UNIVERSIDAD AUTÓNOMA DE SAN LUIS POTOSÍ
FACULTAD DE CIENCIAS

Arrangements of pseudocircles on surfaces and knot shadows

TESIS

Que para obtener el grado
de Doctor en Ciencias Aplicadas

en el

Posgrado en Ciencias Aplicadas

presenta

Carolina MEDINA GRACIANO

Asesor:
Dr. Gelasio SALAZAR ANAYA

Mayo 22, 2017 San Luis Potosí, S.L.P., México.

Declaration of Authorship

I, Carolina MEDINA GRACIANO, declare that this thesis titled, “Arrangements of pseudocircles on surfaces and knot shadows” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

Acknowledgements

I would like to use these lines to thank my advisor for all the time and work that he graciously invested on me. I also thank Conacyt for the financial support and LAISLA for allowing me to further develop my studies and complement the research experience.

Contents

Chapter 1

Introduction

The work reported in this thesis is naturally divided into two parts. In the first one, we present results on the embeddability of arrangements of pseudocircles into surfaces. In the second part, we include our work on the number of unknot diagrams associated to a given shadow. Thus the first part falls into the realm of Topological Graph Theory, whereas the second is mainly concerned with Knot Theory.

At a first glance, this thesis then might appear to be the union of two rather disjoint pieces of work. Our own point of view differs with this possible assessment, and we shall use part of this introductory chapter to explain how we ended up working in these (and other) problems in the course of my Ph.D. studies.

1.1 From drawings to arrangements of pseudocircles

Originally, my thesis project revolved around Graph Drawing, in particular dealing with crossing numbers of complete and complete bipartite graphs. At some point in time we were mostly investigating drawings of complete graphs, aiming to verify the Harary-Hill Conjecture for some wide, interesting family of drawings. We came across Jan Kynčl's interesting question published in MathOverFlow [33], in which he raises several interesting questions on (spherical and) pseudospherical drawings.

We found Kynčl's discussion and questions very enticing, and started to work to get a better understanding of pseudospherical drawings. To continue this discussion, it is worth saying a few words about the notion of a pseudospherical drawing. A drawing of a graph G in the sphere is *spherical* if each edge is contained in a great circle of the sphere. This is a natural generalization of rectilinear drawings in the plane. These are drawings in which each edge is a straight segment; equivalently, each edge is contained in a straight line. Thus spherical drawings are indeed the natural generalization, in the sphere, to the widely studied rectilinear drawings in the plane.

The idea behind the concept of a pseudospherical drawing is to mimic the relation of rectilinear drawings with pseudolinear drawings. We recall that a drawing in the plane is *pseudolinear* if each edge can be extended to a pseudoline, so that the resulting collection is an arrangement of pseudolines. Arrangements of pseudolines are a classical, important notion in combinatorial geometry [26]. In the realm of spherical drawings, the natural generalization is that a *pseudospherical* drawing is one in which each edge can be extended to a *pseudocircle* (simple closed curve in the sphere), so that each pair of these pseudocircles intersect exactly in two points, at which they cross. In other words, the resulting pseudocircles form an *arrangement of pseudocircles*.

The relationship between rectilinear and pseudolinear drawings is an interesting, active field of research in graph drawing, and it has led to some of the most important results known for the rectilinear crossing number of complete graphs [1, 2, 38]. This fruitful interplay between rectilinear and pseudolinear drawings led several researchers to propose and investigate several questions on pseudospherical drawings. Among them, as we mentioned above, is Kynčl's MathOverFlow entry [33].

Coincidentally, as we started out investigating some of the questions raised by Kynčl, we participated in a workshop in which Oswin Aichholzer presented several results and questions on pseudospherical drawings. A natural, important question in which we had already started to work on, is the following: is it true that *every* drawing of a complete graph is pseudospherical? Aichholzer presented a drawing of K_6 that was claimed *not* to be pseudospherical. We took a strong interest in this example, with the goal of trying to characterize which drawings of K_n were pseudospherical.

After spending some time working with this particular example, we realized that this drawing was actually pseudospherical. This drawing is illustrated in Figure 1.1. The edges of K_6 are drawn as solid segments, and the dotted segments are the extensions (of the edges) to pseudocircles, so that the result is an arrangement of pseudocircles.

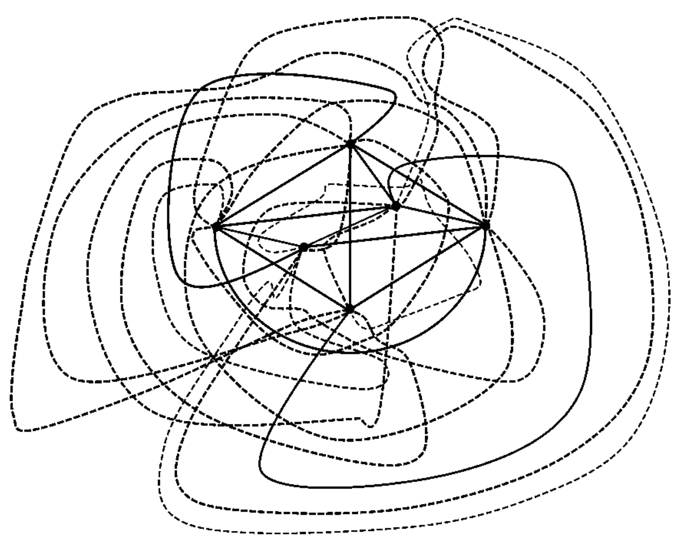


Figure 1.1: A pseudospherical drawing of K_6.

Motivated by this finding, which again left open the question of whether or not every drawing of K_n is pseudospherical, we set out to review the literature in order to learn as much as possible about arrangements of pseudocircles. To our surprise, we found scarcely any results on this topic.

One statement that particularly attracted our attention was a result by Ortner (namely [41, Theorem 10]), claiming that an arrangement of pseudocircles is embeddable into the sphere if and only if all its 4-subarrangements are embeddable into the sphere. We thought that this characterization, or some variant of it, had some potential to be used in our quest to characterize which drawings of K_n are pseudospherical.

As we proceeded to read in detail Ortner's paper, we found the topic highly interesting. A conjecture posed at the end of that paper caught our attention: can these results be generalized to arbitrary compact surfaces? We got seriously engaged in this question, and shifted our focus, from pseudospherical drawings of complete graphs, to the embeddability of arrangements of pseudocircles into compact surfaces.

1.2 Arrangements of pseudocircles in surfaces

There are several variants of the notion of an arrangement of pseudocircles in the literature. These variants are discussed in Section 2.1. As far as we know, the concept originated with Grünbaum, who defined them as collections of simple closed curves (*pseudocircles*) that pairwise intersect in exactly two points, at which they cross. In [41], Ortner also includes the requirement that no three pseudocircles have a point in common.

Working under this definition from [41], we eventually settled Ortner's question by showing that an arrangement of pseudocircles is embeddable into the compact orientable surface Σ_g of genus g if and only if each of its $(4g+4)$-subarrangements embed into Σ. As we tried to streamline our arguments, we realized that the core arguments we were using did not make full use of the conditions in this definition. In the end, we found out that we did not need that every pair of pseudocircles crossed each other exactly twice, or that no three pseudocircles had a common intersection. For the final proof, we only needed that the pseudocircles intersect each other a finite number of times, and that some pseudocircle intersected all the other pseudocircles in the collection. As Ortner had observed in [41], some form of this last condition is absolutely necessary in order to obtain some result along these lines. With this relaxed definition of an arrangement of pseudocircles, we were able to prove our main theorem: an arrangement of pseudocircles is embeddable into the compact orientable surface Σ_g of genus g if and only if each of its $(4g+5)$-subarrangements embed into Σ. As we mentioned above, this can be strengthened to $4g+4$ under the definition of an arrangement given in [41].

