The unbearable hardness of unknotting

Arnaud de Mesmay, Yoav Rieck, Eric Sedgwick, Martin Tancer

Version of Record: https://www.sciencedirect.com/science/article/pii/S0001870821000864
Manuscript_0946a04a6aeb12b1d8e42f90a6557493

The Unbearable Hardness of Unknotting*

Arnaud de Mesmay[1], Yo'av Rieck[2], Eric Sedgwick[3], and Martin Tancer[4]

[1]Universit Paris-Est, LIGM, CNRS, ENPC, ESIEE Paris, UPEM, Marne-la-Valle, France
[2]Department of Mathematical Sciences, University of Arkansas Fayetteville, AR 72701, USA
[3]School of Computing, DePaul University, 243 S. Wabash Ave, Chicago, IL 60604, USA
[4]Department of Applied Mathematics, Charles University, Malostranské nám. 25, 118 00 Praha 1, Czech Republic.

February 1, 2021

Abstract

We prove that deciding if a diagram of the unknot can be untangled using at most k Reidemeister moves (where k is part of the input) is **NP**-hard. We also prove that several natural questions regarding links in the 3-sphere are **NP**-hard, including detecting whether a link contains a trivial sublink with n components, computing the unlinking number of a link, and computing a variety of link invariants related to four-dimensional topology (such as the 4-ball Euler characteristic, the slicing number, and the 4-dimensional clasp number).

Contents

*AdM is partially supported by the French ANR project ANR-16-CE40-0009-01 (GATO) and the CNRS PEPS project COMP3D. AdM and MT are partially supported by the Czech-French collaboration project EMBEDS II (CZ: 7AMB17FR029, FR: 38087RM). This work was partially supported by a grant from the Simons Foundation (grant number 283495 to Yo'av Rieck). MT is partially supported by the project CE-ITI (GAČR P202/12/G061) and by the Charles University projects PRIMUS/17/SCI/3 and UNCE/SCI/004.

1 Introduction

Unknot recognition via Reidemeister moves. The *unknot recognition problem* asks whether a given knot is the unknot. Decidability of this problem was established by Haken [Hak61], and since then several other algorithms were constructed (see, e.g., the survey of Lackenby [Lac17a]).

One can ask, naively, if one can decide whether a given knot diagram represents the unknot simply by untangling the diagram: trying various Reidemeister moves until there are no more crossings. A first issue is that one might need to *increase* the number of crossings at some point in this untangling: examples of "hard unknots" witnessing this necessity can be found in Kaufman and Lambropoulou [KL14]. The problem then, obviously, is knowing when to stop: if we have not been able to untangle the diagram using so many moves, is the knot in question necessarily knotted or should we keep on trying?

In [HL01], Hass and Lagarias gave an explicit (albeit rather large) bound on the number of Reidemeister moves needed to untangle a diagram of the unknot. Lackenby [Lac15] improved the bound to polynomial thus showing that the unknot recognition problem is in **NP** (this was previously proved in [HLP99]). The unknot recognition problem is also in co-**NP** [Lac16] (assuming the Generalized Riemann Hypothesis, this was previously shown by Kuperberg [Kup14]). Thus if the unknot recognition problem were **NP**-complete (or co-**NP**-complete) we would have that **NP** and co-**NP** coincide which is commonly believed not to be the case; see, e.g., [GJ79, Section 7.1]. This suggests that the unknot recognition problem is *not* **NP**-hard.

It is therefore natural to ask if there is a way to use Reidemeister moves leading to a better solution than a generic brute-force search. Our main result suggests that there may be serious difficulties in such an approach: given a 3-SAT instance Φ we construct an unknot diagram and a number k, so that the diagram can be untangled using at most k Reidemeister moves if and only if Φ is satisfiable. Hence any algorithm that can calculate the minimal number of Reidemeister moves needed to untangle unknot diagrams – an algorithm that can actually unknot the diagram optimally – will be robust enough to tackle any problem in **NP**.

The main result of this paper is:

Theorem 1. *Given an unknot diagram D and an integer k, deciding if D can be untangled using at most k Reidemeister moves is* **NP***-complete.*

NP-membership follows from the result of Lackenby [Lac15], so we only show **NP**-hardness. For the reduction in the proof of Theorem 1 we have to construct arbitrarily large diagrams of the unknot. The difficulty in the proof is to establish tools powerful enough to provide useful lower bounds on the minimal number of Reidemeister moves needed to untangle these diagrams. For instance, the algebraic methods of Hass and Nowik [HN10] do not appear strong enough for our reduction. It is also quite easy to modify the construction and give more easily lower bounds on the number of Reidemeister moves needed to untangle unlinks if one allows the use of arbitrarily many components of diagrams with constant size, but those techniques too cannot be used for Theorem 1. We develop the necessary tools in Section 4.

Computational problems for links. Our approach for proving Theorem 1 partially builds on techniques to encode satisfiability instances using Hopf links and Borromean rings, that we previously used in [dMRST18] (though the technical details are very different). With these techniques, we also show that a variety of link invariants are **NP**-hard to compute. Precisely, we prove:

Theorem 2. *Given a link diagram L and an integer k, the following problems are* **NP***-hard:*

(a) deciding whether L admits a trivial unlink with k components as a sublink.

(b) deciding whether an intermediate invariant has value k on L,

(c) deciding whether $\chi_4(L) = 0$,

(d) deciding whether L admits a smoothly slice sublink with k components.

We refer to Definition 12 for the definitions of $\chi_4(L)$, the 4-ball Euler characteristic, and intermediate invariants. These are broadly related to the topology of the 4-ball, and include the unlinking number, the ribbon number, the slicing number, the concordance unlinking number, the concordance ribbon number, the concordance slicing number, and the 4-dimensional clasp number. See, e.g., [Shi74] for a discussion of many intermediate invariants.

Related complexity results. The complexity of computational problems pertaining to knots and links is quite poorly understood. In particular, only very few computational lower bounds are known, and as far as we know, almost none concern classical knots (i.e., knots embedded in S^3): apart from our Theorem 1, the only other such hardness proof we know of [KS, Sam18] concerns counting coloring invariants (i.e., representations of the fundamental group) of knots. More lower bounds are known for classical links. Lackenby [Lac17b] showed that determining if a link is a sublink of another one is **NP**-hard. Our results strengthen this by showing that even finding an n-component unlink as a sublink is already **NP**-hard. Agol, Hass and Thurston [AHT06] showed that computing the genus of a knot in a 3-manifold is **NP**-hard, and Lackenby [Lac17b] showed that computing the Thurston complexity of a link in S^3 is also **NP**-hard. Our results complement this by showing that the 4-dimensional version of this problem is also **NP**-hard.

Regarding upper bounds, the current state of knowledge is only slightly better. While, as we mentioned before, it is now known that the unknot recognition problem is in **NP**$\cap$ co-**NP**, many natural link invariants are not even known to be decidable. In particular, this is the

case for all the invariants for which we prove **NP**-hardness, except for the problem of finding the maximal number of components of a trivial sublink, which is in **NP** (see Theorem 4).

Shortly before we finished our manuscript, Koenig and Tsvietkova posted a preprint [KT18] that also shows that certain computational problems on links are **NP**-hard, with some overlap with the results obtained in this paper (the trivial sublink problem and the unlinking number). They also show **NP**-hardness of computing the number of Reidemeister moves for getting between two diagrams of the unlink, but their construction does not untangle the diagram and requires arbitrarily many components. Theorem 1 of the current paper is stronger and answers Question 17 of [KT18].

Organization. This paper is organized as follows. After some preliminaries in Section 2, we start by proving the hardness of the trivial sublink problem in Part I because it is very simple and provides a good introduction for our other reductions. We then proceed to prove Theorem 1 in Part II and the hardness of the unlinking number and the other invariants in Part III. The three parts are independent and the reader can read any one part alone.

2 Preliminaries

Notation. Most of the notation we use is standard. By *knot* we mean a tame piecewise linear embedding of the circle S^1 into the 3-sphere S^3. By *link* we mean a tame, piecewise linear embedding of the disjoint union of any finite number of copies of S^1. We use interval notation for natural numbers, for example, $[3, 5]$ means $\{3, 4, 5\}$; we use $[12]$ to indicate $[1, 12]$. We assume basic familiarity with computational complexity and knot theory, and refer to basic textbooks such as Arora and Barak [AB09] for the former and Rolfsen [Rol90] for the latter.

Diagram of a knot or a link. All the computational problems that we study in this paper take as input the *diagram* of a knot or a link, which we define here.

A diagram of a knot is a piecewise linear map $D\colon S^1 \to \mathbb{R}^2$ in general position; for such a map, every point in $\mathbb{R}^2$ has at most two preimages, and there are finitely many points in $\mathbb{R}^2$ with exactly two preimages (called *crossings*). Locally at crossing two arcs cross each other transversely, and the diagram contains the information of which arc passes 'over' and which 'under'. This we usually depict by interrupting the arc that passes under. (A diagram usually arises as a composition of a (piecewise linear) knot $\kappa\colon S^1 \to \mathbb{R}^3$ and a generic projection $\pi\colon \mathbb{R}^3 \to \mathbb{R}^2$ which also induces 'over' and 'under' information.) We usually identify a diagram D with its image in $\mathbb{R}^2$ together with the information about underpasses/overpasses at crossings; see, for example, Figure 3, ignoring the notation on the picture. Diagrams are considered up-to isotopy.

Similarly, a diagram of a link is a piecewise linear map $D\colon \coprod S^1 \to \mathbb{R}^2$ in general position, where $\coprod$ denotes a disjoint union of a finite number of circles S^1, and with the same additional information at the crossings.

By an *arc* in the diagram D we mean a set $D(\alpha)$ where α is an arc in S^1, i.e., a subset of S^1 homeomorphic to the closed interval (note that this definition is slightly non-standard but it will be very useful later on).

The size of a knot or a link diagram is its number of crossings plus number of components of the link. Up to a constant factor, this complexity exactly describes the complexity of encoding the combinatorial information contained in a knot or link diagram.

3-satisfiability. A formula in *conjunctive normal form* in variables $x_1, \ldots, x_n$ is a Boolean formula of the form $c_1 \wedge c_2 \wedge \cdots \wedge c_m$ where each c_i is a *clause*, that is, a formula of the form $(\ell_1 \vee \ell_2 \vee \cdots \vee \ell_k)$ where each ℓ_j is a *literal*, that is, a variable x_t or its negation $\neg x_t$. A formula Φ is *satisfiable* if there is an assignment to the variables (each variable is assigned TRUE or FALSE) such that Φ evaluates to TRUE in the given assignment.

A 3-SAT problem is a well-known **NP**-hard problem. On input there is a formula Φ in conjunctive normal form such that every clause contains exactly[1] three variables; see, e.g., [Pap94, Proposition 9.2].

Part I
Trivial sublink

Informally, the trivial sublink problem asks, given a link L and a positive integer n, whether L admits the n-component unlink as a sublink. We define:

Definition 3 (The Trivial Sublink Problem)**.** An *unlink,* or a *trivial link,* is a link in S^3 whose components bound disjointly embedded disks. A *trivial sublink* of a link L is an unlink formed by a subset of the components of L. The *trivial sublink problem* asks, given a link L and a positive integer n, whether L admits an n component trivial sublink.

Theorem 4. *The trivial sublink problem is* **NP***-complete.*

Note that Theorem 4 is just a slight extension of Theorem 2(a), claiming also **NP**-membership. The essential part is **NP**-hardness.

Proof. It follows from Hass, Lagarias, and Pippenger [HLP99] that deciding if a link is trivial is in **NP**. By adding to their certificate a collection of n components of L we obtain a certificate for the trivial sublink problem, showing that it is in **NP** (**NP**-membership of the trivial sublink problem was also established, using completely different techniques, by Lachenby [Lac15]). Thus all we need to show is that the problem is **NP**-hard. We will show this by reducing 3-SAT to the trivial sublink problem.

Given a 3-SAT instance Φ, with n variables (say $x_1, \ldots, x_n$) and m clauses, we construct a diagram D_Φ as follows (see Figure 1): we first mark $n+m$ disjoint disks in the plane. In each of the first n disks we draw a diagram of the Hopf link, marking the components in the ith disk as κ_{x_i} and $\kappa_{\neg x_i}$. In the remaining m disks we draw diagrams of the Borromean rings and label them according to the clauses of Φ. We now band each component of the Borromean rings to the Hopf link component with the same label. Whenever two bands cross we have one move over the other (with no "weaving"); we assume, as we may, that no two bands cross twice. It is easy to see that this can be done in polynomial time. The diagram we obtain is D_Φ and the link it represents is denoted L_Φ (note that L_Φ has exactly $2n$ components). We complete the proof by showing that L_Φ admits an n-component trivial sublink exactly when Φ is satisfiable.

Claim 4.1. *If Φ is satisfiable then L_Φ admits an n-component trivial sublink.*

[1] Here we adopt a convention from [Pap94]. Some other authors define require only 'at most three' literals.

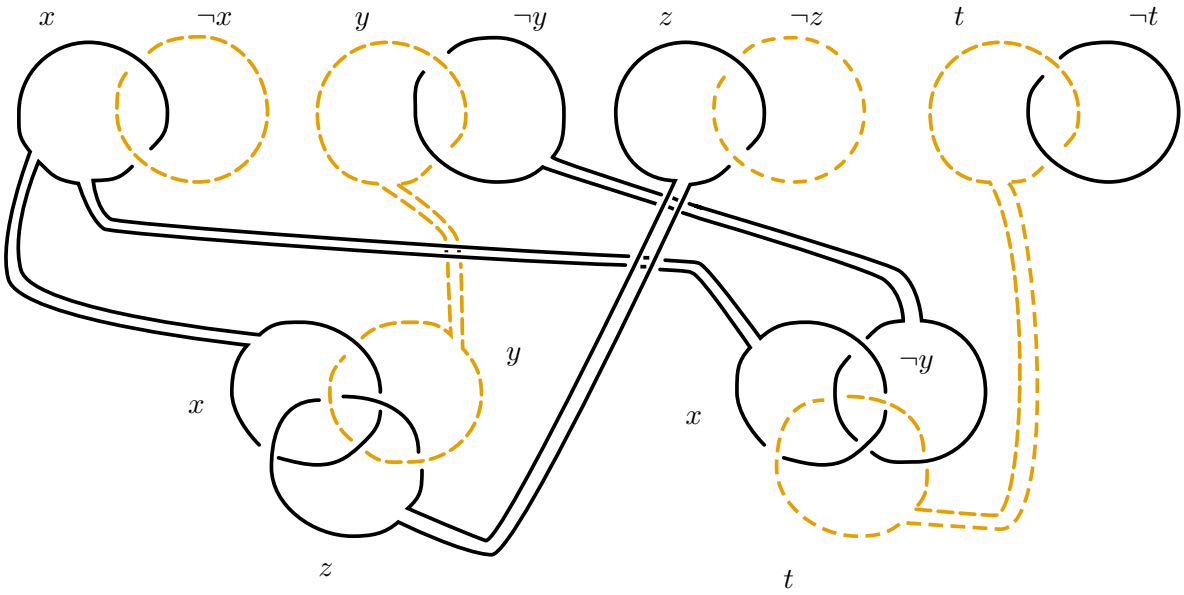

Figure 1: D_Φ for $\Phi = (x \vee y \vee z) \wedge (x \vee \neg y \vee t)$, a satisfying assignment is dashed.

