, embed an edge gadget $g_e(\epsilon)$. By edge configuration and deformation retracts, if $e \neq e'$ then $g_e \cap g_{e'}$ consists of four points inside $\mathbb{B}(0,1)$. For each g_e, remove the arc $\partial\mathbb{B}(0,1) \cap \mathbb{B}(c_e, 3\epsilon\tan(\frac{\epsilon}{2}))$ between its end-points. Fix an orientation of the resulting closed loop and call it γ.

Ensure generic: We now argue that γ may be assumed to be in a generic position.

Transverse intersections: Since no two edges are parallel, and slopes of segments of $g_e \cap \mathbb{B}(0,1)$ are within angle $\epsilon/2$ of the slope of e, we may repeat the construction above after choosing ϵ small enough so that all intersection points between distinct gadgets are transverse double points.

Double-points: What remains are possible isolated multiple points. We compute all intersection points between all gadgets in polynomial time, as well as the minimum distance ϵ' between them, and resolve multiple points locally (inside balls of radius $\epsilon'/2$, say) into sets of transverse double-points so as to preserve the key intersection property below.

The polynomial construction of γ is finished. See Figure 24.

Note that for all $e \in E$ we have the bigon M_e satisfying $e \subset M_e \subset \overline{\mathbb{B}(e,\epsilon)}$ whence by deformation-retracts, for all $\{e_i\}_{i=1}^k \subset E$:

$$\bigcap_{i=1}^k M_{e_i} \quad \text{retracts by deformation onto} \quad \bigcap_{i=1}^k e_i. \qquad \text{(deformation-retract)}$$

Both intersections are nonempty if and only if all the e_i share a common vertex. □

Proof of the equivalence. We now show that the instances $(G = (V,E), k)$ of PLANAR VERTEX COVER and $(\gamma, 6\operatorname{Card}(E) + k)$ of LOOPINNUM yield equivalent problems.

From pinning set to vertex cover. Suppose that γ has a pinning set of cardinal $6\operatorname{Card}(E)+k$. There must be at least $6\operatorname{Card}(E)$ pins (6 per gadget) pinning the embedded monorbigon regions outside $\mathbb{B}(0,1)$, and at most k pins remain. There must also be a pin inside every M_e. We construct a map from the remaining $\leq k$ pins to vertices of G as follows: if the pin is in no M_e, then choose any vertex; if the pin is only in one M_e, then choose either vertex associated to that bigon; if it is in more than one M_e, these intersect in a nonempty region containing a unique vertex of G by the (deformation-retract) property which is the chosen vertex. This yields a vertex cover of the graph using at most k vertices.

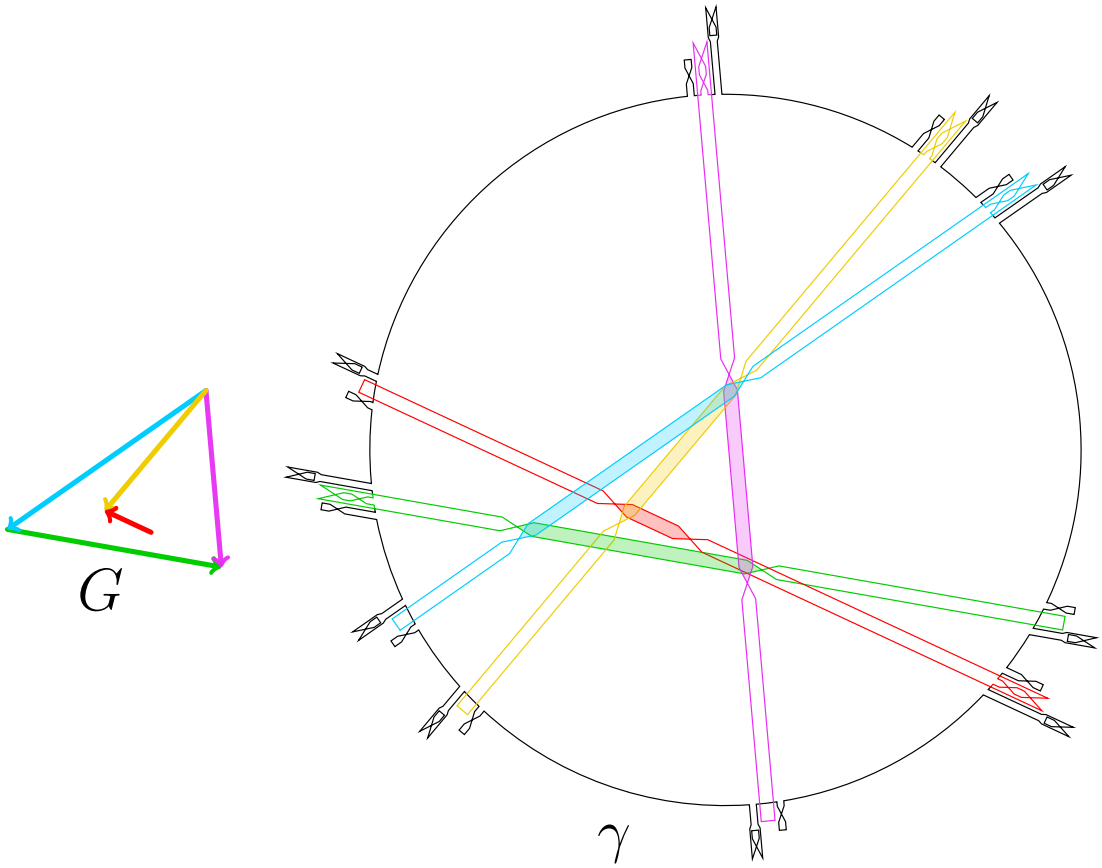


Figure 24: Assembling edge gadgets around the graph.