These results are stated in Chapter 2. At the end of that chapter, we show that these results can be reformulated in terms of the genera of the subarrangements of an arrangement. These statements are proved in Chapter 3. For the proofs, we introduce the concept of a *cluster of graphs*, which is a collection of pairwise edge-disjoint connected graphs embedded in some (the same) surface, with the condition that there is a graph in the collection (an *anchor*) that intersects all the other graphs.

The notion of a cluster of graphs turns the question of the embeddability of an arrangements of pseudocircles into a question about the genera of subclusters of a given cluster of graphs. This is stated as the Main Theorem in Section 3.1: if the genus of the union of a cluster of graphs $\mathcal{H}$ is at least g, then there is a subcluster $\mathcal{H}_g$ of size at most $4g + 5$ such that the genus of the union of the graphs in $\mathcal{H}_g$ is also at least g. As we explain in that section, the results given above on arrangements of pseudocircles follow immediately from this statement.

The proof of the Main Theorem encompasses the rest of Chapter 3. The heart of the proof is the use of an analogue of Thomassen's 3-path-property for trails of an embedded graph. The seeds of this technique (to which we informally refer to as "short-circuiting non-separating cycles") can be traced back to the work in [41].

We close Part I of this thesis with Chapter 4, where we include some concluding remarks and open questions.

1.3 From arrangements of pseudocircles to knot theory

In the last stages of our work on arrangements of pseudocircles, an opportunity came up to visit Prof. Jorge L. Ramírez-Alfonsín (at the Université de Montpellier) for four weeks. As we tried to find some common interests, the topic of arrangements of pseudocircles came up, and Prof. Ramírez-Alfonsín proposed to investigate a totally different aspect of an arrangement of pseudocircles.

In general terms, the proposal was to regard an arrangement of pseudocircles as a shadow of a link, and to investigate how many distinct link types can be obtained from this shadow. If we start with an arrangement of pseudocircles which, regarded as an embedded graph, has n vertices, then there are 2^n link diagrams that have this arrangement as their shadow. This follows simply because there are two distinct ways to assign the over/undercrossing information at each of the n vertices. The question we set out to investigate was: among these 2^n diagrams, how many nonequivalent link types are there?

We focused our attention on arrangements of pseudocircles as defined in [41]: each pair of pseudocircles cross each other exactly twice, and no three pseudocircles meet at a common point. We note that this last condition is quite natural in this context, as in knot theory it is customary to work with regular diagrams.

During this research stay, and in the subsequent months, we obtained results on two particular types of arrangements of pseudocircles. An instance of the first family of arrangements (which we call *ring* arrangements) is illustrated in Figure 1.2. It is easy to see how to generalize this 4-arrangement to an arrangement of k pseudocircles, for any positive integer k. Specifically, we worked on positive link diagrams that had a ring arrangement as their shadow.

Let us recall the definition of a positive crossing, and of a positive link diagram. If as we travel along a link component (following its orientation) pass over a crossing, and the strand underneath (following its orientation) goes from right to left, the crossing is *positive*;

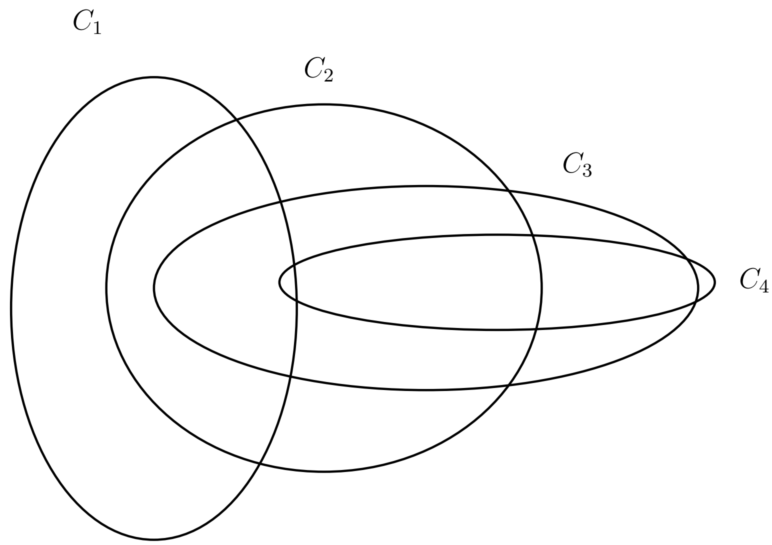


Figure 1.2: The ring arrangement of order 4.

otherwise the crossing is *negative*. Thus a positive crossing looks as in Figure 1.3(a), and a negative crossing looks as in Figure 1.3(b). Therefore every crossing in an oriented link diagram is either positive or negative, and a diagram is *positive* if all its crossings are positive.

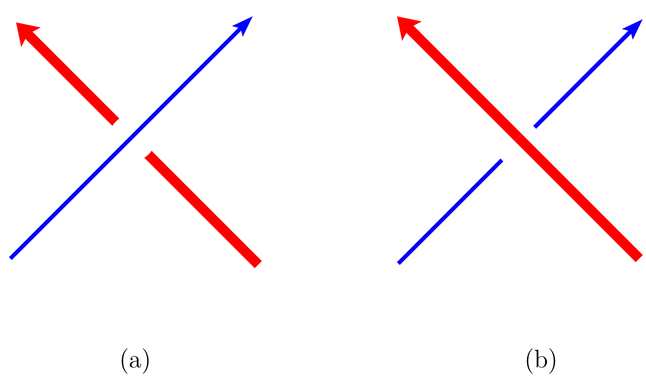


Figure 1.3: In (a) we illustrate a positive crossing, and in (b) a negative crossing.

A ring arrangement with k pseudocircles induces 2^k positive link diagrams, as for each of the k pseudocircles there are two possible ways to choose its orientation, and once the orientation is given the diagram gets determined, by the positiveness condition. The question we investigated was the following. Among these 2^k link diagrams, how many (and which) distinct link types are there? We fully answered this question, both for ring links and for a related family of arrangements of pseudocircles.

Even though we do not include our results on this problem in this thesis, we mention them because later on this investigation this aproach led us to work on a related problem that is a main part of this thesis. As we reviewed the literature on the topic, we found very few results in this direction. We moved on to asking questions of an even more basic nature, such as: how many distinct knot types can one obtain from a given shadow? We finally settled on what is seemingly the most elementary question one could ask in this direction: how many *unknot* diagrams can one obtain from a given shadow?

Our investigations around this basic question turned out to be very fruitful, and constitute Part II of this thesis.

1.4 Unknot and knotted diagrams arising from a given shadow

As we mentioned in the previous section, the leading problem in Part II of this thesis is to take a knot shadow S with n vertices, and investigate, out of the 2^n diagrams that have S as their shadow, how many are unknot diagrams. The basic notions in knot theory that are used in this thesis are reviewed in Chapter 5.

We found in the literature several interesting results on the complement of this question (how many are knotted diagrams). This question arises naturally in the context of long polymer chains, and has been investigated roughly in the following form. If one takes a random knot diagram, what is the probability that it is not a diagram of the unknot? We review the related work around this question in Section 6.1.2.