Given a satisfying assignment we remove from L_Φ the components that correspond to satisfied literals, that is, if $x_i = \text{TRUE}$ we remove κ_{x_i} from L_Φ and if $x_i = \text{FALSE}$ we remove $\kappa_{\neg x_i}$ from L_Φ. We claim that the remaining n components form an unlink. To see this, first note that since the assignment is satisfying, from each copy of the Borromean rings at least one component was removed. Therefore the rings fall apart and (since we did not allow "weaving") the diagram obtained retracts into the first n disks. In each of these disks we had, originally, a copy of the Hopf link; by construction exactly one component was removed. This shows that the link obtained is indeed the n-component unlink; Claim 4.1 follows.

Claim 4.2. *If L_Φ admits an n-component trivial sublink, then Φ is satisfiable.*

Suppose that L_Φ admits an n-component trivial sublink $\mathcal{U}$. Since $\mathcal{U}$ itself does not admit the Hopf link as a sublink, for each i, at most one of κ_{x_i} and $\kappa_{\neg x_i}$ is in $\mathcal{U}$. Since $\mathcal{U}$ has n components we see that *exactly* one of κ_{x_i} and $\kappa_{\neg x_i}$ is in $\mathcal{U}$. If κ_{x_i} is in $\mathcal{U}$ we set $x_i = \text{FALSE}$ and if $\kappa_{\neg x_i}$ is in $\mathcal{U}$ we set $x_i = \text{TRUE}$. Now since $\mathcal{U}$ does not admit the Borromean rings as a sublink, from each copy of the Borromean rings at least one component is not in $\mathcal{U}$. It follows that in each clause of Φ at least one literal is satisfied, that is, the assignment satisfies Φ; Claim 4.2 follows.

This completes the proof of Theorem 4. □

Part II
The number of Reidemeister moves for untangling

3 A restricted form of the satisfiability problem

For the proof of Theorem 1, we will need a slightly restricted form of the 3-SAT problem given by lemma below.

Lemma 5. *Deciding whether a formula Φ in conjunctive normal form is satisfiable is* **NP**-*hard even if we assume the following conditions on Φ.*

- *Each clause contains exactly three literals.*
- *No clause contains both x and $\neg x$ for some variable x.*
- *Each pair of literals $\{\ell_1, \ell_2\}$ occurs in at most one clause.*

Proof. The first condition says that we consider the 3-SAT problem. Any clause violating the second condition can be removed from the formula without affecting satisfiability of the formula as such clause is always satisfied. Therefore, it is sufficient to provide a recipe to build in polynomial time a formula Φ satisfying the three conditions above out of a formula Φ' satisfying only the first two conditions.

First, we consider an auxiliary formula

$$\Psi = \Psi(t, a, b, c) = (t \vee a \vee \neg b) \wedge (t \vee b \vee \neg c) \wedge (t \vee c \vee \neg a) \wedge (a \vee b \vee c) \wedge (\neg a \vee \neg b \vee \neg c).$$

We observe that for any satisfying assignment of Ψ we get that t is assigned TRUE. Indeed, if t were assigned FALSE then $(t \vee a \vee \neg b) \wedge (t \vee b \vee \neg c) \wedge (t \vee c \vee \neg a)$ translates as $(b \Rightarrow a) \wedge (c \Rightarrow b) \wedge (a \Rightarrow c)$, that is all a, b, and c are equivalent. However, then $(a \vee b \vee c) \wedge (\neg a \vee \neg b \vee \neg c)$ cannot be satisfied.

On the other hand, we also observe that there is a satisfying assignment for Ψ where t is assigned TRUE and, for example, it is sufficient to assign a with TRUE, b with FALSE and c arbitrarily.

Now we return to the formula Φ' discussed above. Suppose there exists a pair literals ℓ_1 and ℓ_2 contained in two clauses of Φ', say $(\ell_1 \vee \ell_2 \vee \ell_3)$ and $(\ell_1 \vee \ell_2 \vee \ell_4)$. We replace them with

$$(\ell_1 \vee \ell_2 \vee x) \wedge (\ell_3 \vee \neg x \vee \neg t_1) \wedge (\ell_4 \vee \neg x \vee \neg t_2) \wedge \Psi(t_1, a_1, b_1, c_1) \wedge \Psi(t_2, a_2, b_2, c_2),$$

where $x, t_1, t_2, a_1, a_2, b_1, b_2, c_1$, and c_2 are newly added variables, obtaining a new formula Φ''. We aim to show that Φ'' is satisfiable if and only if Φ' is satisfiable.

Let us first assume that Φ' is satisfiable and fix a satisfying assignment. If $\ell_1 \vee \ell_2$ is TRUE in this assignment, then we may extend it to variables of Φ'' by setting x to FALSE, t_1 and t_2 to TRUE and $a_1, \ldots, c_2$ so that $\Psi(t_1, a_1, b_1, c_1) \wedge \Psi(t_2, a_2, b_2, c_2)$ is satisfied. If $\ell_1 \vee \ell_2$ is FALSE in the assignment for Φ', then ℓ_3 and ℓ_4 must be assigned TRUE. Then we may extend by setting x to TRUE and $t_1, t_2, a_1, \ldots, c_2$ as before.

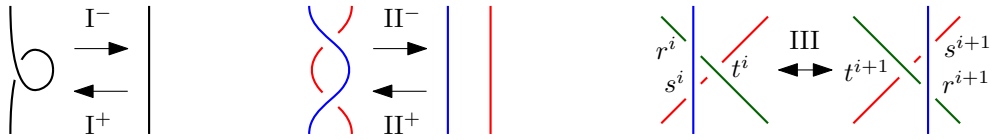

Figure 2: Reidemeister moves.

Now let us assume that Φ'' is satisfiable. Then t_1 and t_2 are set to TRUE due to the properties of Ψ. If x is TRUE, then both ℓ_3 and ℓ_4 must be true and the restriction of the assignment on Φ'' to the variables of Φ' is therefore a satisfying assignment for Φ'. Similarly, if x is FALSE, then $\ell_1 \vee \ell_2$ must be true and the restriction is again a satisfying assignment for Φ'.

We also observe that in Φ'' we have reduced the number of clauses containing the literals ℓ_1 and ℓ_2 simultaneously, and for any other pair of literals we do not increase the number of clauses containing that pair. Therefore, after a polynomial number of steps, when always adding new variables, we arrive at a desired formula Φ satisfying all three conditions. □

4 The defect

Reidemeister moves. Reidemeister moves are local modifications of a diagram depicted in Figure 2 (the labels at the crossings in a III move will be used only later on). We distinguish the I move (left), the II move (middle) and the III move (right). The first two moves affect the number of crossings, thus we further distinguish the I^- and the II^- moves which reduce the number of crossings from the I^+ and the II^+ moves which increase the number of crossings.

The number of Reidemeister moves for untangling a knot. A diagram of the unknot is *untangled* if it does not contain any crossings. The untangled diagram is denoted by U. Given a diagram D of an unknot, an *untangling* of D is a sequence $\mathbf{D} = (D^0, \dots, D^k)$ where $D^0 = D$, $D^k = U$ (recall that diagrams are only considered up to isotopy) and D^i is obtained from D^{i-1} by a single Reidemeister move. The number of Reidemeister moves in $\mathbf{D}$ is denoted by $\mathrm{rm}(\mathbf{D})$, that is, $\mathrm{rm}(\mathbf{D}) = k$. We also define $\mathrm{rm}(D) := \min \mathrm{rm}(\mathbf{D})$ where the minimum is taken over all untanglings $\mathbf{D}$ of D.

The defect. Let us denote by $\mathrm{cross}(D)$ the number of crossings in D. Then the *defect* of an untangling $\mathbf{D}$ is defined by the formula

$$\mathrm{def}(\mathbf{D}) := 2\,\mathrm{rm}(\mathbf{D}) - \mathrm{cross}(D).$$

The *defect* of a diagram D is defined as $\mathrm{def}(D) := 2\,\mathrm{rm}(D) - \mathrm{cross}(D)$. Equivalently, $\mathrm{def}(D) = \min \mathrm{def}(\mathbf{D})$ where the minimum is taken over all untanglings $\mathbf{D}$ of D.

The defect is a convenient way to reparametrize the number of Reidemeister moves due to the following observation.

Observation 6. *For any diagram D of the unknot and any untangling $\mathbf{D}$ of D we have $\mathrm{def}(\mathbf{D}) \geq 0$. Equality holds if and only if $\mathbf{D}$ uses only II^- moves.*

This notion of defect is different from the one that was introduced by Chang and Erickson [CE17] (following Arnold [Arn94a, Arn94b] and Aicardi [Aic94]) to study homotopy moves.

Proof. Every Reidemeister move in $\mathbf{D} = (D^0, \dots, D^k)$ removes at most two crossings and the II^- move is the only move that removes exactly two crossings. Therefore, the number of crossings in $D = D^0$ is at most $2k$ and equality holds if and only if every move is a II^- move. □

Crossings contributing to the defect. Let $\mathbf{D} = (D^0, \dots, D^k)$ be an untangling of a diagram $D = D^0$ of an unknot.

Given a crossing r^i in D^i, for $0 \le i \le k-1$, it may vanish by the move transforming D^i into D^{i+1} if this is a I^- or a II^- move affecting the crossing. In all other cases it *survives* and we denote by r^{i+1} the corresponding crossing in D_{i+1}. Note that in the case of a III move there are three crossings affected by the move and three crossings afterwards. Both before and after, each crossing is the unique intersection between a pair of the three arcs of the knot that appear in this portion of the diagram. So we may say that these three crossings survive the move though they change their actual geometric positions (they swap the order in which they occur along each of the three arcs); see Figure 2.

With a slight abuse of terminology, by a *crossing in* $\mathbf{D}$ we mean a maximal sequence $\mathbf{r} = (r^a, r^{a+1}, \dots, r^b)$ such that r^{i+1} is the crossing in D^{i+1} corresponding to r^i in D^i for any $i \in [a, b-1]$. By maximality we mean that r^b vanishes after the $(b+1)$st move and either $a = 0$ or r^a is introduced by the ath Reidemeister move (which must be a I^+ or II^+ move).

An *initial crossing* is a crossing $\mathbf{r} = (r^0, r^1, \dots, r^b)$ in $\mathbf{D}$. Initial crossings in $\mathbf{D}$ are in one-to-one correspondence with crossings in $D = D^0$. For simplicity of notation, r^0 is also denoted r (as a crossing in D).

A Reidemeister II^- move in $\mathbf{D}$ is *economical* if both crossings removed by this move are initial crossings; otherwise, it is *wasteful*.

Let $m_3(\mathbf{r})$ be the number of III moves affecting a crossing $\mathbf{r}$. The *weight* of an initial crossing $\mathbf{r}$ is defined in the following way.

$$w(\mathbf{r}) = \frac{2}{3} m_3(\mathbf{r}) + \begin{cases} 0 & \text{if } \mathbf{r} \text{ vanishes by an economical } \mathrm{II}^- \text{ move;} \\ 1 & \text{if } \mathbf{r} \text{ vanishes by a } \mathrm{I}^- \text{ move;} \\ 2 & \text{if } \mathbf{r} \text{ vanishes by a wasteful } \mathrm{II}^- \text{ move.} \end{cases}$$

For later purposes, we also define $w(r) := w(\mathbf{r})$ and $w(R) := \sum_{r \in R} w(r)$ for a subset R of the set of all crossings in D.

Lemma 7. *Let* $\mathbf{D}$ *be an untangling of a diagram* D*. Then*

$$\mathrm{def}(\mathbf{D}) \ge \sum_{\mathbf{r}} w(\mathbf{r}),$$

where the sum is over all initial crossings $\mathbf{r}$ *of* $\mathbf{D}$*.*

Proof. In the proof we use the discharging technique, common in graph theory.

Let us put charges on crossings in $\mathbf{D}$ and on Reidemeister moves used in $\mathbf{D}$. The initial charge will be

2 on each Reidemeister move;

-1 on each initial crossing; and

0 on each non-initial crossing.

We remark that the sum of the initial charges equals to $\mathrm{def}(\mathbf{D})$ by the definition of the defect.

Now we start redistributing the charge according to the rules described below. The aim is that, after the redistribution, the charge on each initial crossing will be at least $w(\mathbf{r})$ and it will be at least 0 on non-initial crossings and on Reidemeister moves. This will prove the lemma, as the sum of the charges after the redistribution is still equal to the defect, whereas it will be at least the sum of the weights of initial crossings.

We apply the following rules for the redistribution of the charge.

(R1) Every I^+ move sends charge 2 to the (non-initial) crossing it creates.

(R2) Every I^- move sends charge 2 to the crossing it removes.

(R3) Every II^+ move sends charge 1 to each of the two (non-initial) crossings it creates.

(R4) Every economical II^- move sends charge 1 to each of the two (initial) crossings it removes.

(R5) Every wasteful II^- move that removes exactly one initial crossing sends charge 3 to this initial crossing.

(R6) Every III move sends $\frac{2}{3}$ to every crossing it affects.

(R7) Every non-initial crossing which is removed by a wasteful II^- move sends charge 1 to this move.

Now it is routine to check that the desired conditions are satisfied, which we now explain.

Every move has charge at least 0: The initial charge on moves is 2. The rules are set up so that every move distributes charge at most 2 with exception of the rule (R5). However, in this case, the wasteful II^- move that removes exactly one initial crossing from (R5) gets 1 charge from rule (R7).

Every non-initial crossing has charge at least 0: The initial charge is 0. The only rule that depletes the charge is (R7); however in such case, the charge is replenished by (R1) or (R3).