From vertex cover to pinning set. Now G has a vertex cover of cardinal k. Consider a set of $6\,\mathrm{Card}(E)+k$ pins P, and place $6\,\mathrm{Card}(E)$ pins in the embedded monogons and bigons outside $\mathbb{B}(0,1)$ (that is 6 per gadget), so that k pins remain. For each vertex of the cover, list its incident edges $e_1,\dots,e_m$ and place a pin in the region $\cap_{i=1}^m M_{e_i}$, which is homotopic to a disc by the (deformation-retract) property. We claim that γ is pinned (note that we did not place a pin at infinity).

By contradiction, assume γ is not pinned: it is not taut in the complement of the pins so by Theorem 2.4 it has a singular monogon (with marked point x, say) or bigon (with marked points x, y, say). Let α be the restriction of γ to this monorbigon.

We first argue that α are contained in the coloured cores of the gadgets or the strands. By Lemma 2.6, for all pins p, q we have $\mathrm{lk}(\alpha, [p]-[q]) = 0$. Thus α cannot pass through a portion of $\mathbb{S}^1$ between two distinct gadgets since otherwise it would link with the consecutive monogons. Moreover α cannot have a marked point among the double-points u_i of a gadget (as in Figure 23), since otherwise the neighbouring pin p_i together with at least one pin p'_j in another monogon region of this or an adjacent gadget would yield $\mathrm{lk}(\alpha, [p_i]-[p'_j]) \in \{\pm 1, \pm 2\}$ (this holds regardless of whether α is a monogon or a bigon).

For the rest of this explanation, we now assign colours to portions of γ. Each gadget gets assigned its own colour which appears only on its core while its bark remains black, and the arcs of $\mathbb{S}^1$ between gadgets also remain black (as illustrated in Figure 23). If α is a singular monogon, label the colours of its outgoing edges as c_1 and c_2. If α is a bigon between marked points x and y, label its colours $c_1^x, c_2^x, c_1^y, c_2^y$ where c_1^x joins to c_1^y and c_2^x joins to c_2^y. See Figure 25.

Note that if α is a monogon with marked point x, then $c_1 = c_2$ and are not black, namely x is not a point of the form w_1, w_6 or u_1, u_2. This follows from the argument above, since x does not border a gadget monogon, and α does not contain a portion of $\mathbb{S}^1$ connecting two distinct

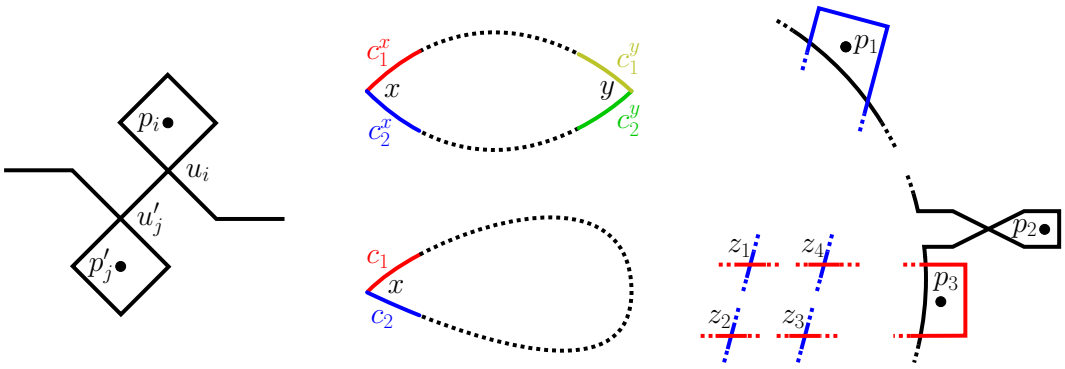


Figure 25: Left: Arguing that α must avoid the portion of γ running between gadgets. Middle: Singular monorbigon α with marked points and colours of initial segments. Right: Ruling out the possibility that α is a "bigon between gadgets".

gadgets.

Similarly, if α is a bigon, then $c_1^x = c_1^y$, $c_2^x = c_2^y$, and neither c_1^x nor c_2^y is black. If $c_1^x \neq c_2^x$, then α is a "bigon between gadgets" and the marked points are among the vertices labelled $\{z_1, z_2, z_3, z_4\}$ in the right side of Figure 25. However, since the coloured portions of distinct gadgets intersects exactly 4 times, the reader may verify that the few possible cases may be ruled out considering the pins p_1, p_2, and p_3 and applying the linking Lemma 2.6.

We have reduced to the case where α is contained within a single gadget, monochromatic, and its marked point(s) of α are among $\{w_1, w_2, w_3, w_4, w_5, w_6\}$. As the reader may check, Lemma 2.6 rules all out remaining cases and the argument is complete.