Back to our question, we did not find any relevant results in the literature. A folklore result that is a staple in every elementary knot theory course is that, given any shadow S, there is always at least *one* unknot diagram that has S as its shadow. The main statement we proved (Theorem 9) is that if S has n vertices, then there are at least $2^{\sqrt[3]{n}}$ unknot diagrams that have S as their underlying shadow. Most of Chapter 6 is devoted to the proof of this result.

Finally, we turned our attention to another quite natural question. Given a knot shadow, is there an over/under assignment at its vertices such that the resulting diagram is a diagram of the trefoil knot? We investigate this, and some related questions, in Chapter 7.

It is easy to see that there exist arbitrarily large shadows that do not have this property (for instance, those shadows that look like the one in Figure 6.3). These shadows, which we call *simple*, have the property that every diagram associated to them is an unknot diagram.

Thus we started out by characterizing which shadows are simple. This turned out to be a straightforward characterization: a shadow is simple if and only if each of its vertices is a cut-vertex (Theorem 14). After presenting this characterization, we proceed to show (Theorem 15) that every shadow that is not simple has an over/under assignment at its vertices that gives a diagram of the trefoil knot.

It seems natural to wonder if similar results can be proved for other shadows, such as the figure-eight knot (after the trefoil knot, the simplest knot). In this direction, we showed that no such statement holds, not only for the figure-eight knot, but for every knot with even crossing number. This is the content of Observation 16, also in Chapter 16.

We close Part II of this thesis with Chapter 8, where we include some concluding remarks and open questions.

Chapter 2

Arrangements of pseudocircles on surfaces

The first part of this thesis is motivated by a question posed in [41]. In that paper, Ortner proved that an arrangement of pseudocircles is embeddable into the sphere if and only if all of its subarrangements of size at most 4 are embeddable into the sphere. Ortner asked if an analogous result held for embeddability into a compact orientable surface Σ_g of genus $g > 0$. We answer this question positively, under an even more general definition of an arrangement of pseudocircles than the one considered in [41]:

Theorem 1. *An arrangement of pseudocircles is embeddable into Σ_g if and only if all of its subarrangements of size at most $4g + 5$ are embeddable into Σ_g.*

As we will see, for the arrangements investigated in [41] (what we will call *strong* arrangements), the size bound $4g + 5$ can be improved to $4g + 4$.

We show that Theorem 1 follows as a consequence of a more general result on the genera of subgraphs of an embedded graph (namely the Main Theorem in Section 3.1). This connection is based on the embedded graph naturally induced by an arrangement of pseudocircles. As we show at the end of this section, Theorem 1 can be equivalently formulated by saying that if the embedded graph induced by an arrangement of pseudocircles has genus greater than g for some $g \geq 0$ (we recall the definition of the genus of an embedded graph in Section 2.2), then there is a subarrangement of size at most $4g + 5$ whose induced embedded graph has genus greater than g.

Before we arrive to this equivalent formulation, we need to review the concept of an arrangement of pseudocircles (Section 2.1), as well as the notion of the embeddability of an arrangement into a surface (Section 2.2).

2.1 Arrangements of pseudocircles

A *pseudocircle* is a simple closed curve on a surface. There exist several variations on the definition of an arrangement of pseudocircles in the literature. These objects were introduced by Grünbaum in [26] (he called them *arrangements of curves*), who required that any two pseudocircles in the collection intersect each other in exactly two points, at which they cross. Under Grünbaum's definition, arrangements of pseudocircles generalize arrangements of circles, in the same way as arrangements of pseudolines generalize arrangements of lines.

Sometimes the pseudocircles in the collection are not required to intersect each other (in [42], an arrangement in which every two pseudocircles intersect is called *complete*). This relaxed definition is used for instance in [28], where Kang and Müller showed (among other results) that every arrangement of at most four pseudocircles in the plane is isomorphic to an arrangement of circles (see also [35, 36]). In addition, sometimes tangential intersections between pseudocircles are allowed. Grünbaum himself proposed this relaxed notion where tangential intersections (or *osculations*) are possible, leading to the concept of a *weak arrangement of curves*. This more general notion is adopted for instance in [5], where Agarwal et al. gave an upper bound on the number of empty lenses in arrangements of pseudocircles, and derived several important applications of this result. Moreover, in the combinatorial formalism of arrangements given in [34], Linhart and Ortner allow pseudocircles to intersect each other more than twice.

The definition used in [41] is in line with the original concept introduced by Grünbaum, with the additional condition that no three pseudocircles meet at a common point. In [41], Ortner defines an arrangement of pseudocircles as a finite collection of pseudocircles in some compact orientable surface, such that:

(i) no three pseudocircles meet each other at the same point;

(ii) each intersection point between pseudocircles is a crossing, rather than tangential; and

(iii) each pair of pseudocircles intersects exactly twice.

We call these collections *strong arrangements of pseudocircles*, to distinguish them from a more general version that we present below.

The motivation behind this version we introduce below is that we realized that our results hold in this more general setting. We need not assume Conditions (i) and (ii). Moreover, we do not need the full strength of (iii), where it is required that every pair of pseudocircles intersect each other: it suffices to ask that there is a pseudocircle intersected by all the other pseudocircles in the collection.

Definition 1. An arrangement of pseudocircles is a finite collection of pseudocircles in some compact orientable surface (the host surface of the arrangement) that pairwise intersect a finite number of times (possibly zero), and such that there exists a pseudocircle that is intersected by all the other pseudocircles in the collection. A pseudocircle with this property (it need not be unique) is an anchor of the arrangement.

This is the definition that we adopt in this work. Clearly, every strong arrangement is also an arrangement according to this definition. A natural generalization of this definition would be to drop the requirement that one pseudocircle is intersected by all the others. However, as we discuss in Chapter 4 (and, as it was pointed out in [41]), without some minimal requirement of this form, no result along the lines of Theorem 1 holds.

2.2 Embeddability of an arrangement of pseudocircles into a surface

Theorem 1 is a statement about the embeddability of an arrangement of pseudocircles into a surface. Since an arrangement is by definition already embedded on a surface, the notion of its embeddability into another surface must be clarified. This concept is based on the isomorphism between arrangements of pseudocircles.

An arrangement of pseudocircles Γ can be naturally regarded as an embedded graph, whose vertices are the points where the pseudocircles intersect each other. Following [41], this embedded graph is the *arrangement graph* of Γ.

We emphasize that an arrangement graph is an *embedded graph*, that is, an abstract graph (a combinatorial entity with vertices and edges) with a *fixed* embedding on some surface. To continue with our discussion on the embeddability of an arrangement into a surface, we need to recall when two embedded graphs are isomorphic. To proceed, we first remind the reader that the *rotation* around a vertex v in an embedded graph G is a cyclic permutation of the edges incident with v; this cyclic rotation records the clockwise order in which these edges leave v in the embedding.

Suppose that G is an embedded graph (on some surface), with vertex set V and edge set E, and G' is an embedded graph (on some surface), with vertex set V' and edge set E'. Then (the embedded graphs) G and G' are *isomorphic* if there is a mapping $\phi : V \cup E \to V' \cup E'$ that is a graph isomorphism when G and G' are regarded as abstract graphs, and in addition the following holds: if the rotation in G of vertex v is $(e_1 e_2 \cdots e_m)$, then the rotation of $\phi(v)$ in G' is $\big(\phi(e_1)\phi(e_2)\cdots\phi(e_m)\big)$. Thus two embedded graphs are isomorphic if their underlying abstract graphs have an isomorphism that preserves and reflects not only the structure of the graphs but also their embeddings.