Every initial crossing $\mathbf{r}$ has charge equal at least $w(\mathbf{r})$: The initial charge is -1. First we observe that (R6) sends the charge $\frac{2}{3}m_3(\mathbf{r})$ to $\mathbf{r}$. If $\mathbf{r}$ vanishes by an economical II^- move, it gets additional charge 1 by (R4). If $\mathbf{r}$ vanishes by a I^- move, it gets additional charge 2 by (R2). Finally, if $\mathbf{r}$ vanishes by a wasteful II^- move then this move removes exactly one initial crossing, namely $\mathbf{r}$. Therefore $\mathbf{r}$ gets an additional charge 3 by (R5). □

We will also need a variant for the previous lemma where we get equality, if we use the I^- and II^- moves only.

Lemma 8. *Let* $\mathbf{D}$ *be an untangling of a digram* D *which uses the* I^- *and* II^- *moves only. Then*

$$\mathrm{def}(\mathbf{D}) = \sum_{\mathbf{r}} w(\mathbf{r}) = \text{ number of } \mathrm{I}^- \text{ moves},$$

where the sum is over all initial crossings $\mathbf{r}$ *of* $\mathbf{D}$.

Proof. Let k_1 be the number of initial crossings removed by a I^- move and k_2 the number of initial crossings removed by a II^- move. Then $\mathrm{rm}(\mathbf{D}) = k_1 + k_2/2$ and from the definition of the defect we get $\mathrm{def}(\mathbf{D}) = 2(k_1 + k_2/2) - (k_1 + k_2) = k_1$. It follows directly from the definition of the weight that $k_1 = \sum_{\mathbf{r}} w(\mathbf{r})$. □

Twins and the preimage of a bigon. Let $\mathbf{r}$ be an initial crossing in an untangling $\mathbf{D} = (D^0, \ldots, D^k)$ removed by an economical II^- move. The *twin* of $\mathbf{r}$, denoted by $t(\mathbf{r})$ is the other crossing in $\mathbf{D}$ removed by the same II^- move. Note that $t(\mathbf{r})$ is also an initial crossing (because the move is economical). We also get $t(t(\mathbf{r})) = \mathbf{r}$. If $\mathbf{r} = (r^0, \ldots, r^b)$, then we also extend the definition of a twin to D^i in such a way that $t(r^i)$ is uniquely defined by $t(\mathbf{r}) = (t(r^0), \ldots, t(r^b))$. In particular, we will often use a twin $t(r)$ of a crossing $r = r^0$ in D (if it exists).

Furthermore, the crossings r^b and $t(r^b)$ in D^b form a bigon that is removed by the forthcoming II^- move. Let $\alpha^b(r)$ and $\beta^b(r)$ be the two arcs of the bigon (with endpoints r^b and $t(r^b)$) so that $\alpha^b(r)$ is the arc that, after extending slightly, overpasses the crossings r^b and $t(r^b)$ whereas a slight extension of $\beta^b(r)$ underpasses these crossings. (The reader may remember this as α is 'above' and β is 'below'.)[2] Now we can inductively define arcs $\alpha^i(r)$ and $\beta^i(r)$ for $i \in [0, b-1]$ so that $\alpha^i(r)$ and $\beta^i(r)$ are the unique arcs between r^i and $t(r^i)$ which are transformed to (already defined) $\alpha^{i+1}(r)$ and $\beta^{i+1}(r)$ by the $(i+1)$st Reidemeister move. We also set $\alpha(r) = \alpha^0(r)$ and $\beta(r) = \beta^0(r)$. Intuitively, $\alpha(r)$ and $\beta(r)$ form a preimage of the bigon removed by the bth move and they are called the *preimage arcs* between r and $t(r)$.

Close neighbors. Let R be a subset of the set of crossings in D. Let r and s be any two crossings in D (not necessarily in R) and let c be a non-negative integer. We say that r and s are *c-close neighbors with respect to R* if r and s can be connected by two arcs α and β such that

- α enters r and s as an overpass;
- β enters r and s as an underpass;
- α and β may have self-crossings; however, neither r nor s is in the interior of α or β; and
- α and β together contain at most c crossings from R in their interiors. (If there is a crossing in the interior of both α and β, this crossing is counted only once.)

Lemma 9. *Let R be a subset of the set of crossings in D, let $c \in \{0, 1, 2, 3\}$. Let r be the crossing in R which is the first of the crossings in R removed by an economical II^- move (we allow a draw). If $w(R) \leq c$, then r and its twin $t(r)$ are c-close neighbors with respect to R.*

Proof. Let $\alpha(r)$ and $\beta(r)$ be the preimage arcs between r and $t(r)$. We want to verify that they satisfy the properties of the arcs from the definition of the close neighbors. The first two properties follow immediately from the definition of preimage arcs.

Next, we want to check that neither r nor $t(r)$ is the interior of $\alpha(r)$ or $\beta(r)$. For contradiction, let us assume that this not the case. For example, suppose that r also lies in the

[2]Note that $\alpha^b(r)$ and $\beta^b(r)$ are uniquely defined by r as well as by $\mathbf{r}$. The choice of r in the notation will be more useful later on.

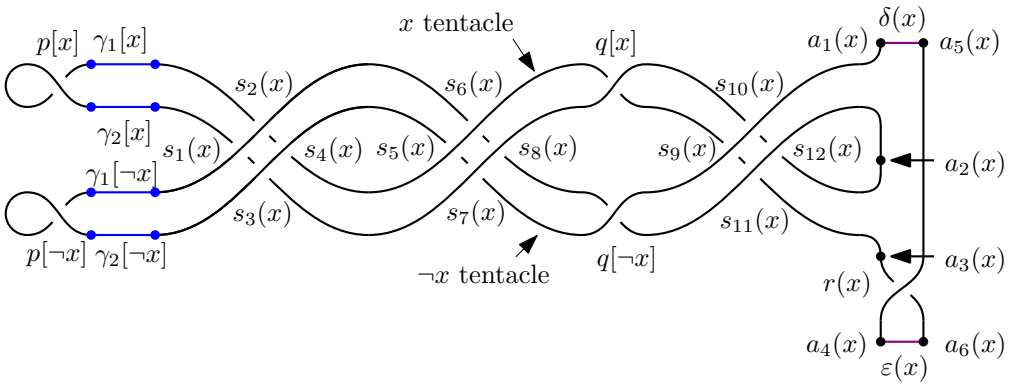

Figure 3: The variable gadget $V(x)$.

interior of $\alpha(r)$. Let $\mathbf{r} = (r, r^1, \ldots, r^b)$ be the initial crossing corresponding to r. The $(b+1)$st Reidemeister move in $\mathbf{D}$ removes $\mathbf{r}$. Any preceding Reidemeister move either does not affect $\mathbf{r}$ at all, or it is a III move swapping the crossing with other crossings. In any case, it preserves the self-crossing of $\alpha(r)$ at $\mathbf{r}$. However, this contradicts the fact that $\alpha(r^b)$ is an arc of a bigon removed by the $(b+1)$st move.

In order to check the last property, let us assume that, for contradiction, $\alpha(r)$ and $\beta(r)$ together contain at least $c+1$ crossings from $R(x)$ in their interiors. These crossings have to be removed from the arcs $\alpha(r)$ and $\beta(r)$ until we reach $\alpha(r^b)$ and $\beta(r^b)$. They cannot be removed by an economical II^- move as r is the first crossing from $R(x)$ removed by such a move. Thus they have to be removed from the arcs either by a I^- move, a wasteful II move or a III move (by swapping with r or $t(r)$). This contradicts $w(x) \leq c$. Indeed, if only I^- and II^- moves are used, we get a total weight at least $c+1$ on the crossings; if at least one III move is used, we get a weight at least $(c+1) \cdot \frac{2}{3}$ on the crossings and an additional $\frac{2}{3}$ on r or $t(r)$. This is in total more than c as $c \leq 3$. □

5 The reduction

Let Φ be a formula in conjunctive normal form satisfying the conditions stated in Lemma 5 and let n be the number of variables. Our aim is to build a diagram $D(\Phi)$ by a polynomial-time algorithm such that $\mathrm{def}(D(\Phi)) \leq n$ if and only if Φ is satisfiable.

The variable gadget. First we describe the variable gadget. For every variable x we consider the diagram depicted at Figure 3 and we denote it $V(x)$.

The gadget contains 17 crossings $p[x], p[\neg x], q[x], q[\neg x]$, $r(x)$ and $s_i(x)$ for $i \in [12]$.

The variable gadget also contains six distinguished arcs $\gamma_i[x]$ and $\gamma_i[\neg x]$ for $i \in [2]$, $\delta(x)$ and $\varepsilon(x)$ and six distinguished auxiliary points $a_1(x), \ldots, a_6(x)$ which will be useful later on in order to describe how the variable gadget is used in the diagram $D(\Phi)$.

We also call the arc between $a_1(x)$ and $a_2(x)$ which contains $\gamma_1[\neg x]$ and $\gamma_2[\neg x]$ the $\neg x$ *tentacle*, and similarly, the arc between $a_2(x)$ and $a_3(x)$ which contains $\gamma_1[x]$ and $\gamma_2[x]$ is the x *tentacle*. Informally, a satisfying assignment to Φ will correspond to the choice whether we will decide to remove first the loop at $p[x]$ by a I^- move and simplify the x tentacle or

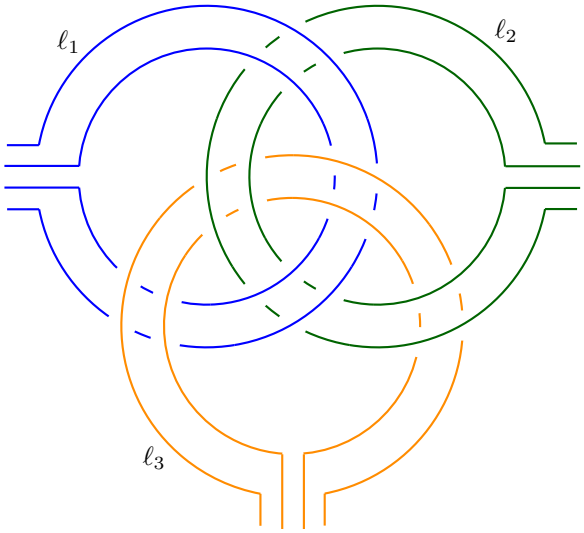

Figure 4: The clause gadget for clause $c = (\ell_1 \vee \ell_2 \vee \ell_3)$.

whether we remove first the loop at $p[\neg x]$ and remove the $\neg x$ tentacle in the final construction of $D(\Phi)$.

We also remark that in the notation, we use square brackets for objects that come in pairs and will correspond to a choice of literal $\ell \in \{x, \neg x\}$. This regards $p[\ell]$, $q[\ell]$, $\gamma_1[\ell]$ and $\gamma_2[\ell]$ whereas we use parentheses for the remaining objects.

The clause gadget. Given a clause $c = (\ell_1 \vee \ell_2 \vee \ell_3)$ in Φ, the *clause gadget* is depicted at Figure 4. The construction is based on the Borromean rings. It contains three pairs of arcs (distinguished by color) and with a slight abuse of notation, we refer to each of the three pairs of arcs as a "ring". Note that each ring has four pendent endpoints (or leaves) as in the picture. Each ring corresponds to one of the literals ℓ_1, ℓ_2, and ℓ_3.

A blueprint for the construction. Now we build a blueprint for the construction of $D(\Phi)$. Let $x_1, \ldots, x_n$ be the variables of Φ and let $c_1, \ldots, c_m$ be the clauses of Φ.

For each clause $c_j = (\ell_1 \vee \ell_2 \vee \ell_3)$ we take a copy of the graph $K_{1,3}$ (also known as the star with three leaves). We label the vertices of degree 1 of such a $K_{1,3}$ by the literals ℓ_1, ℓ_2, and ℓ_3. Now we draw these stars into the plane sorted along a horizontal line; see Figure 5.

Next for each literal $\ell \in \{x_1, \ldots, x_n, \neg x_1, \ldots, \neg x_n\}$ we draw a piecewise linear segment containing all vertices labelled with that literal according to the following rules (follow Figure 5).

- The segments start on the right of the graphs $K_{1,3}$ in the top down order $x_1, \neg x_1$, $x_2, \neg x_2, \ldots, x_n, \neg x_n$.
- They continue to the left while we permute them to the order $x_1, \ldots, x_n, \neg x_1, \ldots, \neg x_n$. We also require that $x_1, \ldots, x_n$ occur above the graphs $K_{1,3}$ and $\neg x_1, \ldots, \neg x_n$ occur below these graphs (everything is still on the right of the graphs).

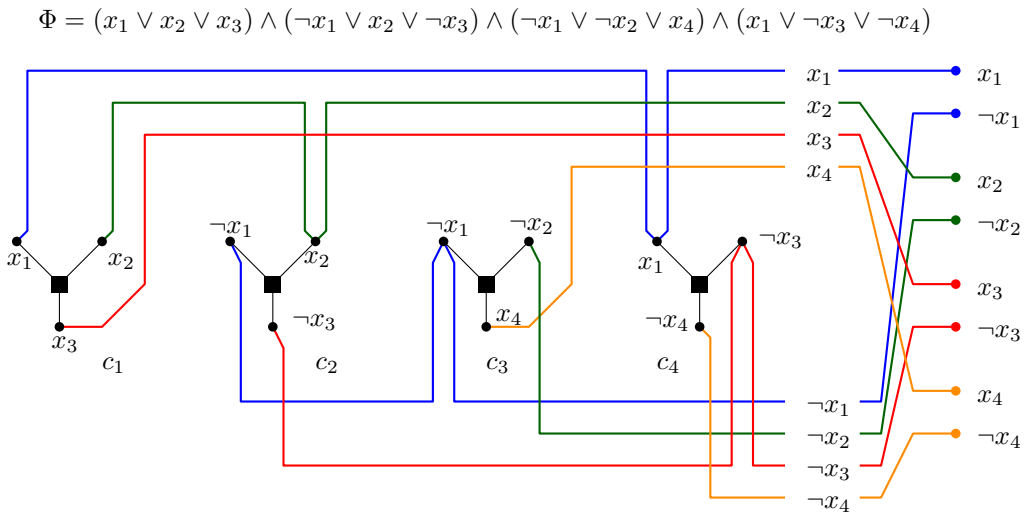

Figure 5: A blueprint for the construction of $D(\Phi)$.

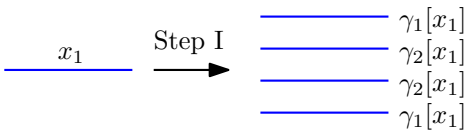

Figure 6: Step I: Replacing segments.

- Next, for each literal ℓ the segment for ℓ continues to the left while it makes a 'detour' to each vertex v labelled ℓ. If v is not the leftmost vertex labelled ℓ, then the detour is performed by a 'finger' of two parallel lines. We require that the finger avoids the graphs $K_{1,3}$ except of the vertex v. If v is the leftmost vertex labelled ℓ, then we perform only a half of the finger so that v becomes the endpoint of the segment.