This completes the proof of the reduction 2.33. □

Corollary 2.35 (LooPinNum is NP-hard)**.** *The LooPinNum problem is NP-hard, even in restriction to loops in the plane.*

Proof. The PLANAR VERTEX COVER problem is NP-complete [GJS74]. The Algorithm 2.33 returns a loop whose size is proportional to the size of its input planar graph. Hence the LooPinNum problem is NP-hard. □

Remark 2.36 (vertex degrees ≤ 3)**.** *The PLANAR VERTEX COVER problem remains NP-complete even for graphs with* max*-degree* 3 *by [GJ77][Lemma 1], but such an assumption would not simplify our construction.*

In fact, our construction could be adapted without significant modification to work for any graph (planar or not), but the proof would be more involved.

3 The pinning ideal of a multiloop

In this last section, we first recall the definition of the pinning ideals and semi-lattices of a multiloop, and ask a few questions relating their combinatorial structure to the topology of multiloops. Then we present counter-examples to certain naïve conjectures about pinning ideals or semi-lattices discovered by systematic exploration. Finally we propose a few heuristics for approximation algorithms.

3.1 Structure of the pinning ideal

Definition 3.1 (pinning ideal)**.** *Consider a multiloop* $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$ *with regions* R*.*

The poset of regions $(\mathcal{P}(R), \subset)$ *forms a semi-lattice with respect to* $\cup$*.*

The pinning sets form the pinning ideal $\mathcal{PI} \subset \mathcal{P}(R)$*: a sub-poset which is absorbing under union.*

The unions of minimal pinning sets form the pinning semi-lattice $\mathcal{PSL} \subset \mathcal{PI}$*.*

Question 3.2 (structure of pinning ideals)**.** *A general question is: what can we say about the join posets which arise as pinning ideals of filling loops or multiloops?*

For instance, for which $s, n, u, v \in \mathbb{N}$ *can we find a multiloop with* $s \in \mathbb{N}$ *strands and* $d \in \mathbb{N}$ *double-points, having* u *optimal pinning sets and* v *minimal pinning sets?*

Remark 3.3 (computation)**.** *By Corollary 2.35, pinning ideals are hard to compute in general, even for loops in the sphere.*

By Theorem 2.25, the pinning ideal of a loop γ *with regions* R *is isomorphic to the solution ideal of its mobidisc formula* $\mathrm{MoB}(\gamma) \subset \mathcal{P}(R)$*.*

Question 3.4 (number of strands)**.** *Is there a topological notion generalising mobidiscs to multiloops, which would yield a basis for the pinning ideal of a multiloop?*

Can we compute or estimate lower bounds for the number of strands of a multiloop from its pinning ideal?

Question 3.5 (topological moves)**.** *How do pinning ideals of multiloops behave under Reidemeister moves, flypes, and crossing resolutions? We will see in Figures 26 and 27 that the pinning number can change under R3-moves and flypes on indecomposible loops in the sphere.*

How do pinning ideals of loops behave under the operations of spheric-sums and toric-sums introduced in [Sim23]?

Question 3.6 (multiloops with the same pin-intersection data)**.** *We have many examples of loops in the sphere with the same pinning ideal.*

We may associate to a multiloop γ *its complete* pin-intersection *data, consisting of its poset of regions* $\mathcal{P}(R)$ *decorated by the cardinal* $P \mapsto \mathrm{Card}(P)$ *and self-intersection* $P \mapsto \mathrm{si}(\gamma, P)$ *functionals. Can we find multiloops with that same data?*

3.2 Variance under Reidemeister triangle moves and flypes

One may hope to interpret certain link invariants in terms of certain numerical invariants of multiloops related to the pinning problem (such as the pinning number, the number of optimal or minimal pinning sets, etc.).

For this one must first choose a map from multiloops to links, and show that if two multiloops yield the same link, then they have equal pinning quantities. The next paragraphs records the failure of two such approaches.

Lifting multiloops to legendrian links: pinning varies under $R3$-moves

A multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{F}$ lifts to a unoriented Legendrian link in the projective tangent bundle $\overline{\gamma}\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \overline{\mathbb{F}}$.

Two such links are isotopic if and only if their projections differ by sequences of certain combinatorial moves, including Reidemeister moves $R3$.

If a quantity related to the pinning problem for multiloops were invariant under the Reidemeister move $R3$, then one may hope to relate it to an isotopy invariant of that Legendrian link. Alas, the pinning number is not, as one can see in Figure 26.

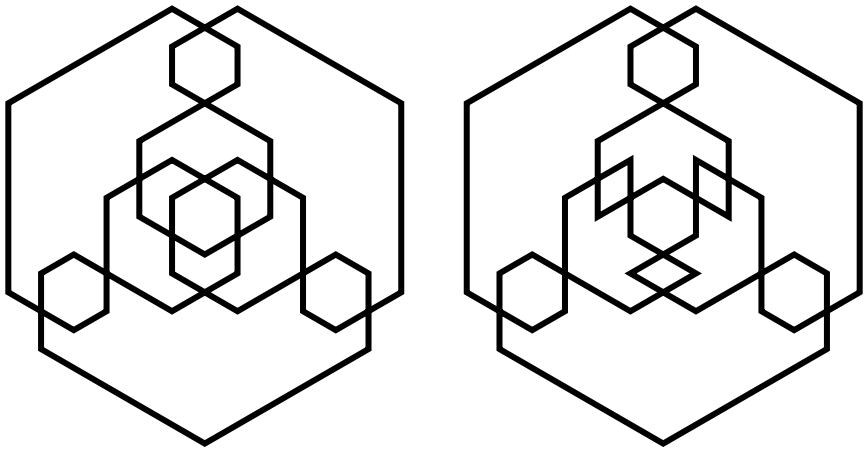


Figure 26: The pinning number is not invariant under $R3$ moves: [SS24a, link].