Remark. Throughout this work, whenever we have two embedded graphs G and G', and mention they are isomorphic, it is tacitly understood that they are isomorphic as embedded graphs, and not only (the weaker, implied fact) that their underlying abstract graphs are isomorphic.

We are now ready to recall when two arrangements of pseudocircles are isomorphic. Let $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_n\}$ and $\Delta = \{\delta_1, \delta_2, \ldots, \delta_n\}$ be arrangements of pseudocircles (note they have the same size). Let G and G' be the arrangement graphs of Γ and Δ, respectively. Then Γ and Δ are isomorphic arrangements if there is an isomorphism from G to G' that maps the pseudocircles in Γ to the pseudocircles in Δ. Formally, Γ and Δ are *isomorphic* if there is an isomorphism $\phi : V \cup E \to V' \cup E'$ from G to G', and a permutation $\rho(1)\,\rho(2)\,\cdots\,\rho(n)$ of $1\,2\,\cdots\,n$ such that the following holds: if the cycle in G corresponding to the pseudocircle γ_i is $v_0 e_1 v_1 \ldots e_m v_0$, then the cycle in G' corresponding to the pseudocircle $\delta_{\rho(i)}$ is $\phi(v_0)\phi(e_1)\phi(v_1)\ldots\phi(e_m)\phi(v_0)$.

At an informal level, this reflects the intuitive notion that two arrangements of pseudocircles are isomorphic if after one removes the whole host surface except for a very thin strip around each edge, and a very small disk around each vertex, the arrangements are undistinguishable.

Definition 2. An arrangement of pseudocircles Γ is embeddable into Σ_g if there is an arrangement Δ isomorphic to Γ such that the host surface of Δ is Σ_g.

We prove Theorem 1 under an equivalent form we give below, which is given in terms of the genus of embedded graphs. We refer the reader to [39] for basic concepts on graph embeddings, such as the facial walks ([39, Sec. 4.1]) and the genus ([39, Eq. (4.2)]) of an embedded graph.

For this discussion we recall that if G is an embedded graph with vertex set V, edge set E, and set of facial walks $\mathcal{W}$, then the *genus* $\text{gen}(G)$ of G is $\text{gen}(G) := (1/2)(2-|V|+|E|-|\mathcal{W}|)$. The essential property of the genus of an embedded graph that we will use is that G is isomorphic to a graph embedded in Σ_g if and only if $\text{gen}(G) \leq g$.

Let Γ be an arrangement of pseudocircles, and let G be its arrangement graph. It follows immediately from Definition 2 that Γ is embeddable into Σ_g if and only if G is isomorphic to a graph embedded in Σ_g. Now from our previous remark, this last condition holds if and only if $\text{gen}(G) \leq g$. Thus we obtain that Γ is embeddable into Σ_g if and only if $\text{gen}(G) \leq g$. Equivalently, Γ is *not* embeddable into Σ_g if and only if $\text{gen}(G) > g$.

With this observation in hand, we note that Theorem 1 can be equivalently interpreted by saying that if Γ is not embeddable into Σ_g, then Γ has a subarrangement Γ' of bounded size (at most $4g+5$) that already witnesseses this non-embeddability; that is, the arrangement graph G' of Γ' satisfies $\text{gen}(G') > g$. We now write this equivalent formulation of Theorem 1 formally, as this is the version under which we work in the next chapter.

Theorem 1 (Equivalent form)**.** *Let* Γ *be an arrangement of pseudocircles with arrangement graph* G*, and let* $g \geq 0$ *be an integer. Then* $\text{gen}(G) > g$ *if and only if* Γ *has a subarrangement* Γ'*, with* $|\Gamma'| \leq 4g+5$*, such that the arrangement graph* G' *of* Γ' *satisfies* $\text{gen}(G') > g$*.*

Chapter 3

Proof of Theorem 1

We devote this chapter to the proof of Theorem 1, in its equivalent form given at the end of Chapter 2.

3.1 Clusters of graphs, the Main Theorem, and proof of Theorem 1

As we came up with the proof of Theorem 1, we realized that our arguments held in a more general setting, and so we ended up obtaining it as a consequence of a more general result. In this section we state this result (the Main Theorem of this chapter), and show that Theorem 1 follows as a corollary.

In its equivalent formulation given at the end of Chapter 2, Theorem 1 can be interpreted as saying that if an embedded graph G with $\mathrm{gen}(G) > g$ can be decomposed into a collection $\mathcal{C}$ of edge-disjoint cycles, where one cycle in $\mathcal{C}$ intersects all the other cycles in $\mathcal{C}$, then there is a subcollection $\mathcal{C}'$ of $\mathcal{C}$, with $|\mathcal{C}'| \leq 4g + 5$, such that $\mathrm{gen}\Big(\bigcup_{C\in\mathcal{C}'} C\Big) > g$. When we proved this statement, we realized that our arguments did not depend on the assumption that the elements of $\mathcal{C}$ were cycles; we only needed their connectedness, and the property that some element in $\mathcal{C}$ intersects all the other elements of $\mathcal{C}$.

This led us to the following concept. A collection $\mathcal{H}$ of pairwise edge-disjoint connected graphs simultaneously embedded in a surface is a *cluster of graphs* if there is a graph H in $\mathcal{H}$ (an *anchor* of $\mathcal{H}$) that intersects every graph in $\mathcal{H}$.

The arrangement graph associated to an arrangement of pseudocircles can thus be naturally regarded as (the union of) a cluster of graphs: each pseudocircle corresponds to a cycle in the cluster, where the cycle that corresponds to an anchor pseudocircle is an anchor of the cluster.

Our main result in this chapter is the following statement, which is thus a generalization of Theorem 1. Throughout this chapter, if $\mathcal{H}$ is a family of graphs embedded on the same surface (such as a cluster), we use $\bigcup\mathcal{H}$ to denote the embedded graph that is the union of the elements of $\mathcal{H}$.

Main Theorem. (Implies Theorem 1). *Let $\mathcal{H}$ be a cluster of graphs such that* $\text{gen}(\bigcup \mathcal{H}) > g$, *for some $g \geq 0$. Then there is an $\mathcal{H}_g \subseteq \mathcal{H}$ with $|\mathcal{H}_g| \leq 4g + 5$, such that* $\text{gen}(\bigcup \mathcal{H}_g) > g$.

In Section 3.2 we state two lemmas and show that they imply the Main Theorem. The rest of the chapter is then almost entirely devoted to the proofs of these lemmas.

We close this section by showing that Theorem 1 is an easy consequence of the Main Theorem. From the previous discussion this could be seen as a mere formality, but we write it for completeness.

Proof of Theorem 1. As we have mentioned, we will prove Theorem 1 in its equivalent formulation given at the end of Section 2.2. The "only if" part is trivial: if a subgraph G' of an embedded graph G satisfies $\text{gen}(G') > g$, then obviously $\text{gen}(G) > g$.

For the "if" part, let $\Gamma = \{\gamma, \gamma_1, \ldots, \gamma_n\}$ be an arrangement of pseudocircles, where γ is an anchor of Γ. Let G be the arrangement graph of Γ, and suppose $\text{gen}(G) > g$ for some $g \geq 0$. Now let C be the cycle in G induced by γ, and let C_i be the cycle in G induced by γ_i, for $i = 1, 2, \ldots, n$.