Note that the segments often intersect each other; however, for any $i \in [n]$ the segments for x_i and $\neg x_i$ do not intersect (using the assumption that no clause contains both x_i and $\neg x_i$).

The final diagram. Finally, we explain how to build the diagram $D(\Phi)$ from the blueprint above.

Step I (four parallel segments): We replace each segment for a literal ℓ with four parallel segments; see Figure 6. The outer two will correspond to the arc $\gamma_1[\ell]$ from the variable gadget and the inner two will correspond to $\gamma_2[\ell]$; compare with Figure 3.

Step II (clause gadgets): We replace each copy of $K_{1,3}$ by a clause gadget for the corresponding clause c; see Figure 7. Now we aim to describe how is the clause gadget connected to the quadruples of parallel segments obtained in Step I. Let v be a degree 1 vertex of the $K_{1,3}$ we are just replacing. Let ℓ be the literal which is the label of this vertex. Then c may or may not be the leftmost clause containing a vertex labelled ℓ.

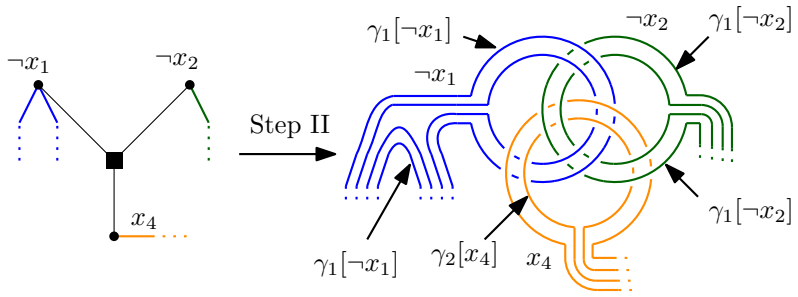

Figure 7: Step II: Replacing the $K_{1,3}$.

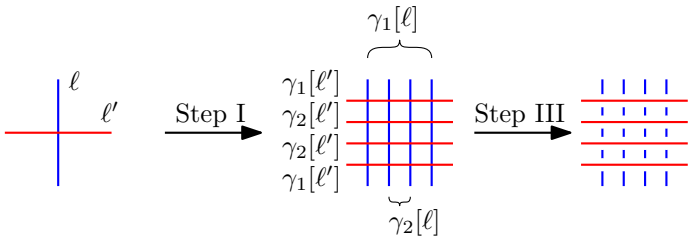

Figure 8: Step III: Resolving crossings.

If c is the leftmost clause containing a vertex labelled ℓ, then there are four parallel segments for ℓ with pendent endpoints (close to the original position of v) obtained in Step I. We connect them to the pendent endpoints of the clause gadget (on the ring for ℓ); see $\neg x_2$ and x_4 in Figure 7. Note also that at this moment the two $\gamma_1[\ell]$ arcs introduced in Step I merge as well as the two $\gamma_2[\ell]$ arcs merge.

If c is not the leftmost clause labelled ℓ then there are four parallel segments passing close to v (forming a tip of a finger from the blueprint). We disconnect the two segments closest to the tip of the finger and connect them to the pendent endpoints of the clause gadget (on the ring for ℓ); see $\neg x_1$ in Figure 7.

Step III (resolving crossings): If two segments in the blueprint, corresponding to literals ℓ and ℓ' have a crossing, Step I blows up such a crossing into 16 crossing of corresponding quadruples. We resolve overpasses/underpasses at all these crossings in the same way. That is, one quadruple overpasses the second quadruple at all 16 crossings; see Figure 8.

However, we require one additional condition on the choice of overpasses/underpasses. If ℓ and ℓ' appear simultaneously in some clause c we have 8 crossings on the rings for ℓ and ℓ' in the clause gadget for c. We can assume that the ring of ℓ passes over the ring of ℓ' at all these crossings (otherwise we swap ℓ and ℓ'). Then for the 16 crossings on segments for ℓ and ℓ' we pick the other option, that is we want that the $\gamma_1[\ell]$ and $\gamma_2[\ell]$ arcs underpass the $\gamma_1[\ell']$ and $\gamma_2[\ell']$ arcs at these crossings. This is a globally consistent choice because we assume that there is at most one clause containing both ℓ and ℓ', this is the third condition in the statement of Lemma 5.

Step IV (the variable gadgets): Now, for every variable x_i, the segments $\gamma_1[x_i], \gamma_2[x_i], \gamma_1[\neg x_i]$ and $\gamma_2[\neg x_i]$ do not intersect each other. We extend them to a variable gadget as in Figure 9.

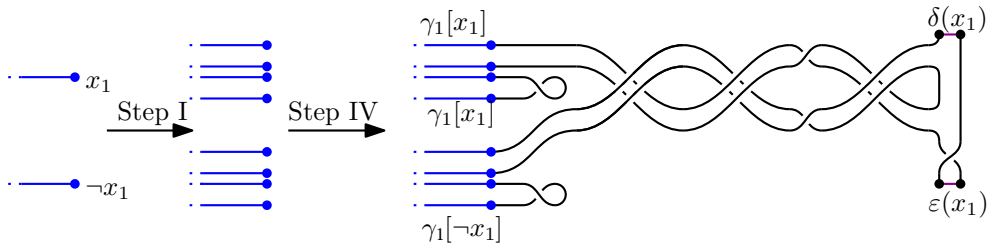

Figure 9: Step IV: Adding the variable gadgets.

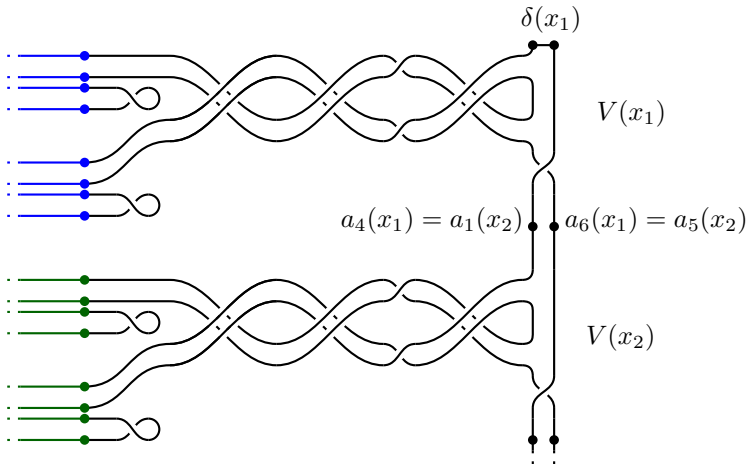

Figure 10: Step V: Interconnecting the variable gadgets.

Namely, to the bottom right endpoints of $\gamma_1[x_i], \gamma_2[x_i], \gamma_1[\neg x_i]$ and $\gamma_2[\neg x_i]$ we glue the parts of the variable gadget containing the crossings $p[x_i]$ and $p[\neg x_i]$ and to the top right endpoints of $\gamma_1[x_i], \gamma_2[x_i], \gamma_1[\neg x_i]$ and $\gamma_2[\neg x_i]$ we glue the remainder of the variable gadget. At this moment, we obtain a diagram of a link, where each link component has a diagram isotopic to the diagram on Figure 3.

Step V (interconnecting the variable gadgets): Finally, we form a connected sum of individual components. Namely, for every $i \in [n-1]$ we perform the knot sum along the arcs $\delta(x_i)$ and $\varepsilon(x_{i+1})$ by removing them and identifying $a_4(x_i)$ with $a_1(x_{i+1})$ and $a_6(x_i)$ with $a_5(x_{i+1})$ as on Figure 10. The arcs $\delta(x_1)$ and $\varepsilon(x_n)$ remain untouched. This way we obtain the desired unknot diagram $D(\Phi)$; see Figure 11.

The core of the **NP**-hardness reduction is the following theorem.

Theorem 10. *Let Φ be a formula in conjunctive normal form with n variables satisfying the conditions in the statement of Lemma 5. Then* $\mathrm{def}(D(\Phi)) \leq n$ *if and only if Φ is satisfiable.*

Theorem 1 immediately follows from Theorem 10 and Lemma 5:

Proof of Theorem 1 modulo Theorem 10. Due to the definition of the defect, the minimum number of Reidemeister moves required to untangle D equals $\frac{1}{2}(\mathrm{def}(D(\Phi)) + \mathrm{cross}(D(\Phi)))$. Therefore, setting $k = \frac{1}{2}(n + \mathrm{cross}(D(\Phi))$, Theorem 10 gives that $D(\Phi)$ can be untangled

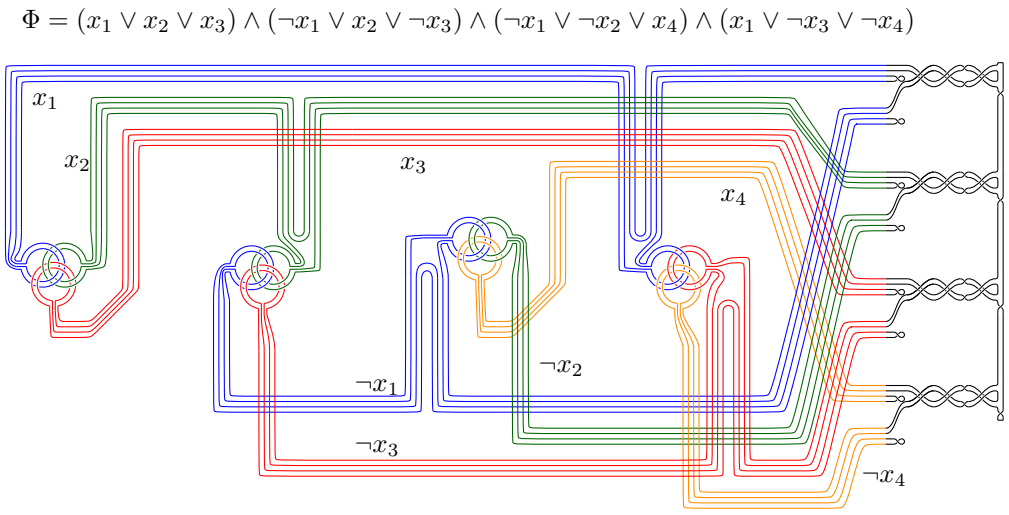

Figure 11: The final construction for the formula $\Phi = (x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee x_2 \vee \neg x_3) \wedge (\neg x_1 \vee \neg x_2 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee \neg x_4)$. For simplicity of the picture, we do not visualize how the crossings are resolved in Step III. (Unfortunately, we cannot avoid tiny pictures of gadgets.)

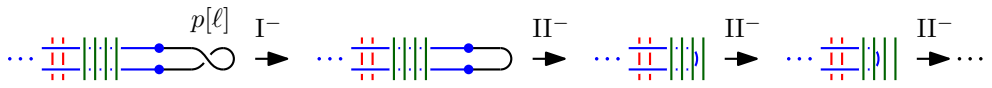

Figure 12: Initial simplifications.

with at most k moves if and only if Φ is satisfiable. This gives the required **NP**-hardness via Lemma 5. (Note also, that $D(\Phi)$ and k can be constructed in polynomial time in the size of Φ.) □

The remainder of this section is devoted to the proof of Theorem 10.

5.1 Satisfiable implies small defect

The purpose of this subsection is to prove the 'if' implication of Theorem 10. That is we are given a satisfiable Φ and we aim to show that $\mathrm{def}(D(\Phi)) \leq n$.

Let us consider a satisfying assignment. For any literal ℓ assigned TRUE we first remove the loop at the $p[\ell]$ vertex in the variable gadget (see Figure 3) by a I$^-$ move. This way, we use one I$^-$ move on each variable gadget, that is n such moves. Next we aim to show that it is possible to finish the untangling of the diagram by II$^-$ moves only. As soon as we do this, we get an untangling with defect n by Lemma 8 which will finish the proof.

Thus it remains to finish the untangling with II$^-$ moves only. We again pick ℓ assigned TRUE and we start shrinking the ℓ tentacle by II$^-$ moves. This way we completely shrink $\gamma_1[\ell]$ and $\gamma_2[\ell]$ as due to the construction as all arcs that meet $\gamma_1[\ell]$ simultaneously meet $\gamma_2[\ell]$ and vice versa. See Figure 12 for the initial I$^-$ move and a few initial II$^-$ moves. Furthermore we can continue shrinking the ℓ tentacle until we get a loop next to the $q[\ell]$ vertex; see Figure 13.

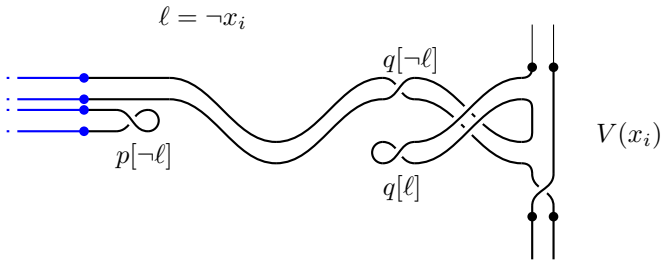

Figure 13: The ℓ tentacle was shrunk to a loop next to $q[\ell]$. In this example we have $\ell = \neg x_i$.

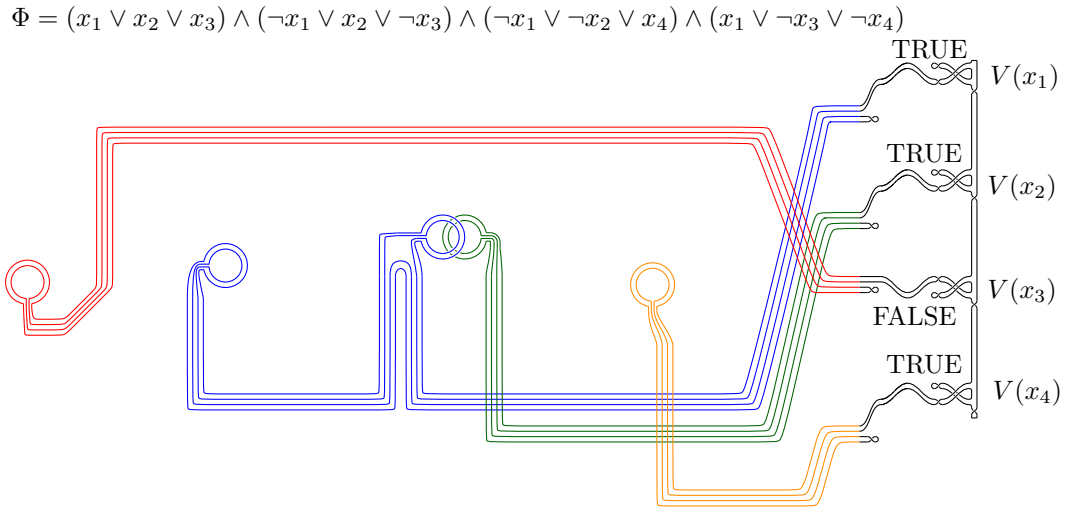

Figure 14: Simplifying the tentacles according to a satisfying assignment $x_1 = x_2 = x_4 =$ TRUE, $x_3 =$ FALSE.