Lifting multiloops to alternating links: pinning varies under flypes

A multiloop $\gamma\colon \sqcup_1^s \mathbb{S}^1 \looparrowright \mathbb{R}^2$ yields two alternating diagrams for a link $\hat{\gamma}\colon \sqcup_1^s \mathbb{S}^1 \hookrightarrow \mathbb{R}^2 \times \mathbb{R}$ which are related by a mirror image reflection.

An oriented link may have several alternating diagrams, by the Tait flyping conjecture proved by Thistlethwaite–Menasco [MT91, MT93], the indecomposible alternating diagrams of a prime link are related by sequences of flypes.

Hence if a pinning-quantity of indecomposible multiloops in the sphere is invariant under flypes moves, then one may hope to relate it to an isotopy invariant of prime alternating links. Alas, the pinning number is not, as one can see in Figure 27.

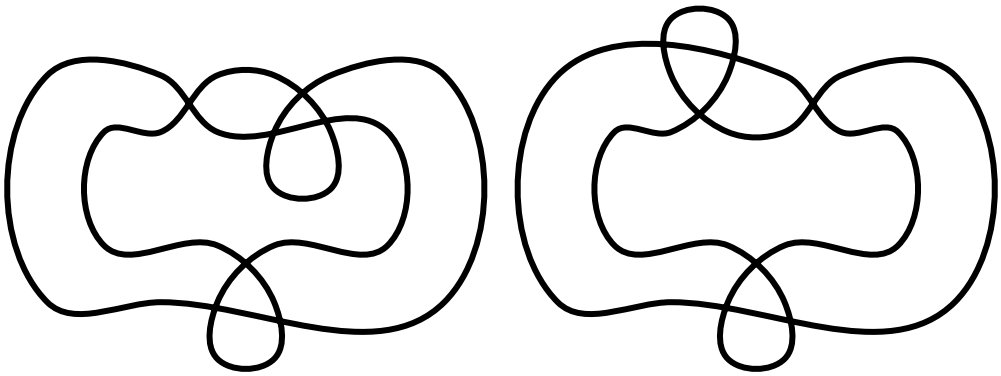


Figure 27: The pinning number may change under flypes & mutations: [SS24a, link].

3.3 Databases of pinning ideals and related quantities

Using Algorithm 0.4 and some functionalities from `plantri` [BM07], we computed the pinning ideals and semi-lattices of all irreducible indecomposible spherical multiloops with at most 12

regions (unoriented and up to reflection), and the statistics of certain numerical parameters. The number of such loops follows [OEI24, A264759] and the number of multiloops follows [OEI24, A113201]. The results are available in the LooPindex [SS24b]. Let us mention a few observations.

Remark 3.7 (region degrees)**.** *On average, there is a strong correlation between the region's degrees and probability of being pinned. It is false however that the average degree increases from optimal pinning sets to minimal pinning sets. See Figure 28.*

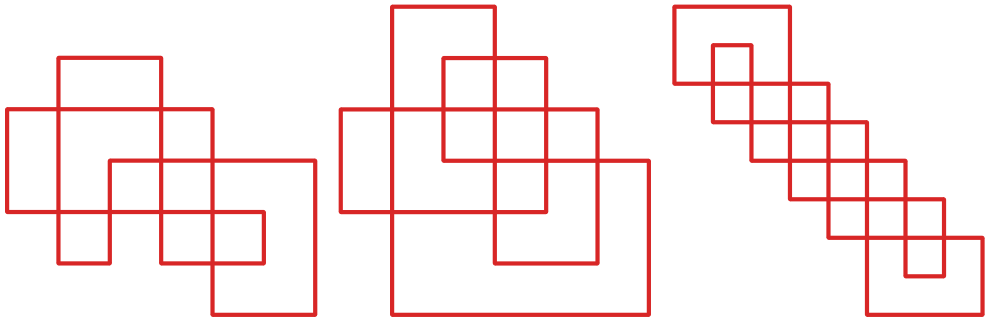


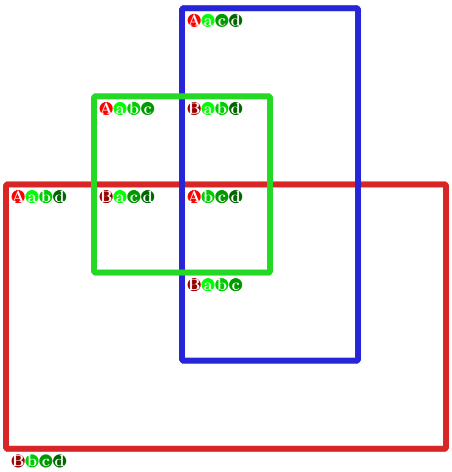
Figure 28: Left and center: Loops with the the property that the average degree of every optimal pinning set is larger than the average degree of every minimal, suboptimal pinning set.
Right: A loop with the property that the sum of degrees of some optimal pinning set is greater than the sum of degrees of some minimal suboptimal pinning set.
See [SS24a, link].