Then clearly $\mathcal{C} = \{C, C_1, \ldots, C_n\}$ is a cluster of graphs with anchor C. By the Main Theorem, there exists a $\mathcal{C}_g \subseteq \mathcal{C}$, with $|\mathcal{C}_g| \leq 4g + 5$, such that $\text{gen}(\bigcup \mathcal{C}_g) > g$. Now let Γ' be the subcollection of Γ that consists of the pseudocircles that induce the cycles in $\mathcal{C}_g$. Then Γ' satisfies the required conditions, since $|\Gamma'| = |\mathcal{C}_g| \leq 4g + 5$, $\bigcup \mathcal{C}_g$ is the arrangement graph of Γ', and $\text{gen}\big(\bigcup \mathcal{C}_g\big) > g$. □

As we have already mentioned, the size bound $4g+5$ in Theorem 1 can be slightly refined (to $4g+4$) for the class of arrangements considered in [41]. This improvement relies not only on the Main Theorem, but on its proof. Thus we prove this refinement of Theorem 1 in the next section, immediately after the proof of the Main Theorem (see Remark at the end of Section 3.2.1).

3.2 Proof of the Main Theorem

As we shall see shortly, the Main Theorem follows easily by an induction based on the following statement. We note that if H is an anchor of a cluster $\mathcal{H}$, then in particular H is an embedded graph (subgraph of the embedded graph $\bigcup \mathcal{H}$), and as such, H has a genus. We encourage the reader to follow our own custom, which is to informally think of our next statement as "at most 4 graphs of the cluster need to be added to the anchor, to obtain a graph whose genus is greater than the genus of the anchor".

Theorem 2. *Let $\mathcal{H}$ be a cluster of graphs with anchor H. Suppose that* $\text{gen}(H) < \text{gen}(\bigcup \mathcal{H})$. *Then there is a subcollection $\overline{\mathcal{H}} \subseteq \mathcal{H}$, with $H \in \overline{\mathcal{H}}$ and $|\overline{\mathcal{H}}| \leq 5$, such that* $\text{gen}(H) < \text{gen}\big(\bigcup \overline{\mathcal{H}}\big)$.

In Section 3.2.2 we state two lemmas and show that they imply Theorem 2. Before we proceed to that, we show that the Main Theorem follows from this statement.

3.2.1 The Main Theorem follows from Theorem 2

Proof of the Main Theorem. Assuming Theorem 2, we prove the Main Theorem by induction on g, for a fixed cluster of graphs $\mathcal{H}$ with anchor H.

For the base case $g = 0$ the assumption is that $\text{gen}(\bigcup \mathcal{H}) > 0$. If $\text{gen}(H) > 0$ then we are done by taking $\mathcal{H}_0 := \{H\}$, and so we may assume that $\text{gen}(H) = 0$. In this case we apply Theorem 2, to obtain a subcollection $\overline{\mathcal{H}} \subseteq \mathcal{H}$, with $H \in \overline{\mathcal{H}}$ and $|\overline{\mathcal{H}}| \leq 5$ such that $\text{gen}(H) < \text{gen}(\bigcup \overline{\mathcal{H}})$. Thus in particular $\text{gen}(\bigcup \overline{\mathcal{H}}) > 0$, and so we are done by setting $\mathcal{H}_0 := \overline{\mathcal{H}}$.

We now suppose that the Main Theorem holds for $g = h$ for some $h \in \{0, 1, \ldots, \text{gen}(\bigcup \mathcal{H}) - 2\}$, and show that then it holds for $g = h + 1$. (Note that the largest value of g for which the statement of the Main Theorem makes sense is $g = \text{gen}(\bigcup \mathcal{H}) - 1$).

The assumption that the statement holds for $g = h$ means that there is a subcollection $\mathcal{H}_h$ of $\mathcal{H}$ with $|\mathcal{H}_h| \leq 4h + 5$ such that $\text{gen}(\bigcup \mathcal{H}_h) > h$. If $\text{gen}(\bigcup \mathcal{H}_h) > h + 1$ then we are done by setting $\mathcal{H}_{h+1} := \mathcal{H}_h$, so we may assume that $\text{gen}(\bigcup \mathcal{H}_h) = h + 1$. Note that since $h \leq \text{gen}(\bigcup \mathcal{H}) - 2$, it follows that $\text{gen}(\bigcup \mathcal{H}) > h + 1$.

Let $K := \bigcup \mathcal{H}_h$, and let $\mathcal{K}$ be the collection $\{K\} \cup (\mathcal{H} \setminus \mathcal{H}_h)$. We claim that $\mathcal{K}$ is acluster of graphs with anchor K. To see this, note that every graph in $\mathcal{H}$ intersects H, and so in particular every graph in $\mathcal{H} \setminus \mathcal{H}_h$ intersects H. Since H is contained in K, it follows that every graph in $\mathcal{H} \setminus \mathcal{H}_h$ intersects K. This proves the claim.

Note that $\bigcup \mathcal{K} = \bigcup \mathcal{H}$. Thus $\text{gen}(\bigcup \mathcal{K}) = \text{gen}(\bigcup \mathcal{H}) > h+1$. Since $\text{gen}(K) = \text{gen}(\bigcup \mathcal{H}_h) = h+1$, it follows that $\text{gen}(K) < \text{gen}(\bigcup \mathcal{K})$. Thus by Theorem 2 there is a subcollection $\overline{\mathcal{K}} \subseteq \mathcal{K}$, with $K \in \overline{\mathcal{K}}$ and $|\overline{\mathcal{K}}| \leq 5$, such that $\text{gen}(\bigcup \overline{\mathcal{K}}) > \text{gen}(K)$. Let $\mathcal{H}_{h+1} := \mathcal{H}_h \cup \left(\overline{\mathcal{K}} \setminus \{K\}\right)$. Note that $\bigcup \overline{\mathcal{K}} = \bigcup \mathcal{H}_{h+1}$. Thus $\mathcal{H}_{h+1}$ is a subcollection of $\mathcal{H}$ such that $|\mathcal{H}_{h+1}| = |\mathcal{H}_h| + |\overline{\mathcal{K}}| - 1 \leq |\mathcal{H}_h| + 4 \leq (4h + 5) + 4 = 4(h + 1) + 5$. Since $\text{gen}(\bigcup \mathcal{H}_{h+1}) = \text{gen}(\bigcup \overline{\mathcal{K}}) > \text{gen}(K) = h + 1$, it follows that $\mathcal{H}_{h+1}$ satisfies the required properties. □

Remark. As we mentioned in Section 2.2, the size bound $4g+5$ in Theorem 1 can be improved to $4g + 4$ if the arrangement of pseudocircles under consideration is strong. To see this, we note that the size bound $4g + 5$ in the Main Theorem can be improved to $4g + 4$ if for the base case in the proof we can guarantee the existence of an $\mathcal{H}_0$ with $|\mathcal{H}_0| \leq 4$ and such that $\text{gen}(\mathcal{H}_0) > 0$. Now if Γ is a strong arrangement of pseudocircles that cannot be embedded into the sphere, [41, Theorem 10] guarantees that there is a subarrangement Γ_0 of size at most 4 that cannot be embedded into the sphere. Thus in this case the collection $\mathcal{H}_0$ of those cycles (in the cluster of graphs $\mathcal{H}$ associated to Γ) that correspond to the pseudocircles in Γ_0 satisfies $|\mathcal{H}_0| \leq 4$ and $\text{gen}(\mathcal{H}_0) > 0$, as required.