We continue the same process for every literal ℓ assigned TRUE. In the intermediate steps, some of the other arcs meeting $\gamma_1[\ell]$ and $\gamma_2[\ell]$ might have already been removed. However, it is still possible to simplify the ℓ tentacle as before. See Figure 14 for the result after shrinking all tentacles assigned TRUE.

Because we assume that we started with a satisfying assignment, in each clause gadget at least one ring among the three Borromean rings disappears. Consequently, if there are two remaining rings in some clause gadget, then they can be pulled apart from each other by II^- moves as in Figure 15.

After this step, for each ℓ assigned TRUE, the $\gamma_1[\neg\ell]$ and $\gamma_2[\neg\ell]$ form 'fingers' of four parallel curves. These fingers can be further simplified by II^- moves so that any crossings among different fingers are removed; see Figure 16. For each variable gadget $V(x)$ we get one of the two possible pictures at Figure 17 left. Both of them simplify to the picture on the right by three further II^- moves.

Finally, we recall how the variable gadgets are interconnected (compare the right picture at Figure 17 with Figure 10). Then it is easy to remove all remaining $2n$ crossings by II^- moves gradually from top to bottom. This finishes the proof of the 'if' part of Theorem 10.

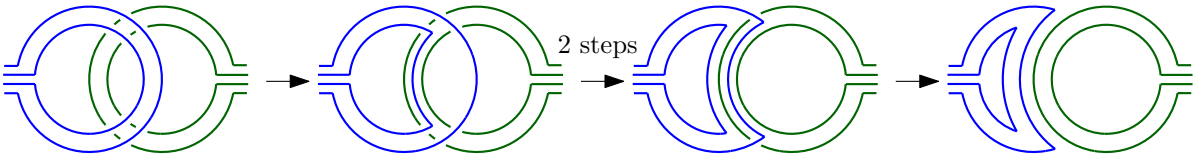

Figure 15: Untangling two rings in the clause gadget via II$^-$ moves.

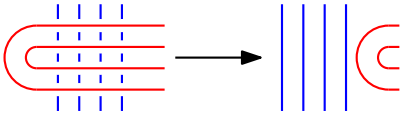

Figure 16: Simplifying $\gamma_1[\neg\ell]$ and $\gamma_2[\neg\ell]$ via II$^-$ moves. First, we untangle the inner (horizontal) 'finger' and then we untangle the outer (horizontal) 'finger'.

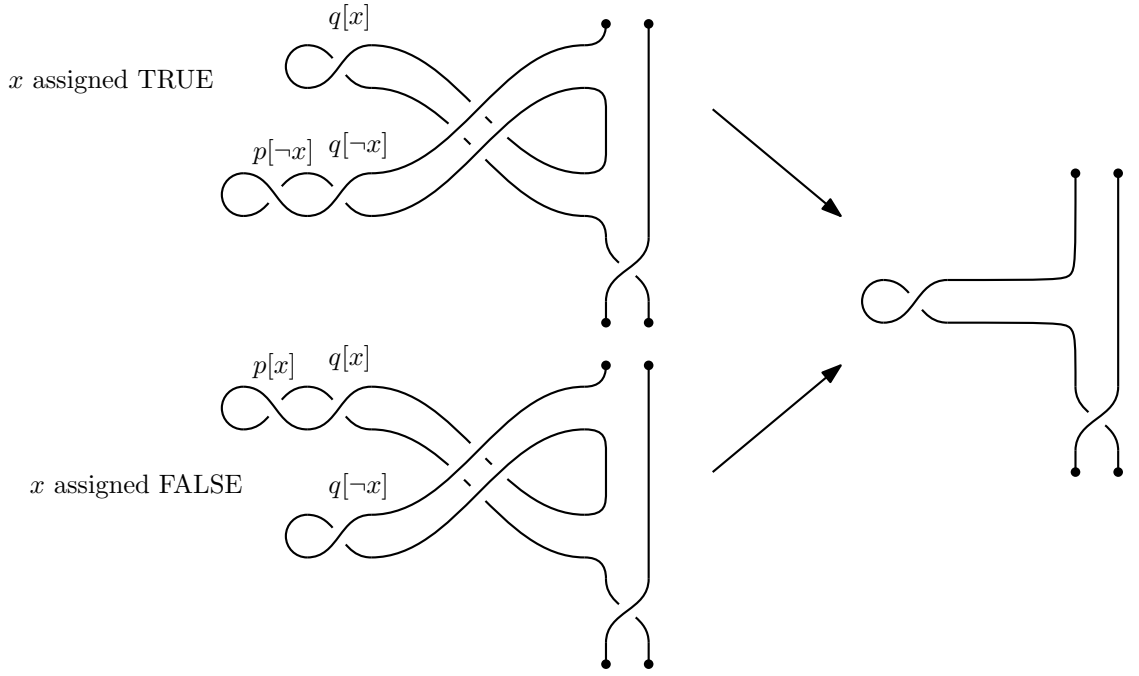

Figure 17: Results of the simplifications on the previous picture on the level of variable gadgets.

5.2 Small defect implies satisfiable

The purpose of this subsection is to prove the 'only if' part of the statement of Theorem 10. Recall that this means that we assume $\mathrm{def}(D(\Phi)) \leq n$ and we want to deduce that Φ is satisfiable. (Along the way we will actually also deduce that $\mathrm{def}(D(\Phi)) = n$.) In this subsection, we heavily use the terminology introduced in Section 4.

Let $\mathbf{D} = (D^0, \ldots, D^k)$ be an untangling of D with $\mathrm{def}(\mathbf{D}) \leq n$. For a variable x let $R(x) = \{p[x], p[\neg x], q[x], q[\neg x], s_1(x), \ldots, s_{12}(x)\}$ be the set of 16 out of the 17 self-crossings in the variable gadget $V(x)$ (we leave out $r(x)$) and let the *weight of* x, denoted by $w(x)$, be the sum of weights of the crossings in $R(x)$.

Our first aim is to analyze the first economical II^- move that removes some of the crossings in $R(x)$, using Lemma 9.

Claim 10.1. *Let x be a variable with $w(x) \leq 1$. Let r be the first crossing in $R(x)$ which is removed by an economical II^- move (we allow a draw). Then one of the following cases holds:*

(i) $\{r, t(r)\} = \{s_1(x), s_2(x)\}$, $w(p[x]) = w(x) = 1$ *and $p[x]$ is removed by a I^- move prior to removing r and $t(r)$.*

(ii) $\{r, t(r)\} = \{s_1(x), s_3(x)\}$, $w(p[\neg x]) = w(x) = 1$ *and $p[\neg x]$ is removed by a I^- move prior to removing r and $t(r)$.*

Before we start the proof, we remark that the condition $w(x) \leq 1$ implies that there are (at least 15) crossings in $R(x)$ removed by economical II^- moves. In particular, r in the statement always exists.

Proof. First, we need to identify the possible pairs $\{r, t(r)\}$ where $r \in R(x)$. Such pairs are found by a case analysis, using Lemma 9 with $c = 1$ and $R = R(x)$. We emphasize that here r is a variable, not to be confused with the crossings labelled $r(x)$, for which we have $r(x) \notin R(x)$ (by definition of $R(x)$).

The general strategy is the following. For each element of $R(x)$ we consider whether it may be r. We analyze possible arcs α and β from the definition of c-close neighbors.

There are two directions in which α may emanate from r. For each direction we allow up to one internal crossing from $R(x)$ on α, getting a candidate position for $t(r)$ (even if α passes through other variable gadgets we count only the crossings from $R(x)$). We immediately disregard the cases when α passes through r again (this is not allowed by the third item of the definition of close neighbors). We also emphasize that we are interested only in the cases when α enters the candidate $t(r)$ as an overpass.

Next, we refocus on β; again there are two possible directions and we again identify possible the possible position of $t(r)$ (this time entered as an underpass).

Finally, we compare the lists of candidate positions for $t(r)$ obtained for α and for β; it must be possible to obtain $t(r)$ in both ways.

The candidate positions of $t(r)$ as an endpoint of α and as an endpoint of β are summarized in Table 1; and they can be easily found with the aid of Figure 3. Note that it follows from the construction of $D(\Phi)$ that $\delta(x)$ and $\varepsilon(x)$ are (usually) not in $D(\Phi)$. For considerations in the table, we denote by $\delta'(x)$ the arc in $D(\Phi)$ between the points $a_1(x)$ and $a_5(x)$ which avoids $r(x)$ (equivalently, any other crossing in $V(x)$). Similarly, we let $\varepsilon'(x)$ denote the arc

Choice of r	$t(r)$ in α	$t(r)$ in β	Overlap
$p[x]$	$\gamma_2[x]$	$\gamma_1[x], s_2(x), s_4(x)$	$\emptyset$
$p[\neg x]$	$\gamma_2[\neg x], s_3(x), s_4(x)$	$\gamma_1[\neg x]$	$\emptyset$
$s_1(x)$	$\gamma_1[\neg x], \gamma_2[\neg x], s_3(x), s_2(x)$	$\gamma_2[x], \gamma_1[x], s_2(x), s_3(x)$	$s_2(x), s_3(x)$
$s_2(x)$	$s_1(x), \gamma_1[\neg x]$	$\gamma_1[x], p[x], \gamma_2[x], s_1(x), s_4(x)$	$s_1(x)$
$s_3(x)$	$\gamma_2[\neg x], p[\neg x], \gamma_1[\neg x], s_1(x), s_4(x)$	$s_1(x), \gamma_2[x]$	$s_1(x)$
$s_4(x)$	$s_3(x), \gamma_2[\neg x], p[\neg x]$	$s_2(x), \gamma_1[x], p[x]$	$\emptyset$
$s_5(x)$	$s_6(x)$	$s_7(x)$	$\emptyset$
$s_6(x)$	$s_5(x)$	$s_8(x), q[\neg x]$	$\emptyset$
$s_7(x)$	$s_8(x), q[x]$	$s_5(x)$	$\emptyset$
$s_8(x)$	$s_7(x), q[x]$	$s_6(x), q[\neg x]$	$\emptyset$
$q[x]$	$s_7(x), s_8(x)$	$s_9(x), s_{11}(x)$	$\emptyset$
$q[\neg x]$	$s_9(x), s_{10}(x)$	$s_6(x), s_8(x)$	$\emptyset$
$s_9(x)$	separate analysis		
$s_{10}(x)$	$s_9(x), q[\neg x], \delta'(x), r(x), \varepsilon'(x)$	$s_{12}(x)$	$\emptyset$
$s_{11}(x)$	$s_{12}(x)$	$s_9(x), q[x], r(x), \varepsilon'(x), \delta'(x)$	$\emptyset$
$s_{12}(x)$	$s_{11}(x)$	$s_{10}(x)$	$\emptyset$

Table 1: Possible positions for $t(r)$ depending on r in Claim 10.1.

in $D(\Phi)$ between the points $a_4(x)$ and $a_6(x)$ which avoids $r(x)$. The table also misses values for $r = s_9(x)$ which requires a separate analysis.

In order to avoid any ambiguity, we explain how the first row of the table is obtained, considering the case $r = p[x]$ (follow Figure 3):

We know that α may emanate to the left or to the right. Emanating to the left is immediately ruled out as we reach $p[x]$ again in the next crossing. Emanating to the right allows $t(r)$ to be some crossing on $\gamma_2[x]$ or seemingly it may be $s_1(x)$ or $s_3(x)$; however $s_1(x)$ and $s_3(x)$ are entered by α as an underpass, so they cannot be $t(r)$. Therefore the only option, from the point of view of α, is that $t(r)$ belongs to $\gamma_2[x]$, as marked in Table 1. Similarly when focusing on β, emanating to the left is immediately ruled out whereas emanating to the right allows $t(r)$ to belong $\gamma_1[x]$ or to be $s_2(x)$ or $s_4(x)$ (as in the table). We conclude that for $r = p[x]$ there is no $t(r)$ suitable both for α and β, using the fact that $\gamma_1[x]$ and $\gamma_2[x]$ do not intersect (in the whole $D(\Phi)$). (This is marked by the $\emptyset$ in the overlap column.) We deduce that r cannot be $p[x]$.

In general, for identifying the overlaps for other options of r, we use that no two of the arcs $\gamma_1[x], \gamma_2[x], \gamma_1[\neg x], \gamma_2[\neg x]$ intersect.

Next, we want to rule out the case $r = s_9(x)$ because this is not covered by Table 1. Considering the possible arcs α, we get the following options for $t(r)$: $q[\neg x]$, $s_{10}(x)$, $\delta'(x)$, $r(x)$, $\varepsilon'(x)$, and from the point of view of β, we have the following options for $t(r)$: $q[x]$, $s_{11}(x)$, $r(x)$, $\varepsilon'(x)$, $\delta'(x)$. Therefore, there are $r(x)$, $\delta'(x)$ and $\epsilon'(x)$ in the overlap (in addition $\delta'(x)$ and $\epsilon'(x)$ may intersect). However, if we want to reach $r(x)$, $\delta'(x)$ or $\epsilon'(x)$ with α we have to pass through $s_{10}(x)$. Similarly, β has to pass through $s_{11}(x)$. But this violates the condition that α and β together have at most $c = 1$ point of $R(x)$ in their interiors.

By checking Table 1 and by the paragraph above we deduce that the only two options for $\{r, t(r)\}$ are $\{s_1(x), s_2(x)\}$ and $\{s_1(x), s_3(x)\}$. By checking possible α and β we get:

- $\{r, t(r)\} = \{s_1(x), s_2(x)\}$, α is the arc directly connecting $s_1(x)$ and $s_2(x)$ containing

no other crossings and β is the arc connecting $s_1(x)$ and $s_2(x)$ passing (twice) through $p[x]$; or

- $\{r, t(r)\} = \{s_1(x), s_3(x)\}$, α is the arc connecting $s_1(x)$ and $s_3(x)$ passing (twice) through $p[\neg x]$ and β is the arc directly connecting $s_1(x)$ and $s_3(x)$ containing no other crossings.