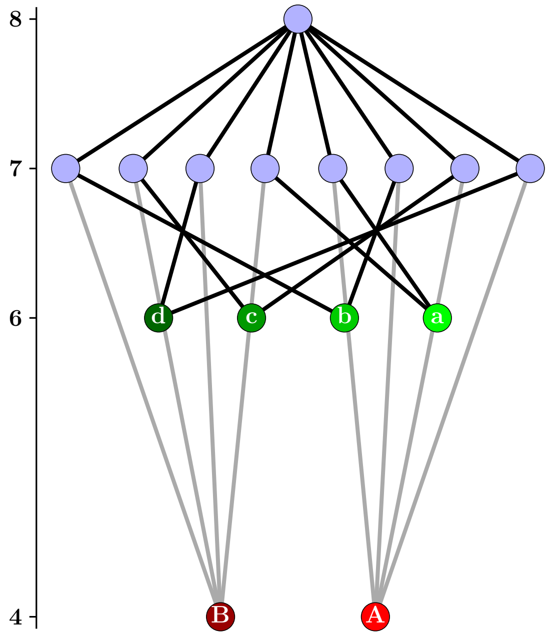
One may turn the previous remark into a heuristic leading to efficient algorithms computing almost optimal pinning sets of most multiloops.

Remark 3.8 (heuristics)**.** *The strategy consisting in solving the boolean formula whose clauses correspond to the regions bounded by embedded monorbigons often yields an almost pinning set.*

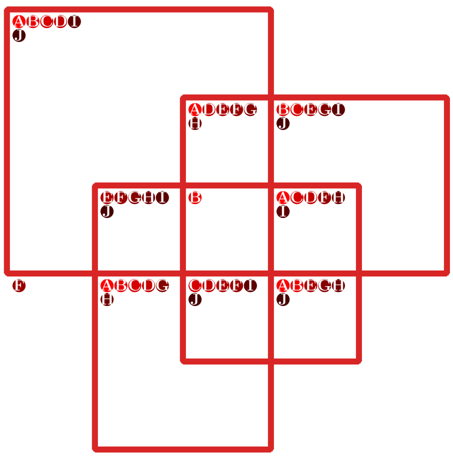
One may construct arbitrarily complicated loops with fixed pinning number $\varpi \geq 3$ by taking long words in $\mathcal{F}_{\varpi-1}$ and corresponding geodesics in a ϖ-punctured sphere. However for long geodesics, the regions to be pinned appear obvious...

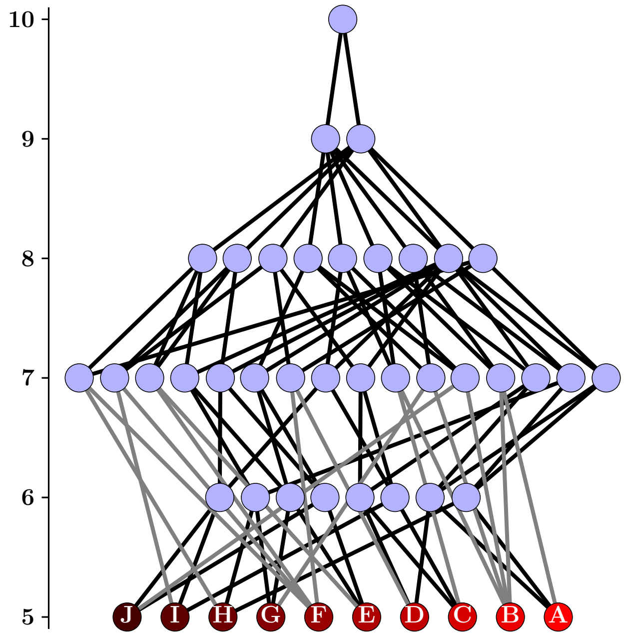
Example 3.9 (smallest multiloops with deg ≥ 3)**.** *Using lemma 1.4 and an exhaustive computation we found all multiloops in the sphere with at most* 12 *regions, all of which have degree ≥ 3 (see Figure 4 and the front page of the LooPindex). Their pinning data is illustrated in the figures that follow. Optimal pinning sets are labelled with capital letters and shades of red, and the other minimal pinning sets are labelled with lowercase letters and shades of green. For better visibility, we do not plot the entire pinning ideal but the pinning semi-lattice, together with the set of all regions. The heights of vertices in the poset (and the labels therein) correspond to their cardinals. A lighter edge emphasizes a greater difference between its endpoint's cardinals.*