3.2.2 Reducing Theorem 2 to two lemmas

We now show that Theorem 2 is an easy consequence of two lemmas we state below, and whose proofs encompass most of the rest of this chapter. These lemmas involve the concept of the degeneracy of a face of an embedded graph, which we now proceed to explain.

First we recall that if G is an embedded graph, then a *face* of G is a connected component of $\mathbb{R}^2 \setminus G$. If a graph is *cellularly embedded* (that is, if each face is homeomorphic to an open disk), then the collection of facial walks determines the embedding, but this is not true for a non-cellularly embedded graph. In the general case, each face is homeomorphic to a compact surface of some genus $g \geq 0$ from which a finite number $m \geq 1$ of points have been removed; here g is the *genus* of the face, and $d := m - 1$ its *degeneracy*. For a face with degeneracy d, there are $d + 1$ facial walks that bound the face. In a cellular embedding, both the genus and the degeneracy of each face are equal to zero. Indeed, each face is bounded by a single facial walk (that is, its degeneracy is zero), and each face is homeomorphic to an open disk, that is, to a sphere (compact surface of genus 0) with 1 point removed. If a face has positive degeneracy, then we say it is *degenerate*; otherwise it is *non-degenerate*.

We illustrate the concepts of genus and degeneracy of a face in Figure 3.1. In this embedded graph, each of the faces F_2, F_3, and F_4 is homeomorphic to an open disk (that is, to a sphere minus one point), and so it has both genus and degeneracy zero. Face F_1 is homeomorphic to a torus minus one point, so it has genus 1 and degeneracy zero. Finally, F_5 is homeomorphic to a sphere minus two points (note that it is bounded by two facial walks), and so it has genus zero and degeneracy 1.

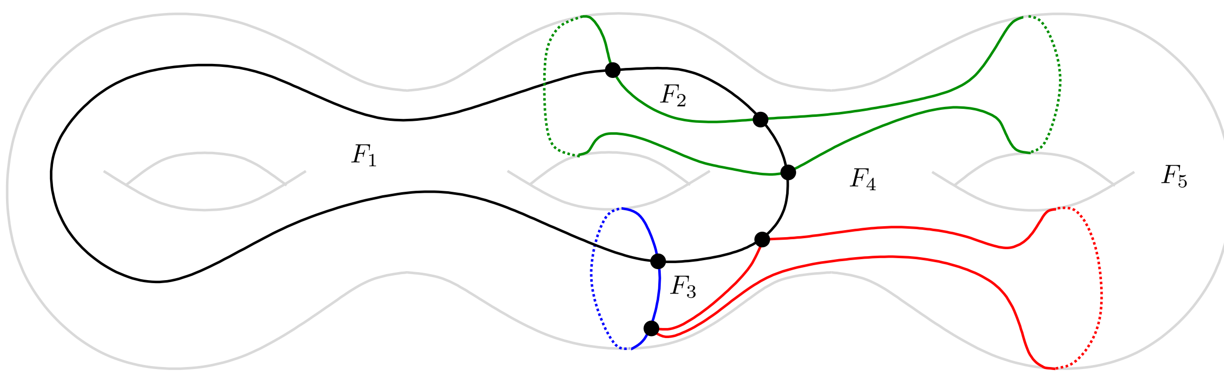


Figure 3.1: Illustration of the genus and the degeneracy of the faces of an embedded graph. Faces F_2, F_3, and F_4 have both genus and degeneracy zero. The fact that F_4 has degeneracy zero may not be immediately obvious, but it is readily verified since it does not contain any non-contractible simple closed curve. Face F_1 has genus 1 and degeneracy zero, and F_5 has genus zero and degeneracy 1.

We are ready to state the lemmas that, put together, imply Theorem 2. We note that in both lemmas we assume that $\bigcup \mathcal{H}$ is cellular. This is an essential assumption for the proofs of these lemmas but, as we shall see shortly, Theorem 2 will follow even if its statement does not include this as a hypothesis.

The first key lemma is the following, which we informally capture by saying that "if the anchor has a degenerate face, then at most 2 graphs of the cluster need to be added to the anchor, to obtain a graph whose genus is greater than the genus of the anchor".

Lemma 3. *Let $\mathcal{H}$ be a cluster of graphs with anchor H, such that $\bigcup \mathcal{H}$ is cellular. Suppose that* $\text{gen}(H) < \text{gen}(\bigcup \mathcal{H})$*, and that H has a degenerate face. Then there is a collection $\mathcal{H}' \subseteq \mathcal{H}$, that includes H and satisfies $|\mathcal{H}'| \leq 3$, such that* $\text{gen}\Big(H\Big) < \text{gen}\Big(\bigcup \mathcal{H}'\Big)$.

We now state the second key lemma. Informally speaking, this says that "if all the faces of the anchor are non-degenerate, then at most two graphs of the cluster need to be added to the anchor, so that the resulting graph either (i) has greater genus than the anchor; or (ii) has a degenerate face". Formally:

Lemma 4. *Let $\mathcal{H}$ be a cluster of graphs with anchor H, such that $\bigcup \mathcal{H}$ is cellular. Suppose that* $\text{gen}(H) < \text{gen}(\bigcup \mathcal{H})$*, and that every face of H is non-degenerate. Then there is a collection $\mathcal{H}'' \subseteq \mathcal{H}$, that includes H and satisfies $|\mathcal{H}''| \leq 3$, such that either (i)* $\text{gen}(H) < \text{gen}(\bigcup \mathcal{H}'')$*; or (ii) $\bigcup \mathcal{H}''$ has a degenerate face.*

Most of the rest of this section (and chapter) is devoted to proving these lemmas. We close this section by showing that they imply Theorem 2.

Proof of Theorem 2, assuming Lemmas 3 and 4. First we show that if Theorem 2 holds when $\bigcup \mathcal{H}$ is cellular, then it always holds. Suppose that $\mathcal{I}$ is a cluster of graphs with anchor I, where $\text{gen}(I) < \text{gen}(\bigcup \mathcal{I})$, and $\bigcup \mathcal{I}$ is not cellular. Every embedded graph is isomorphic to a cellularly embedded graph, and in particular there exists a cluster of graphs $\mathcal{H}$ such that $\bigcup \mathcal{H}$ is cellular, and an isomorphism $\phi : \bigcup \mathcal{I} \to \bigcup \mathcal{H}$ that maps each element of $\mathcal{I}$ to an element of $\mathcal{H}$. The image H of I under ϕ is then an anchor of $\mathcal{H}$, and $\text{gen}(H) < \text{gen}(\bigcup \mathcal{H})$.

Suppose that Theorem 2 holds for $\mathcal{H}$. Thus there is an $\overline{\mathcal{H}} \subseteq \mathcal{H}$, with $H \in \overline{\mathcal{H}}$ and $|\overline{\mathcal{H}}| \leq 5$, such that $\text{gen}(H) < \text{gen}(\bigcup \overline{\mathcal{H}})$. Then the collection $\overline{\mathcal{I}} := \{\phi^{-1}(K) \mid K \in \overline{\mathcal{H}}\}$ contains I, satisfies $|\overline{\mathcal{I}}| \leq 5$, and $\text{gen}(I) < \text{gen}(\bigcup \overline{\mathcal{I}})$. That is, Theorem 2 also holds for $\mathcal{I}$. Therefore, as claimed, it suffices to prove the theorem under the assumption that $\bigcup \mathcal{H}$ is cellular.