Let us first focus on the first case above. Before removing $s_1(x)$ and $s_2(x)$ the crossing $p[x]$ has to be removed from the arc β. This can be done, in principle, by a I^- move, a wasteful II^- move, or $p[x]$ can be swapped with $s_1(x)$ or $s_2(x)$ by a III move before removing $s_1(x)$ and $s_2(x)$.

Removing $p[x]$ by a I^- move is the desired conclusion (i) given that in this case we have $w(p[x]) \geq 1$ (because of the I^- move) as well as $w(p[x]) \leq w(x) \leq 1$ (from the assumptions of the claim); therefore $w(p[x]) = w(x) = 1$ as required.

Removing $p[x]$ by a wasteful II^- move is impossible as we would have $w(p[x]) \geq 2$ whereas $w(x) \leq 1$ by the assumption of the claim.

Similarly, swapping $p[x]$ with $s_1(x)$ or $s_2(x)$ is impossible as we would have $w(p[x]) + w(s_1(x)) + w(s_2(x)) \geq 2 \cdot \frac{2}{3}$ but $w(x) \leq 1$ again.

Conclusion (ii) follows analogously from the second case. □

Now, let us set $\ell := \ell(x) := x$ if the conclusion (i) of Claim 10.1 holds and $\ell := \ell(x) := \neg x$ if the conclusion (ii) holds (assuming $w(x) \leq 1$). (We identify $\neg\neg x$ with x, that is, if $\ell = \neg x$, then $\neg\ell = x$.)

Claim 10.2. *If $w(x) \leq 1$, then $p[\neg\ell]$ and $q[\neg\ell]$ are twins. In addition, the preimage arcs α and β between $p[\neg\ell]$ and $q[\neg\ell]$ contain $\gamma_1[\neg\ell]$ and $\gamma_2[\neg\ell]$.*

Proof. Let us set $R'(x) := \{p[\neg\ell], q[\ell], q[\neg\ell], s_9(x), s_{10}(x), s_{11}(x), s_{12}(x)\}$. Intuitively, the crossings of $R'(x)$ are those crossings of $R(x)$ that do not meet the arc from $q[\ell]$ to $q[\ell]$ containing $p[\ell]$; see Figure 18.

All $r' \in R'(x)$ have $w(r') = 0$ as $w(x) = w(p[\ell]) = 1$ by Claim 10.1. In particular, all such r' are removed by an economical II^- move. Let r be the first of these elements removed by an economical II^- move. Let α and β be the arcs from Lemma 9 with $R = R'$ and $c = 0$.

Now we perform a similar inspection as in the proof of Claim 10.1. This time $c = 0$, thus we do not allow any internal crossing on arcs α and β. Most of the cases are straightforward and we refer to Table 2 for the possible α and β; $\delta'(x)$ and $\varepsilon'(x)$ play the same role as in the proof of the previous claim. The only exception is that we need to rule out the case $r = q[\ell]$ separately:

Let us therefore assume that $r = q[\ell]$. First we want to observe that both α and β emanate from $q[\neg\ell]$ to the left. Indeed, if α emanates to the right, then necessarily $\ell = \neg x$ and $t(r) = s_9(x)$ but we do not have a suitable β for this case. Similarly, if β emanates to the right, then $\ell = x$ and $t(r) = s_9(x)$ but we do not have a suitable α.

Now we know that both α and β emanate to the left. This means that both α and β are subarcs of the arc with both endpoints $q[\ell]$ left from $q[\ell]$ (if $\ell = x$, this is the grey arc on the Figure 18). In particular, $t(r)$ has to be a selfcrossing of this arc, and there is only one option, namely $t(r) = p[\ell]$. However, here we crucially use Claim 10.1, the twin of $q[\ell]$ cannot be $p[\ell]$ as $p[\ell]$ is removed by a I^- move. This rules out the case $r = q[\ell]$.

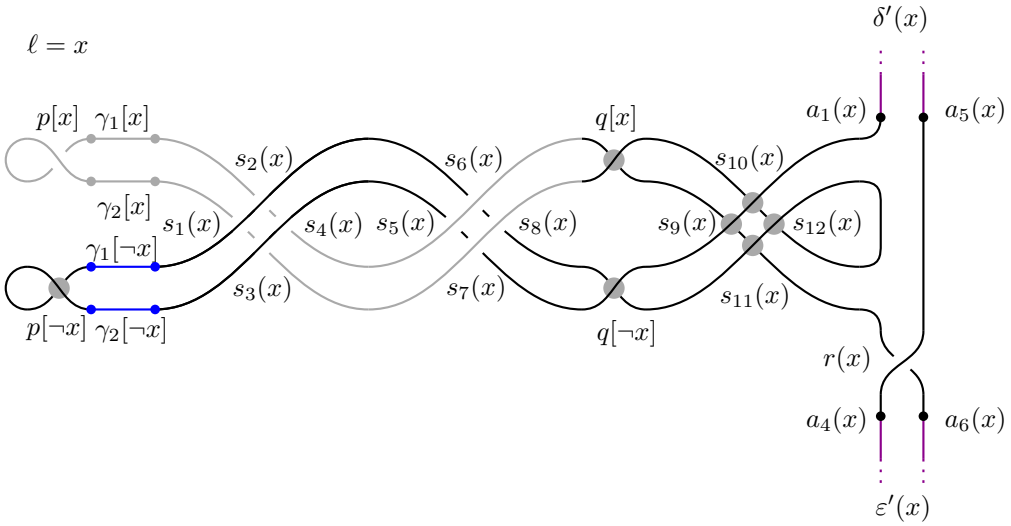

Figure 18: The crossings of $R'(x)$ in the variable gadget in case $\ell = x$.

Choice of r	$t(r)$ in α	$t(r)$ in β	Overlap
$p[\neg\ell]$ if $\ell = x$	$\gamma_2[\neg\ell], s_3(x), s_4(x), q[\neg\ell]$	$\gamma_1[\neg\ell], s_6(x), s_8(x), q[\neg\ell]$	$q[\neg\ell]$
$p[\neg\ell]$ if $\ell = \neg x$	$\gamma_2[\neg\ell], s_7(x), s_8(x), q[\neg\ell]$	$\gamma_1[\neg\ell], s_2(x), s_4(x), q[\neg\ell]$	$q[\neg\ell]$
$q[\ell]$	separate analysis		
$q[\neg\ell]$ if $\ell = x$	$s_4(x), s_3(x), \gamma_2[\neg\ell], p[\neg\ell], s_9(x)$	$s_8(x), s_6(x), \gamma_1[\neg\ell], p[\neg\ell]$	$p[\neg\ell]$
$q[\neg\ell]$ if $\ell = \neg x$	$s_8(x), s_7(x), \gamma_2[\neg\ell], p[\neg\ell]$	$s_4(x), s_2(x), \gamma_1[\neg\ell], p[\neg\ell], s_9(x)$	$p[\neg\ell]$
$s_9(x)$	$q[\neg x], s_{10}(x)$	$q[x], s_{11}(x)$	$\emptyset$
$s_{10}(x)$	$s_9(x), \delta(x), r(x), \epsilon(x)$	$s_{12}(x)$	$\emptyset$
$s_{11}(x)$	$s_{12}(x)$	$s_9(x), r(x), \epsilon(x), \delta(x)$	$\emptyset$
$s_{12}(x)$	$s_{11}(x)$	$s_{10}(x)$	$\emptyset$

Table 2: Possible positions for $t(r)$ depending on r in Claim 10.2.

Therefore, it follows from Table 2 that $\{r, t(r)\} = \{p[\neg\ell], q[\neg\ell]\}$. In addition, a further inspection of the variable gadget also reveals that α contains $\gamma_2[\neg\ell]$ and β contains $\gamma_1[\neg\ell]$ as desired. □

Now, we have acquired enough tools to finish the proof of the theorem. By Claim 10.1, we have $w(x) \geq 1$ for any variable x. By Lemma 7, we deduce

$$\mathrm{def}(\mathbf{D}) \geq \sum_x w(x) \geq n,$$

where the sum is over all variables. On the other hand, we assume $\mathrm{def}(\mathbf{D}) \leq n$. Therefore both inequalities above have to be equalities and in particular $w(x) = 1$ for any variable x. In particular, the assumptions of Claims 10.1 and 10.2 are satisfied for any variable x.

Given a variable x, we assign x with TRUE if the conclusion (i) of Claim 10.1 holds (that is, if $x = \ell(x)$). Otherwise, if the conclusion (ii) of Claim 10.1 holds (i.e. $\neg x = \ell(x)$), we set x to FALSE. It remains to prove that we get a satisfying assignment this way.

For contradiction, suppose there is a clause $c = (\ell_1 \vee \ell_2 \vee \ell_3)$ which is not satisfied with this assignment. Let x_i be the variable of ℓ_i, that is, $\ell_i = x_i$ or $\ell_i = \neg x_i$. The fact that c is not satisfied with the assignment above translates as $\ell(x_i) = \neg\ell_i$ for any $i \in \{1, 2, 3\}$.

By Claim 10.2, we get that $p[\ell_i]$ and $q[\ell_i]$ are twins for any $i \in \{1, 2, 3\}$. Let $R''(c)$ be the set of crossings in these Borromeans union the sets $\{p[\ell_i], q[\ell_i]\}$ for $i \in \{1, 2, 3\}$. All the crossings in $R''(c)$ have weight 0 and they have to be removed by economical II^- moves as all defect is realized on points $p[\ell(x)]$ for all variables x (but $p[\ell_i] = p[\neg\ell(x_i)]$ are not among these points).

Let r be the first removed crossing among the crossings in $R''(c)$. First, we observe that r cannot be any of $p[\ell_i]$ or $q[\ell_i]$ for $i \in \{1, 2, 3\}$. This follows from Claim 10.2 as the arcs $\gamma_1[\ell_i]$ and $\gamma_2[\ell_i]$ contain some crossings in $R''(c)$.

Next we apply Lemma 9 (for 0-close neighbors) with $R = R''(c)$. By symmetry of the clause gadget, it is sufficient to consider the cases that r is one of the crossings $u_1, \ldots, u_8$ on Figure 19 between the rings for ℓ_1 and ℓ_2.

Let α and β be the arcs between r and $t(r)$ from the definition of 0-close neighbors. We can immediately rule out $r \in \{u_4, u_5, u_6, u_7, u_8\}$ by an easy inspection as in Claim 10.2 (this is easy because we always hit a crossing from $R''(c)$ by possible α and β). Therefore it remains to consider the case $r \in \{u_1, u_2, u_3\}$.

Now let us consider the case $r = u_3$. The only option for α is to emanate to the left reaching the crossing u_1 as emanating to the right reaches a crossing from $R''(c)$ as an underpass. Consequently, β has to emanate to the right since emanating to the left would reach u_4. However, before β reaches u_1, it has to pass through $p[\ell_2]$ or $q[\ell_2]$ which rules out this option.

The case $r = u_2$ is ruled out analogously.

It remains to consider the case $r = u_1$. We have already ruled out the case that the twin $t(r)$ would be u_2 or u_3 (it is sufficient to swap r and $t(r)$ in the previous considerations). Thus α has to emanate to the left from u_1 whereas β has to emanate to the right. The first point of $R''(c)$ that α reaches is $p[\ell_1]$ or $q[\ell_1]$ while the first point of $R''(c)$ that β reaches is $p[\ell_2]$ or $q[\ell_2]$. Therefore $t(r)$ must be some crossing of $\gamma_1[\ell_1]$ and $\gamma_1[\ell_2]$ while α is a subarc of $\gamma_1[\ell_1]$ and β is a subarc of $\gamma_1[\ell_2]$. On the one hand, such crossing may exist. On the other hand, α reaches such a crossing always as an underpass and β as an overpass due to our convention in Step III of the construction of $D(\Phi)$. Therefore, we do not get admissible α and β. This

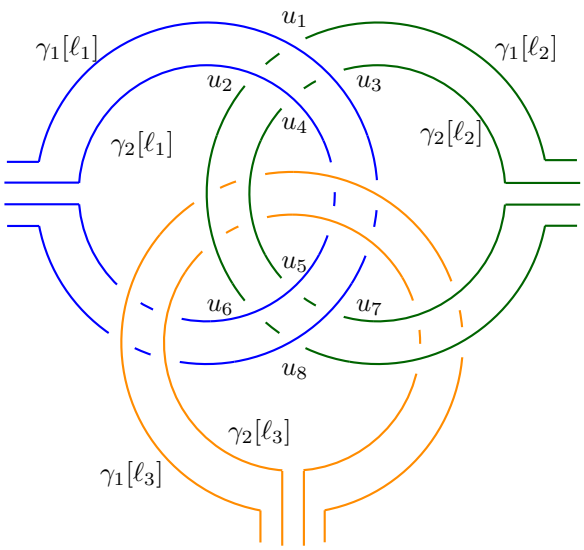

Figure 19: The clause gadget with crossings $u_1, \ldots, u_8$.

contradicts the existence of r. Therefore, the suggested assignment is satisfying. This finishes the proof.

Part III
Hard link invariants

6 Intermediate invariants

In this section we describe the family of link invariants from the statement of Theorem 2. The material presented here is standard and details can be found in various textbooks; since every piecewise linear knot can be smoothed in a unique way, we assume that the knots discussed are smooth. In Part III we work in the smooth category. We first define:

Definition 11. Let L be a link in the 3-sphere. We now give a list of the invariants that we will be using; for a detailed discussion see, for example, [Rol90].

1. A *smooth slice surface* for L is an orientable surface with no closed components, properly and smoothly embedded in the 4-ball, whose boundary is L. Recall that every link bounds an orientable surface in S^3 (a Seifert surface); by pushing the interior of that Seifert surface into the 4-ball we see that every link bounds a smooth slice surface.

2. The *4-ball Euler characteristic* of L, denoted $\chi_4(L)$, is the largest integer so that L bounds a smooth slice surface of Euler characteristic $\chi_4(L)$. Since a smooth slice surface has no closed components, its Euler characteristic is at most the number of components of L; in particular, $\chi_4(L)$ exists.

3. A link is called *smoothly slice* if it bounds a slice surface that consists entirely of disks;

equivalently, the $\chi_4(L)$ equals the number of components of L. Note that unlinks are smoothly slice (but not only unlinks).

4. The *unlinking number*, denoted $u(L)$, is the smallest nonnegative integer so that L admits some diagram D so that after $u(L)$ crossing changes on D a trivial link is obtained.

5. The *ribbon number*, denoted $u_r(L)$, is the smallest nonnegative integer so that L admits some diagram D so that after $u_r(L)$ crossing changes on D a ribbon link is obtained (see [Rol90] for the definition of ribbon link).