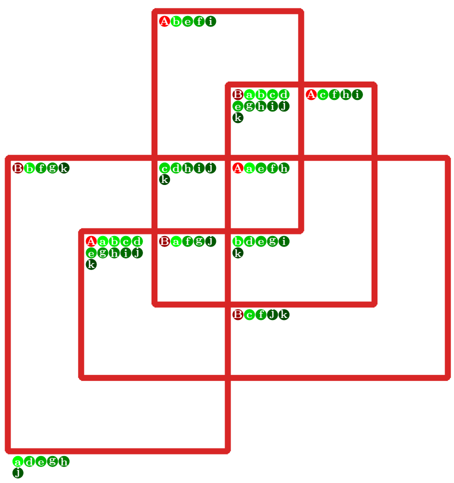


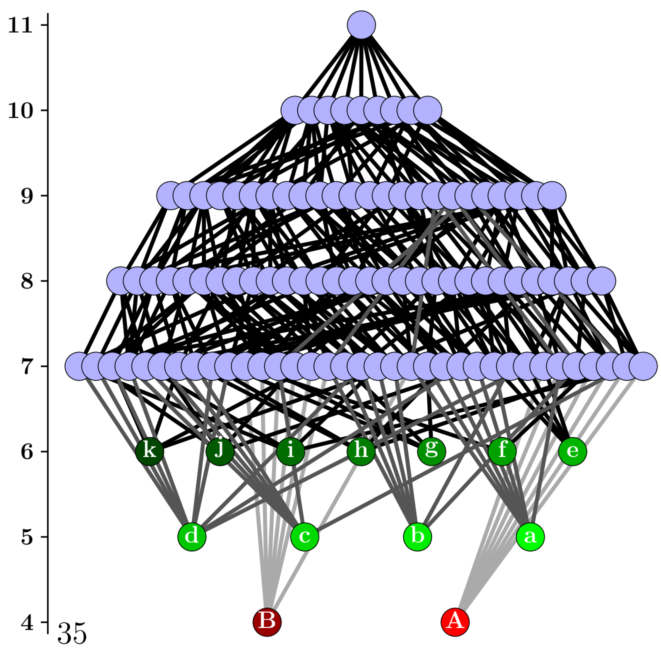
8
7
6
4
d
c
b
a
B
A



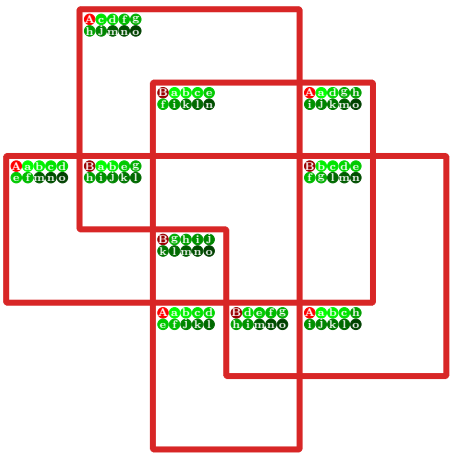


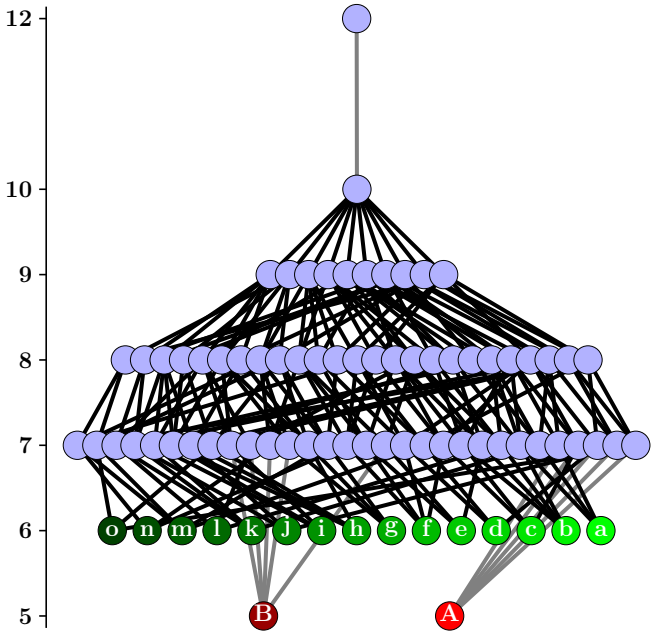
10
9
8
7
6
5
J
I
H
G
F
E
D
C
B
A

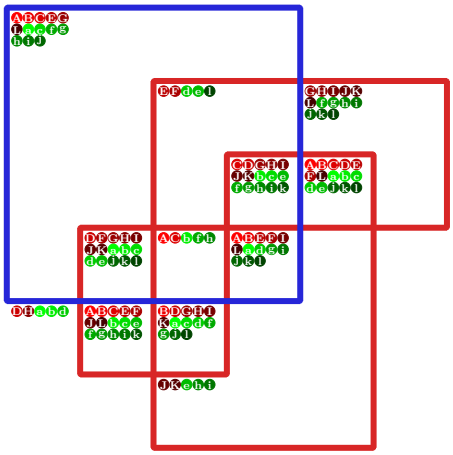


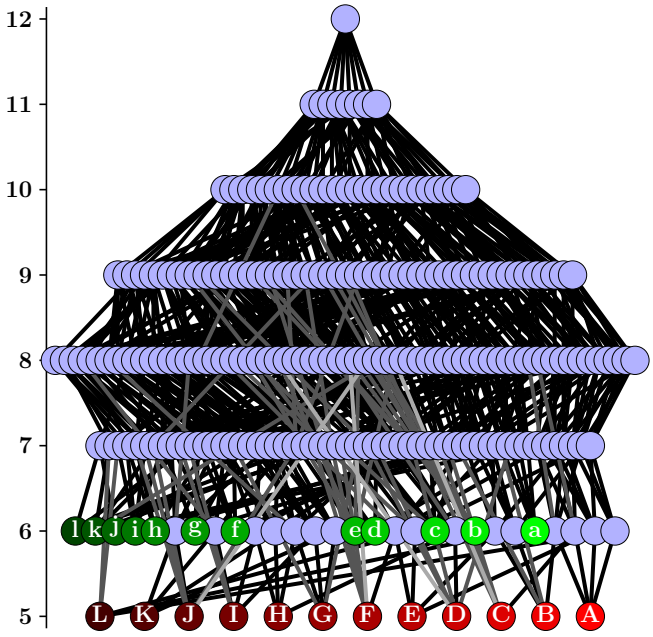


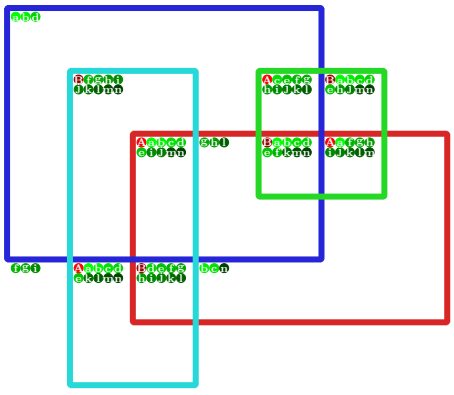
11
10
9
8
7
6
5
4
k
j
i
h
g
f
e
d
c
b
a
B
A

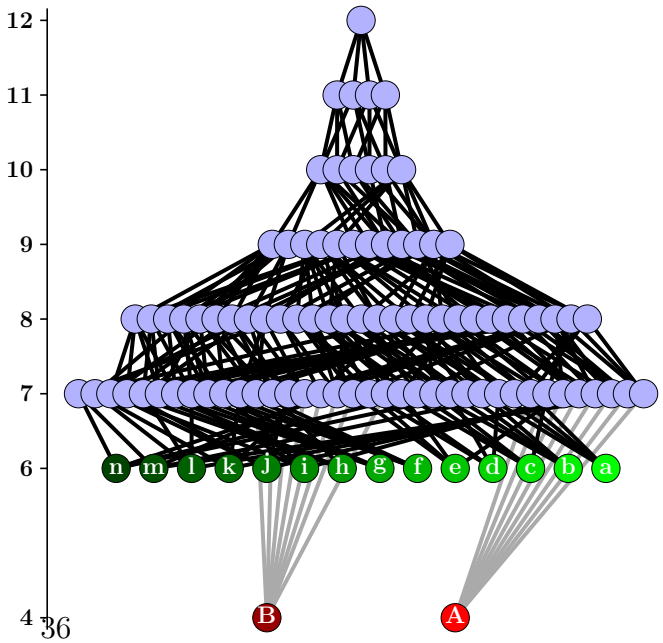



12
10
9
8
7
6
5
o n m l k j i h g f e d c b a
B
A




12
11
10
9
8
7
6
5
l k j i h g f e d c b a
L K J I H G F E D C B A




12
11
10
9
8
7
6
4
n m l k j i h g f e d c b a
B
A

Declarations

Ethical approval: This declaration is not applicable.

Funding: This declaration is not applicable.

Availability of data and materials: The experimental results [SS24a] cited throughout this work (including implementations of our algorithms, catalogs of multiloops and their pinning ideals, and counterexamples to various naïve conjectures) are available at our Github repository.

References

[AHT22] Ian Agol, Joel Hass, and William Thurston. 3-manifold knot genus is NP-complete. In *Collected works of William P. Thurston with commentary. Vol. II. 3-manifolds, complexity and geometric group theory*, pages 461–466. Amer. Math. Soc., Providence, RI, [2022] ©2022. Reprint of [2121524]. 7

[BHW22] Jonathan Bowden, Sebastian Wolfgang Hensel, and Richard Webb. Quasi-morphisms on surface diffeomorphism groups. *J. Amer. Math. Soc.*, 35(1):211–231, 2022. 7

[Bla67] S. J. Blank. Extending immersions of the circle. Ph.D. dissertation, 1967. Brandeis University, Waltham, MA. 4, 19