Thus we let $\mathcal{H} = \{H, H_1, \ldots, H_n\}$ be a cluster of graphs with anchor H, such that $\text{gen}(H) < \text{gen}(\bigcup \mathcal{H})$, and $\bigcup \mathcal{H}$ is cellular. If H has a degenerate face, then Theorem 2 follows immediately from Lemma 3. Thus we suppose that all the faces of H are non-degenerate, and apply Lemma 4. Thus there exist (not necessarily distinct) graphs $H_i, H_j \in \mathcal{H} \setminus \{H\}$ such that for the embedded graph that is the union $H \cup H_i \cup H_j$ either (i) $\text{gen}(H \cup H_i \cup H_j) > \text{gen}(H)$; or (ii) $H \cup H_i \cup H_j$ has a degenerate face. In the first case we are done by letting $\overline{\mathcal{H}} := \{H, H_i, H_j\}$. Thus we assume that (ii) holds, and (i) does not, that is, $\text{gen}(H \cup H_i \cup H_j) = \text{gen}(H)$.

Since $\text{gen}(H) < \text{gen}(\bigcup \mathcal{H})$, it then follows that $\text{gen}(H \cup H_i \cup H_j) < \text{gen}(\bigcup \mathcal{H})$, and so $\mathcal{J} := \{H_r \in \mathcal{H} \mid r \notin \{i, j\}\}$ is not empty. The collection $\mathcal{K} := \{H \cup H_i \cup H_j\} \cup \mathcal{J}$ is then a cluster of graphs with anchor $H \cup H_i \cup H_j$, since the anchor property of H in $\mathcal{H}$ is obviously inherited to $H \cup H_i \cup H_j$ in $\mathcal{K}$.

Since $\bigcup \mathcal{K} = \bigcup \mathcal{H}$, it follows that $\text{gen}(H \cup H_i \cup H_j) < \text{gen}(\bigcup \mathcal{K})$. Recall that the anchor $H \cup H_i \cup H_j$ of $\mathcal{K}$ has a degenerate face. Thus we can apply Lemma 3 to $\mathcal{K}$, to obtain that there exist (not necessarily different) graphs H_k, H_ℓ in $\mathcal{J}$ such that $\text{gen}\Big((H \cup H_i \cup H_j) \cup H_k \cup H_\ell\Big) > \text{gen}\Big(H \cup H_i \cup H_j\Big)$. Therefore we have $\text{gen}\Big(H \cup H_i \cup H_j \cup H_k \cup H_\ell\Big) > \text{gen}(H)$, and so we are done by setting $\overline{\mathcal{H}} := \{H, H_i, H_j, H_k, H_\ell\}$. □

3.3 Proof of Lemma 3

Let $\mathcal{H} = \{H, H_1, \ldots, H_n\}$ be a cluster of graphs such that $\bigcup \mathcal{H}$ is cellular, H is an anchor of $\mathcal{H}$, and H has a degenerate face F. Let $\mathcal{W}$ be the set of facial walks of H that bound F. The degeneracy of F means that $|\mathcal{W}| \geq 2$. We refer the reader to Figure 3.2(a), where we illustrate an anchor H in the double torus, and a face F that is homeomorphic to a disk minus two points; thus the degeneracy of the face F in this example is exactly 1.

We let I denote the subgraph of $\bigcup \mathcal{H}$ induced by the edges contained in the face F. Since $\bigcup \mathcal{H}$ is cellular and F is a degenerate face of H (in particular, F is not homeomorphic to an open disk), it follows that I is not a null graph, that is, I has at least one edge.

To help comprehension, we say that the edges of H_i are *of colour* i, for $i = 1, \ldots, n$. Thus every edge of $\mathcal{H} \setminus \{H\}$ has (exactly) one colour; in particular, every edge in I is in $\mathcal{H} \setminus \{H\}$, and so it has one colour. A subgraph of $\mathcal{H}$ is *monochromatic* if all its edges are of the same colour.

The graph I can be decomposed as the edge-disjoint union of graphs $G_1, G_2, \ldots, G_m$ such that, for $k = 1, \ldots, m$, G_k is a connected monochromatic subgraph of I, and it is maximal with respect to these properties. Note that it may be that $m > n$; indeed, even though each element of $\mathcal{H}$ is connected, the intersection of H_i with F may be disconnected for some $i \in \{1, 2, \ldots, n\}$.

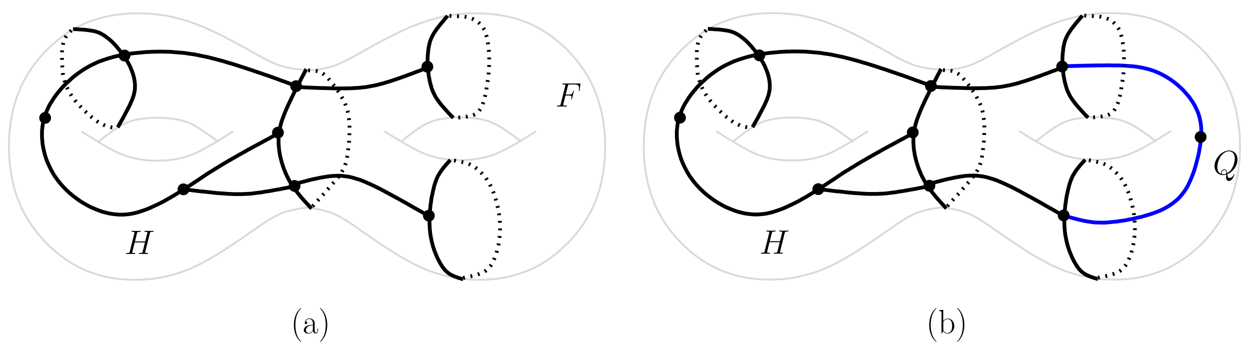


Figure 3.2: In (a) we depict the anchor H of a cluster of graphs $\mathcal{H}$ (the other graphs of $\mathcal{H}$ are not shown). Here H has a face F bounded by two facial walks. Thus F has degeneracy 1. In (b) we illustrate a path Q contained in F, except for its endpoints, one of which lies on W_1, and the other one lies on W_2.

The connectedness of each $H_i \in \mathcal{H}$ implies that G_k has at least one vertex in common with some walk in $\mathcal{W}$, for $k = 1, \ldots, m$. Indeed, suppose that some $G_k \in \{G_1, \ldots, G_m\}$ has no vertex in common with any walk in $\mathcal{W}$; thus G_k is completely (including its vertices) contained in F. Let i be the colour of the edges in G_k. Since H_i is connected, it follows that H_i must equal G_k, and in particular, that H_i does not intersect H. But this is impossible, since H is an anchor of $\mathcal{H}$. From this observation it follows that $\mathcal{G} := \{H, G_1, \ldots, G_m\}$ is a cluster of graphs with anchor H.