6. The *slicing number*, denoted $u_s(L)$, is the smallest nonnegative integer so that L admits some diagram D so that after $u_s(L)$ crossing changes on D a smoothly slice link is obtained.

7. Links $L_0, L_1 \subset S^3$ are called *concordant* if there exists a smooth embedding $f : L_0 \times [0,1] \to S^n \times [0,1]$, so that $f(L_0 \times \{0\}) = L_0 \times \{0\}$ and $f(L_0 \times \{1\}) = L_1 \times \{1\}$. Equivalently, there exists a union of smooth annuli A, properly embedded in $S^3 \times [0,1]$, so that $A \cap (S^3 \times \{0\}) = L_0 \times \{0\}$ and $A \cap (S^3 \times \{1\}) = L_1 \times \{1\}$.

8. The *concordance unlinking number*, denoted $u_c(L)$, is the minimum of the unlinking number over the concordance class of L.

9. The *concordance ribbon number*, denoted $u_{cr}(L)$, is the minimum of the ribbon number over the concordance class of L.

10. The *concordance slicing number*, denoted $u_{cs}(L)$, is the minimum of the slicing number over the concordance class of L.

11. By transversality, every link L bounds smoothly immersed disks in B^4 with finitely many double points. The 4-*dimensional clasp number* (sometimes called the 4-*ball crossing number*) of L, denoted $c_s(L)$, is the minimal number of double points for such disks.

Finally, we define intermediate invariants, whose existence follows from Theorem 2 of [Shi74].

Definition 12 (intermediate invariant)**.** A real valued link invariant $i(L)$ is called an *intermediate invariant* if

$$u(L) \geq i(L) \geq c_s(L).$$

Many invariants are known to be intermediate (see, for example, [Shi74]). We list a few here:

Lemma 13. *The invariants u, u_r, u_s, u_c, u_{cr}, u_{cs} and c_s are all intermediate.*

Proof. It is well known that the unlink is ribbon and a ribbon link is slice, and therefore

$$u \geq u_r \geq u_s \text{ and } u_c \geq u_{cr} \geq u_{cs}$$

Since any link is in its own concordance class we have

$$u_s \geq u_{cs} \text{ and } u \geq u_c$$

Combining these we see that

$$u \geq u_r \geq u_s \geq u_{cs} \text{ and } u \geq u_c \geq u_{cr} \geq u_{cs}$$

Therefore it suffices to show that $u_{cs} \geq c_s$. To see this, decompose B^4 (the unit ball in $\mathbb{R}^4$) as $B^4 = X_1 \cup X_2 \cup X_3$ where here:

$$X_1 = \left\{ \vec{x} \in \mathbb{R}^4 \;\middle|\; 0.6 < |\vec{x}| \leq 1 \right\} \; (\cong S^3 \times (0,1])$$

$$X_2 = \left\{ \vec{x} \in \mathbb{R}^4 \;\middle|\; 0.3 < |\vec{x}| < 0.7 \right\} \; (\cong S^3 \times (0,1))$$

and

$$X_3 = \left\{ \vec{x} \in \mathbb{R}^4 \;\middle|\; |\vec{x}| < 0.4 \right\} \; (\cong \text{int} B^4)$$

We used open intersections to guarantee smoothness.

Let L' be a link in the concordance class of L that minimizes u_s, that is, $u_s(L') = u_{cs}(L)$. Since L and L' are concordant there exists disjoint annuli A, smoothly embedded in $S^3 \times [0,1]$, so that $A \cap (S^3 \times \{0\}) = L \times \{0\}$ and $A \cap (S^3 \times \{1\}) = L' \times \{1\}$. Since $X_1 \cong S^3 \times (0,1]$, there exist disjoint annuli A_1, smoothly embedded in X_1, so that $A_1 \cap \partial X_1 = L$ (note that $\partial X_1 = S^3$) and

$$A_1 \cap (S^3 \times (0.6, 0.7)) = L' \times (0.6, 0.7)$$

Suppose that L' has μ components and let k denote $u_s(L')$. Let L'' be a slice link obtained from L' after k crossing changes, and let S denote the disjoint union of μ circles (so L' is an embedding of S into S^3). Then there is a smooth homotopy $F : S \times (0.3, 0.7) \to S^3$ realizing k crossing changes, that is:

(i) $F(\cdot, t) = L'$ for $0.6 < t < 0.7$.

(ii) $F(\cdot, t) = L''$ for $0.3 < t < 0.4$.

(iii) There are k values $0.4 < t_1 < \cdots < t_k < .06$ so that $F(\cdot, t_i)$ has exactly one transverse double point.

(iv) For any other value t we have that $F(\cdot, t)$ is a smooth embedding of S.

Define $\widehat{F} : S \times (0.3, 0.7) \to X_2$ by

$$\widehat{F}(p, t) = (F(p, t), t)$$

Denote the image of $\widehat{F}$ by A_2. Then A_2 are μ smoothly immersed annuli with exactly k transverse double points. Note that

$$A_1 \cap (X_1 \cap X_2) = A_2 \cap (X_1 \cap X_2)$$

Since L'' is a smoothly slice link with μ components it bounds μ smooth disks disjointly embedded in B^4. Since $X_3 \cong \text{int} B^4$, this induces a smooth embedding $D \subset X_3$, where here D are μ disks, so that $D \cap (S^3 \times (0.3, 0.4)) = L'' \times (0.3, 0.4)$. Note that

$$A_2 \cap (X_2 \cap X_3) = D \cap (X_2 \cap X_3)$$

It is now clear that

$$A_1 \cup A_2 \cup D$$

is a smooth immersion of μ disks with exactly k double points, showing that

$$u_{cs}(L) \leq c_s(L)$$

□

Finally we prove:

Lemma 14. *Let L be a link with μ components. Then*

$$\chi_4(L) \geq \mu - 2c_s(L).$$

Sketch of proof. This was shown by Shibuya [Shi74]; for the convenience of the reader we sketch the proof here. Given μ disks with $c_s(L)$ double points, we endow each disk with an orientation. One can replace a neighborhood of a double point with an annulus in a way that agrees with the chosen orientation. The result is a smooth orientable surface F whose Euler characteristic is $\mu - 2c_s(L)$. □

7 A certain signature calculation.

In this section we calculate the signature of certain links; this will be used in the next section. The signature is an integer valued link invariant, defined for knots by Trotter in [Tro62] and generalized for links by Murasugi in [Mur65].

Since the signature is covered in many standard texts about knots and links we will only summarize how to calculate it:

1. Given a link a L, first construct a Seifert surface for L, that is, an embedded, orientable surface F, with no closed components, whose boundary is L. F need not be connected.

2. Next construct embedded oriented curves $\{a_i\}$ on F that form a basis for $H_1(F;\mathbb{Z})$.

3. Arbitrarily fix a co-orientation on each component of F (that is, directions "above" and "below" F). We define a_i^+ and a_i^- to be parallel copies of a_i pushed slightly above and below F.

4. The *Seifert matrix* M is the square matrix whose ijth entry is $\mathrm{lk}(a_i^+, a_j^-)$, the linking number of a_i^+ and a_j^-. Note that although the linking number is a symmetric function, the matrix M need not be symmetric. Note also that whenever $a_i \cap a_j = \emptyset$, the ijth entry of M is simply $\mathrm{lk}(a_i, a_j)$.

5. The *signature* of L, denoted $\sigma(L)$, is the signature of the symmetric bilinear form on $H_1(F;\mathbb{Z})$ defined by $M + M^T$; explicitly, it is the number of positive entries minus the number of negative entries after diagonalizing $M + M^T$.

It is quite surprising that $\sigma(L)$ is a link invariant, that is, independent of the choices made. Nevertheless, it is known to be an invariant and a very useful one at that, as we shall see in the next section.

We now define:

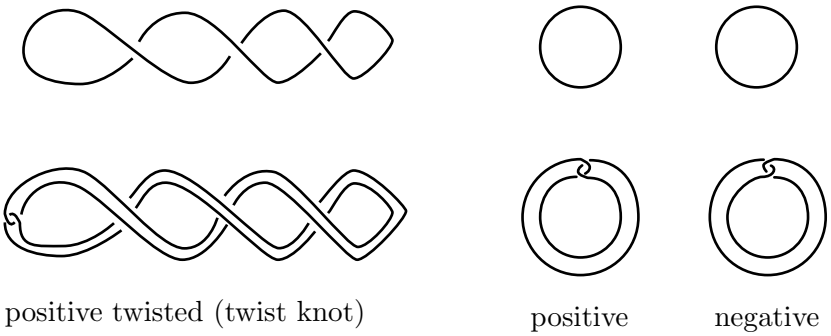

Figure 20: Some Whitehead doubles

Definition 15 (Whitehead Double)**.** Let L be a link. A *Whitehead double* of L is a link obtained by taking two parallel copies of each component of L and joining them together with a clasp (see Figure 20). A Whitehead double is called *positive* if the crossings at the clasp are positive. If the linking number of the two copies of each component is zero the Whitehead double is called *untwisted*. It is easy to see that the untwisted positive Whitehead double is uniquely determined by L.

Lemma 16. *The signature of the positive untwisted Whitehead double of the Hopf link is* 2.

Proof. An untwisted Whitehead double of the Hopf link bounds F, two disjointly embedded once punctured tori (see Figure 21). In that figure each torus is seen as a "flat" annulus with a twisted band attached near the top. The tori are co-oriented to the positive side above the "flat" annulus. In that figure we marked a_1, a_2, a_3, a_4, ordered generators for $H_1(F, \mathbb{Z})$. This gives the Seifert matrix:

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} \tag{1}$$

Symmetrizing M we get

$$M + M^T = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 2 & 1 & 0 \end{pmatrix}$$

A straightforward calculation shows that the signature is 2. □

Before stating the next lemma we describe the type of link we will be dealing with. Let L be a link with an even number of components, say $2n$, so that L can be written as $L = L_1 \cup \cdots \cup L_n$ satisfying the following conditions:

1. $(\forall i)$ L_i is the positive untwisted Whitehead double of the Hopf link.
2. Each L_i bounds F_i, a co-oriented disjoint union of two once punctured tori, as in Figure 21. Let $a_1^{(i)}, a_2^{(i)}, a_3^{(i)}, a_4^{(i)}$ denote the generators for $H_1(F_i; \mathbb{Z})$, again as in the figure.
3. $(\forall i \neq i')$ $F_i \cap F_{i'} = \emptyset$.

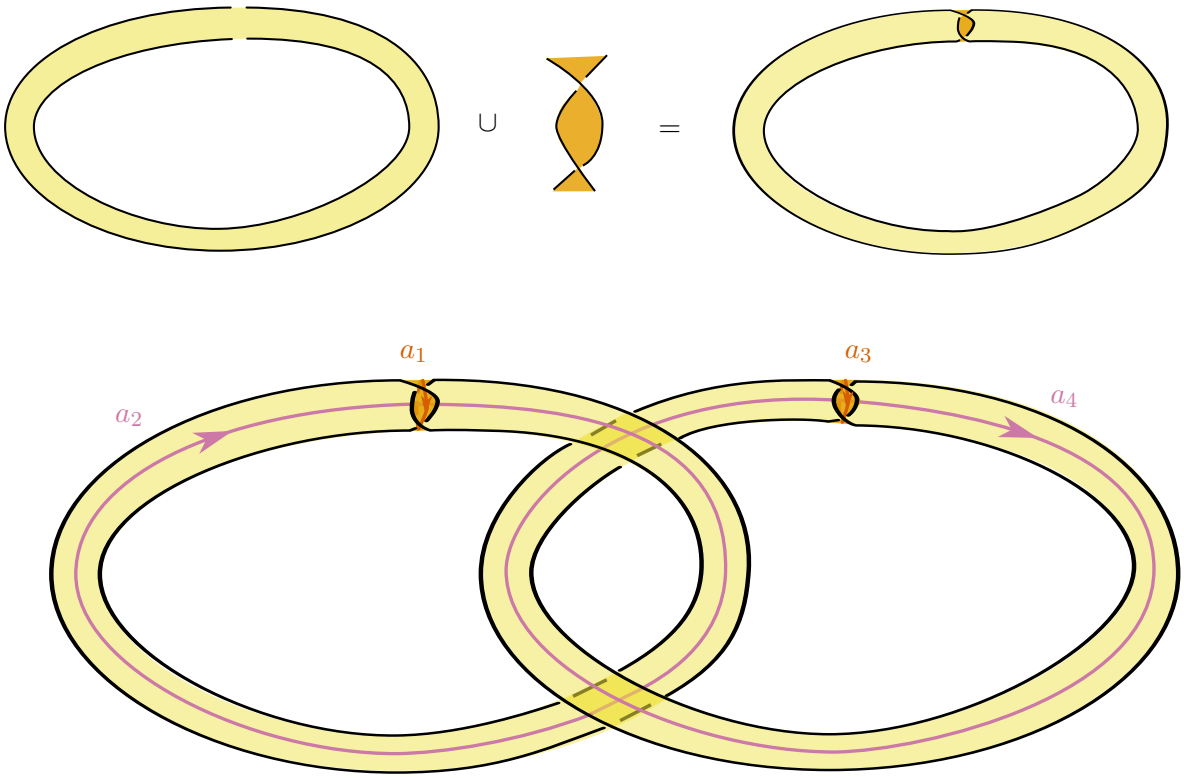

Figure 21: Top: the untwisted Whitehead double of an unknot bounds a flat annulus with a twisted band. Bottom: the positive untwisted Whitehead double of the Hopf link, a Seifert surface and generators of its homology.