[BM07] G. Brinkmann and B.D. McKay. Fast generation of planar graphs. *MATCH Commun. Math. Comput. Chem.*, 58(2):323–357, 2007. see https://users.cecs.anu.edu.au/~bdm/plantri/. 5, 33

[BS84] Joan S. Birman and Caroline Series. An algorithm for simple curves on surfaces. *J. London Math. Soc. (2)*, 29(2):331–342, 1984. 13

[BS87] Joan S. Birman and Caroline Series. Dehn's algorithm revisited, with applications to simple curves on surfaces. In *Combinatorial group theory and topology (Alta, Utah, 1984)*, volume 111 of *Ann. of Math. Stud.*, pages 451–478. Princeton Univ. Press, Princeton, NJ, 1987. 13

[CDGW] Marc Culler, Nathan M. Dunfield, Matthias Goerner, and Jeffrey R. Weeks. SnapPy, a computer program for studying the geometry and topology of 3-manifolds. Available at `http://snappy.computop.org` (08/04/2024). 1

[CL87] Marshall Cohen and Martin Lustig. Paths of geodesics and geometric intersection numbers. I. In *Combinatorial group theory and topology (Alta, Utah, 1984)*, volume 111 of *Ann. of Math. Stud.*, pages 479–500. Princeton Univ. Press, Princeton, NJ, 1987. 13

[CZ16] Robert Coquereaux and Jean-Bernard Zuber. Maps, immersions and permutations. *Journal of Knot Theory and Its Ramifications*, 25(08):1650047, 2016. 9

[DL19] Vincent Despré and Francis Lazarus. Computing the geometric intersection number of curves. *J. ACM*, 66(6):Art. 45, 49, 2019. 6

[dMRST21] Arnaud de Mesmay, Yo'av Rieck, Eric Sedgwick, and Martin Tancer. The unbearable hardness of unknotting. *Adv. Math.*, 381:Paper No. 107648, 36, 2021. 7