4. $(\forall i \neq i')(\forall j, j')\ \ \mathrm{lk}(a_j^{(i)}, a_{j'}^{(i')}) = 0.$

For a link L as above we have:

Lemma 17. $\sigma(L) = 2n.$

Proof. With the conventions above we have that $F = F_1 \cup \cdots \cup F_n$ is an oriented Seifert surface for L. We will use

$$a_1^{(1)}, a_2^{(1)}, a_3^{(1)}, a_4^{(1)}, \ldots, a_3^{(n)}, a_4^{(n)}$$

as an ordered set of generators for $H_1(F;\mathbb{Z})$. We obtain a Seifert matrix that along the diagonal has 4×4 *identical* blocks, each identical to the Seifert matrix of the positive untwisted Whitehead double of the Hopf link (given in the proof of Lemma (16)). For generators corresponding to $i \neq i'$ we have, by assumption, that $\mathrm{lk}(a_i^{(j)}, a_{i'}^{(j')}) = 0$. Thus all the remaining entries are zero. This shows that the signature is the sum of the signatures of the 4×4 blocks, and the lemma follows from Lemma 16. □

8 Unlinking, 4-ball Euler characteristic, and intermediate invariants.

In this section we will show that several link invariants are **NP**-hard (see the invariants defined in Definition 11). Recall the definition of the positive untwisted Whitehead double (Definition 15 and Figures 20 and 21). One last piece of background we will need is a result of A. Levine [Lev12]. Theorem 1.1 of [Lev12] implies, in particular:

Lemma 18. *The untwisted positive Whitehead double of the Hopf link, and that of the Borromean rings, are not smoothly slice.*

We are now ready to describe our construction:

The construction of $L_\Phi^{\mathbf{WH}}$. Given a 3-SAT instance Φ, recall the link L_Φ from Part I (Figure 1), and let L_Φ^{WH} be its positive untwisted Whitehead double. Note that there is a natural bijection between components before and after taking a Whitehead double; let $\kappa_{x_i}^{\mathrm{WH}}$ denote the component corresponding to κ_{x_i} and let $\kappa_{\neg x_i}^{\mathrm{WH}}$ denote the component corresponding to $\kappa_{\neg x_i}$

Remark 19. If we in addition assume[3] that no clause of Φ is of the form $(\ell \vee \ell \vee \ell)$ then, by construction, every component of L_Φ is unknotted. Since the Whitehead double of the unknot is unknotted, we may assume that the components of L_Φ^{WH} in Theorem 20 are all unknotted.

Recall the definition of intermediate invariants (Definition 12) and the examples given in Lemma 13. The goal of this section is to prove:

Theorem 20. *Given a 3-SAT instance Φ with n variables, let L_Φ^{WH} be the link constructed above. Then the following are equivalent, where here i is any intermediate invariant:*

1. *Φ is satisfiable.*

[3]This can be easily assumed without affecting **NP**-hardness. Similarly as in the proof of Lemma 5, we can replace $(\ell \vee \ell \vee \ell)$ with $(\ell \vee \neg t_1 \vee \neg t_2)$ where t_1 and t_2 are new variables forced to be TRUE via the formula Ψ from the proof of Lemma 5.

2. $u(L_\Phi^{WH}) = n$.

3. $i(L_\Phi^{WH}) = n$.

4. $c_s(L_\Phi^{WH}) = n$.

5. $\chi_4(L_\Phi^{WH}) = 0$.

6. L_Φ^{WH} *admits a smoothly slice sublink with* n *components.*

Theorem 2(b)–(d) directly follows from Theorem 20.

Proof. The proof of this Theorem is split in the following steps:

(a) Φ is satisfiable implies that $u(L_\Phi^{\mathrm{WH}}) \leq n$.

(b) $c_s(L_\Phi^{\mathrm{WH}}) \leq i(L_\Phi^{\mathrm{WH}}) \leq u(L_\Phi^{\mathrm{WH}})$.

(c) $\chi_4(L_\Phi^{\mathrm{WH}}) \geq 2n - 2c_s(L_\Phi^{\mathrm{WH}})$.

(d) If $\chi_4(L_\Phi^{\mathrm{WH}}) \geq 0$ then the following two conditions hold:

(d.I) $\chi_4(L_\Phi^{\mathrm{WH}}) = 0$

(d.II) L_Φ^{WH} admits a smoothly slice sublink with n components.

(e) If L_Φ^{WH} admits a smoothly slice sublink with n components, then Φ is satisfiable.

We first show how (a)—(e) prove Theorem 20. Assume first that Φ is satisfiable. Then by (a) and (b) we have that $c_s(L_\Phi^{\mathrm{WH}}) \leq n$, and by (c) we have that $\chi_4(L_\Phi^{\mathrm{WH}}) \geq 0$. Then (d.I) shows that $\chi_4(L_\Phi^{\mathrm{WH}}) = 0$. Working our way back, we see that $c_s(L_\Phi^{\mathrm{WH}}) = i(L_\Phi^{\mathrm{WH}}) = u(L_\Phi^{\mathrm{WH}}) = n$, establishing (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (5). In addition, (d.II) shows directly that (5) $\Rightarrow$ (6). Finally, (e) establishes (6) $\Rightarrow$ (1).

We complete the proof of Theorem 20 by establishing (a)—(e):

(a) Suppose we have a satisfying assignment for Φ (for this implication, *cf.* the proof of Theorem 4). By a single crossing change we resolve the clasp of every component that correspond to a satisfied literal, that is, if $x_i = \textsc{true}$ we change one of the crossings of the clasp of $\kappa_{x_i}^{\mathrm{WH}}$ and if $x_i = \textsc{false}$ we change one of the crossings of the clasp of $\kappa_{\neg x_i}^{\mathrm{WH}}$; as a result, the components corresponding to satisfied literals now form unlinks that are not linked with the remaining components, and we can isotope them away. Since the assignment is satisfying, from each copy of the Borromean rings at least one ring is removed, and the remaining components retract into the first n disks that contained the Hopf links. In each disk we have an untwisted Whitehead double of the unknot which is itself an unknot. Thus we see that the unlink on $2n$ component is obtained, showing that $u(L_\Phi^{\mathrm{WH}}) \leq n$.

(b) By definition of intermediate invariant we have that $c_s \leq i \leq u$.

(c) This is Lemma 14.

(d) Let F be a slice surface for L_Φ^{WH}; recall that F has no closed components. We will use the following notation:

- ν is the number of components of F;
- $\{F_i\}_{i=1}^{\nu}$ are the components of F;
- $g(F_i)$ and $\#\partial F_i$ denote the genus of F_i and the number of its boundary components.

Murasugi [Mur65, Equation 9.4 on Page 416] proved:

$$\left|\sigma(L_\Phi^{\mathrm{WH}})\right| \leq 2\Big(\sum_{i=1}^{\nu} g(F_i)\Big) + \mu - \nu \tag{2}$$

where here μ denotes the number of components of L_Φ^{WH}; applying Murasugi's Theorem with $\mu = 2n$ allows us we estimate $\beta_1(F)$, the first Betti number of F:

Claim 20.1. $\beta_1(F) \geq 2n$.

Proof. Since L_Φ^{WH} has $2n$ components we have that $\sum_{i=1}^{\nu} \#\partial F_i = 2n$; this is used in $\star$ below: Since no component of F is closed, $\beta_1(F_i) = 2g(F_i) + \#\partial F_i - 1$; this is used in $\star\star$ below.

$$\begin{aligned}
2n &= \left|\sigma(L_\Phi^{\mathrm{WH}})\right| && \text{Lemma 17}\\
&\leq 2\Big(\sum_{i=1}^{\nu} g(F_i)\Big) + 2n - \nu && \text{Equation (2)}\\
&= \sum_{i=1}^{\nu} \Big(2g(F_i) + \#F_i - 1\Big) && \star\\
&= \sum_{i=1}^{\nu} \beta_1(F_i) && \star\star\\
&= \beta_1(F)
\end{aligned}$$

This completes the proof of the claim. □

Next we prove:

Claim 20.2. $\nu \geq 2n + \chi(F)$.

Proof. Let $F_{\leq 0}$ be the $\nu - r$ components of F that have nonpositive Euler characteristic; after reordering we may assume that the components of $F_{\leq 0}$ are $F_1, \ldots, F_{\nu-r}$. Since the disk components of F contribute exactly $+r$ to $\chi(F)$ we have

$$\sum_{i=1}^{\nu-r} \chi(F_i) = -r + \chi(F)$$

Solving for r we see:

$$\begin{aligned}
r &= \sum_{i=1}^{\nu-r} \Big(-\chi(F_i)\Big) + \chi(F)\\
&= \sum_{i=1}^{\nu-r} \Big(2g(F_i) + \#F_i - 2\Big) + \chi(F) \qquad (3)
\end{aligned}$$

Disk components contribute zero to $\beta_1(F)$, so $\beta_1(F) = \sum_{i=1}^{\nu-r} \beta_1(F_i)$. Thus we have (here $\star\star$ is as in Claim 20.1):

$$\begin{aligned}
\nu &= r + (\nu - r) \\
&= \sum_{i=1}^{\nu-r} \big(2g(F_i) + \#\partial F_i - 2\big) + \chi(F) + (\nu - r) && \text{Equation (3)} \\
&= \sum_{i=1}^{\nu-r} \big(2g(F_i) + \#\partial F_i - 1\big) + \chi(F) \\
&= \sum_{i=1}^{\nu-r} \beta_1(F_i) + \chi(F) && \star\star \\
&= \beta_1(F) + \chi(F) \\
&\geq 2n + \chi(F) && \text{Claim 20.1}
\end{aligned}$$

This completes the proof of the claim. □

Since F has no closed components we have that $\nu \leq 2n$. Thus by Claim 20.2 we have that $\chi(F) \leq 0$; this establishes Conclusion (d.I).

Now suppose that $\chi(F) = 0$. By Claim 20.2 we have that $\nu = 2n$, that is, F has exactly $2n$ components. Thus each component of F has exactly one boundary component. This means that $F_{\leq 0}$ has exactly $2n - r$ components and each has strictly negative Euler characteristic; thus $\chi(F_{\leq 0}) \leq -(2n - r)$. The r disk components of F contribute exactly $+r$ to $\chi(F)$ we have that

$$0 = \chi(F) \leq -(2n - r) + r = 2r - 2n$$

that is,

$$n \leq r$$

By construction, the link formed by the boundaries of the disks $\{F_{\nu-r+1}, \ldots, F_\nu\}$ is an r component smoothly slice sublink of L_Φ^{WH}, and since $r \geq n$, this establishes Conclusion (d.II).

(e) Finally, assume that L_Φ^{WH} admits an n component smoothly slice sublink L_{SLICE}. By Lemma 18 we have that the positive untwisted Whitehead double of the Hopf link is not a sublink of L_{SLICE} and therefore for each i exactly one of $\kappa_{x_i}^{\text{WH}}$ and $\kappa_{\neg x_i}^{\text{WH}}$ is in L_{SLICE}. If $\kappa_{x_i}^{\text{WH}}$ is in L_{SLICE} we set $x_i = \text{FALSE}$ and if $\kappa_{\neg x_i}^{\text{WH}}$ is in L_{SLICE} we set $x_i = \text{TRUE}$.

Using Lemma 18 again we see that L_{SLICE} does not admit the positive untwisted Whitehead double of the Borromean rings as a sublink. Therefore, from every set of Borromean rings, at least one component does *not* belong to L_{SLICE}; it follows that the assignment above satisfies Φ. □

Acknowledgment

We thank Amey Kaloti and Jeremy Van Horn Morris for many helpful conversations.

References

[AB09] Sanjeev Arora and Boaz Barak. *Computational complexity: a modern approach*. Cambridge University Press, Cambridge, 2009.

[AHT06] Ian Agol, Joel Hass, and William Thurston. The computational complexity of knot genus and spanning area. *Transactions of the American Mathematical Society*, 358:3821–3850, 2006.

[Aic94] Francesca Aicardi. Tree-like curves. *Singularities and bifurcations*, 21:1–31, 1994.

[Arn94a] Vladimir I Arnold. Plane curves, their invariants, perestroikas and classifications. *Advances in Soviet Mathematics*, 21:33–91, 1994.

[Arn94b] Vladimir Igorevich Arnold. *Topological invariants of plane curves and caustics*, volume 5. American Mathematical Soc., 1994.

[CE17] Hsien-Chih Chang and Jeff Erickson. Untangling planar curves. *Discrete & Computational Geometry*, 58(4):889–920, 2017.

[dMRST18] Arnaud de Mesmay, Yo'av Rieck, Eric Sedgwick, and Martin Tancer. Embeddability in $\mathbb{R}^3$ is NP-hard. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1316–1329. Society for Industrial and Applied Mathematics, 2018. Full version on arXiv:1604.00290.

[GJ79] Michael R. Garey and David S. Johnson. *Computers and intractability*. W. H. Freeman and Co., San Francisco, Calif., 1979. A guide to the theory of NP-completeness, A Series of Books in the Mathematical Sciences.

[Hak61] Wolfgang Haken. Theorie der Normalflächen. *Acta Mathematica*, 105(3-4):245–375, 1961.

[HL01] Joel Hass and Jeffrey Lagarias. The number of Reidemeister moves needed for unknotting. *Journal of the American Mathematical Society*, 14(2):399–428, 2001.

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

[HN10] Joel Hass and Tahl Nowik. Unknot diagrams requiring a quadratic number of Reidemeister moves to untangle. *Discrete & Computational Geometry*, 44(1):91–95, 2010.

[KL14] Louis H Kauffman and Sofia Lambropoulou. Hard unknots and collapsing tangles. In *Introductory Lectures on Knot Theory*, 2014.

[KS] Greg Kuperberg and Eric Samperton. Coloring invariants of knots and links are often intractable. Manuscript.

[KT18] Dale Koenig and Anastasiia Tsvietkova. NP-hard problems naturally arising in knot theory. arXiv:1809.10334, 2018.

[Kup14] Greg Kuperberg. Knottedness is in **NP**, modulo GRH. *Adv. Math.*, 256:493–506, 2014.

[Lac15] Marc Lackenby. A polynomial upper bound on Reidemeister moves. *Ann. Math. (2)*, 182(2):491–564, 2015.

[Lac16] Marc Lackenby. The efficient certification of knottedness and Thurston norm. arXiv:1604.00290, 2016.

[Lac17a] Marc Lackenby. Elementary knot theory. In *Lectures on Geometry (Clay Lecture Notes)*. Oxford University Press, 2017.

[Lac17b] Marc Lackenby. Some conditionally hard problems on links and 3-manifolds. *Discrete & Computational Geometry*, 58(3):580–595, 2017.

[Lev12] Adam Simon Levine. Slicing mixed Bing–Whitehead doubles. *Journal of Topology*, 5(3):713–726, 2012.

[Mur65] Kunio Murasugi. On a certain numerical invariant of link types. *Transactions of the American Mathematical Society*, 117:387–422, 1965.

[Pap94] Christos H. Papadimitriou. *Computational complexity*. Addison-Wesley Publishing Company, Reading, MA, 1994.

[Rol90] Dale Rolfsen. *Knots and links*, volume 7 of *Mathematics Lecture Series*. Publish or Perish, Inc., Houston, TX, 1990. Corrected reprint of the 1976 original.

[Sam18] Eric Samperton. Computational complexity of enumerative 3-manifold invariants. *arXiv preprint arXiv:1805.09275*, 2018.

[Shi74] Tetsuo Shibuya. Some relations among various numerical invariants for links. *Osaka J. Math.*, 11:313–322, 1974.

[Tro62] Hale F Trotter. Homology of group systems with applications to knot theory. *Annals of Mathematics*, pages 464–498, 1962.