[dMSS20] Arnaud de Mesmay, Marcus Schaefer, and Eric Sedgwick. Link crossing number is NP-hard. *J. Knot Theory Ramifications*, 29(6):2050043, 15, 2020. 7

[Fá48] István Fáry. On straight line representation of planar graphs. *Acta Univ. Szeged. Sect. Sci. Math.*, 11:229–233, 1948. 28

[FM12] Benson Farb and Dan Margalit. *A primer on mapping class groups*, volume 49 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 2012. 6

[Fri10] Dennis Frisch. Classification of immersions which are bounded by curves in surfaces. Ph.D. dissertation, 2010. 4, 19, 24

[GJ77] M. R. Garey and D. S. Johnson. The rectilinear Steiner tree problem is NP-complete. *SIAM J. Appl. Math.*, 32(4):826–834, 1977. 31

[GJS74] M. R. Garey, D. S. Johnson, and L. Stockmeyer. Some simplified NP-complete problems. In *Sixth Annual ACM Symposium on Theory of Computing (Seattle, Wash., 1974)*, pages 47–63. Association for Computing Machinery, New York, 1974. 31

[GJS76] M. R. Garey, D. S. Johnson, and L. Stockmeyer. Some simplified NP-complete graph problems. *Theoret. Comput. Sci.*, 1(3):237–267, 1976. 27

[HLP99] Joel Hass, Jeffrey C. Lagarias, and Nicholas Pippenger. The computational complexity of knot and link problems. *J. ACM*, 46(2):185–211, 1999. 7

[HS85] Joel Hass and Peter Scott. Intersections of curves on surfaces. *Israel J. Math.*, 51(1-2):90–120, 1985. 4, 18, 20, 23, 26

[HS94] Joel Hass and Peter Scott. Shortening curves on surfaces. *Topology*, 33(1):25–43, 1994. 6, 20

[HS99] Joel Hass and Peter Scott. Configurations of curves and geodesics on surfaces. In *Proceedings of the Kirbyfest (Berkeley, CA, 1998)*, volume 2 of *Geom. Topol. Monogr.*, pages 201–213. Geom. Topol. Publ., Coventry, 1999. 6

[HT74] John Hopcroft and Robert Tarjan. Efficient planarity testing. *J. Assoc. Comput. Mach.*, 21:549–568, 1974. 27

[KT21] Dale Koenig and Anastasiia Tsvietkova. NP-hard problems naturally arising in knot theory. *Trans. Amer. Math. Soc. Ser. B*, 8:420–441, 2021. 7

[Lac21] Marc Lackenby. The efficient certification of knottedness and Thurston norm. *Adv. Math.*, 387:Paper No. 107796, 142, 2021. 7

[LZ04] Sergei K. Lando and Alexander K. Zvonkin. *Graphs on surfaces and their applications*, volume 141 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004. With an appendix by Don B. Zagier, Low-Dimensional Topology, II. 8

[MT91] William W. Menasco and Morwen B. Thistlethwaite. The Tait flyping conjecture. *Bull. Amer. Math. Soc. (N.S.)*, 25(2):403–412, 1991. 33

[MT93] William Menasco and Morwen Thistlethwaite. The classification of alternating links. *Ann. of Math. (2)*, 138(1):113–171, 1993. 33

[NC01] Max Neumann-Coto. A characterization of shortest geodesics on surfaces. *Algebr. Geom. Topol.*, 1:349–368, 2001. 6

[OEI24] OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences, 2024. Published electronically at http://oeis.org. 10, 34

[Pap96] Christian Pappas. Extensions of codimension one immersions. *Trans. Amer. Math. Soc.*, 348(8):3065–3083, 1996. 19

[Poé95] Valentin Poénaru. Extension des immersions en codimension 1 (d'après Samuel Blank). In *Séminaire Bourbaki, Vol. 10*, pages Exp. No. 342, 473–505. Soc. Math. France, Paris, 1995. 4, 19

[Rei62] Bruce L. Reinhart. Algorithms for Jordan curves on compact surfaces. *Ann. of Math. (2)*, 75:209–222, 1962. 13

[Sim23] Christopher-Lloyd Simon. Loops in surfaces, chord diagrams, interlace graphs: operad factorisations and generating grammars, 2023. Submitted for publication, arxiv version. 5, 32

[SS24a] Christopher-Lloyd Simon and Ben Stucky. Loopin: Code to accompany The pinning ideal of a multiloop, 2024. see https://github.com/ChristopherLloyd/LooPin. 33, 34, 37

[SS24b] Christopher-Lloyd Simon and Ben Stucky. The LooPindex : a digital catalog of pinning data for small multiloops, 2024. Self-published electronically, Github. 5, 34

[Ste51] S. K. Stein. Convex maps. *Proc. Amer. Math. Soc.*, 2:464–466, 1951. 28

[SV92] Peter W. Shor and Christopher J. Van Wyk. Detecting and decomposing self-overlapping curves. *Computational Geometry*, 2(1):31–50, 1992. 19

[Tho59] R. Thom. Remarques sur les problèmes comportant des inéquations différentielles globales. *Bulletin de la Société Mathématique de France*, 87:455–461, 1959. 19

[Wag36] K. Wagner. Bemerkungen zum vierfarbenproblem. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 46:26–32, 1936. 28