We claim that to prove the lemma it is enough to show that there exists a subcollection $\mathcal{G}'$ of $\mathcal{G}$, that includes H and satisfies $|\mathcal{G}'| \leq 3$, such that $\text{gen}(H) < \text{gen}(\bigcup \mathcal{G}')$. For suppose such a $\mathcal{G}'$ exists, then $\mathcal{G}' = \{H, G_k, G_\ell\}$ for some (non-necessarily distinct) $k, \ell \in \{1, \ldots, m\}$. Now let i (respectively, j) be the colour of the edges in G_k (respectively, G_ℓ). Thus G_k is a subgraph of H_i, and G_ℓ is a subgraph of H_j. Let $\mathcal{H}' := \{H, H_i, H_j\}$. Since $G_k \cup G_\ell \subseteq H_i \cup H_j$, then $\text{gen}(\bigcup \mathcal{G}') \leq \text{gen}(\bigcup \mathcal{H}')$, and so $\text{gen}(H) < \text{gen}(\bigcup \mathcal{G}')$ implies that $\text{gen}(H) < \text{gen}(\bigcup \mathcal{H}')$. Thus the lemma follows, since $\mathcal{H}'$ satisfies the required properties.

Thus we devote the rest of the proof to show that there exists a subcollection $\mathcal{G}'$ of $\mathcal{G}$, that includes H and satisfies $|\mathcal{G}'| \leq 3$, such that $\text{gen}(H) < \text{gen}(\bigcup \mathcal{G}')$.

Let $G_i \in \mathcal{G}$. If for a walk $W \in \mathcal{W}$ the graph G_i has a vertex in common with W, we say that G_i *attaches to* W. We recall that each $G_i \in \mathcal{G}$ attaches to at least one walk in $\mathcal{W}$.

We first deal with the case in which there is a $G_k \in \mathcal{G}$ that attaches to two distinct facial walks $W_1, W_2 \in \mathcal{W}$. In this case there is a path P from a vertex $u \in W_1$ to a vertex $v \in W_2$ that is contained in F except for its endpoints, and such that P is contained in G_k. We claim that $\text{gen}(H) < \text{gen}(H \cup P)$. Note that this settles the lemma in this case by setting $\mathcal{G}' = \{H, G_k\}$, as the fact that P is contained in G_k implies that $\text{gen}(H \cup P) \leq \text{gen}(H \cup G_k)$.

To see that $\text{gen}(H) < \text{gen}(H \cup P)$, we note that if q is the length (number of edges) of P, then $P \cup H$ has q more edges, $q - 1$ more vertices, and one facial walk less than H (P collapses W_1 and W_2 into a single facial walk). Thus an elementary counting gives that $\text{gen}(P \cup H) = \text{gen}(H) + 1$.

In the remaining case, each $G_i \in \mathcal{G}$ attaches to exactly one facial walk in $\mathcal{W}$. Let $W_1, W_2, \ldots, W_r$ be the elements of $\mathcal{W}$. Thus if we say that $G_i \in \mathcal{G}$ is *of type* s if it attaches to walk W_s, then each $G_i \in \mathcal{G}$ is of type s for exactly one $s \in \{1, 2, \ldots, r\}$. We claim that there is a $G_i \in \mathcal{G}$ of type 1 that intersects a graph of type s for some $s \geq 2$. Seeking a contradiction, suppose that this is not the case. Then there exists a simple closed curve α contained in F, with the following properties:

(a) α does not intersect $G_1 \cup G_2 \cup \cdots \cup G_m$ (and therefore does not intersect $\bigcup \mathcal{H}$); and

(b) $F \setminus \alpha$ has two components, one that contains all the edges of type 1 and one that contains all the edges of type s for $s \geq 2$.

Now α is a non-separating curve in the host surface of $\mathcal{H}$, and since it does not intersect $\bigcup \mathcal{H}$ it follows that the face of $\bigcup \mathcal{H}$ that contains α is not homeomorphic to an open disk, contradicting the cellularity of $\bigcup \mathcal{H}$.

Hence there exist a $G_k \in \mathcal{G}$ of type 1, and a $G_\ell \in \mathcal{G}$ of type $s \geq 2$, such that G_k and G_ℓ have at least one common vertex. It follows that there is a path Q contained in $G_k \cup G_\ell$, with one endpoint in W_1 and the other endpoint in W_s, and that except for these endpoints is contained in F. Let r denote the length of Q. Then $H \cup Q$ has r more edges, $r - 1$ more vertices, and facial walk less than H (here Q collapses W_1 and W_2 into a single facial walk). Again an elementary counting yields that $\text{gen}(H) < \text{gen}(H \cup Q)$, and so $\text{gen}(H) < \text{gen}(H \cup G_k \cup G_\ell)$. Thus in this case we are done by setting $\mathcal{G}' := \{H, G_k, G_\ell\}$. $\square$

3.4 Proof of Lemma 4

The proof of Lemma 4 consists of two steps, which are stated as Claims A and B below. For the proof of Claim B we assume the following statement on non-separating cycles in clusters of graphs whose anchor consists of a single vertex. We recall that if D is a cycle in a graph embedded in a surface Σ such that $\Sigma \setminus D$ is connected, then D is *non-separating*.

Proposition 5. *Let $\mathcal{G}$ be a cluster of graphs, where the anchor G consists of a single vertex, and $\bigcup \mathcal{G}$ is cellularly embedded in Σ_g for some $g \geq 1$. Then there is a subcollection $\mathcal{G}_\circ \subseteq \Big(\mathcal{G} \setminus \{G\}\Big)$, with $|\mathcal{G}_\circ| \leq 2$, such that $\bigcup \mathcal{G}_\circ$ contains a non-separating cycle.*

Deferring the proof of this proposition to Section 3.6, we now move on to reducing Lemma 4 to Claims A and B mentioned above.

Throughout this section, $\mathcal{H} = \{H, H_1, \ldots, H_n\}$ is a cluster of graphs with anchor H as in the statement of Lemma 4. Thus $\bigcup \mathcal{H}$ is cellular, $\text{gen}(H) < \text{gen}(\bigcup \mathcal{H})$, and every face of H is non-degenerate. We let Σ_g be the host surface of $\mathcal{H}$.

The assumption $\text{gen}(H) < \text{gen}(\bigcup \mathcal{H})$ implies that H is not cellularly embedded in Σ_g. Now any graph that is not cellularly embedded has a face that either is degenerate or has positive genus. Since by assumption every face of H is non-degenerate, it follows that H has a face F with positive genus (and degeneracy zero). That is, F is homeomorphic to a compact surface of positive genus from which a single point has been removed. Let W be the unique facial walk of H that bounds F.

The situation is illustrated in (a) of Figure 3.3. The anchor H is contained in the left handle of the double torus, and it is easy to see that $\text{gen}(H) = 1$ (that is, H "fills the left handle"). In this case F has genus 1, as it is homeomorphic to a torus minus one point.

Now suppose that P is a path that is contained in F except for its endpoints u and v, which lie on W. We say that such a path is *F-non-separating* if there is a path R from u to v, contained in W, such that $R \cup P$ is a non-separating cycle. An F-non-separating path is illustrated in (b) of Figure 3.3.

Lemma 4 is an immediate consequence of the following two claims, whose proofs encompass the rest of this section.

Claim A. *Suppose that there exists a subcollection $\mathcal{H}_\circ$ of $\mathcal{H}$, with $|\mathcal{H}_\circ| \leq 2$, such that $\bigcup \mathcal{H}_0$ contains one of the following:*

(i) An F-non-separating path.

(ii) A non-separating cycle contained in F, except perhaps for a single vertex that lies on W.

Then Lemma 4 holds by setting $\mathcal{H}'' = \{H\} \cup \mathcal{H}_\circ$.

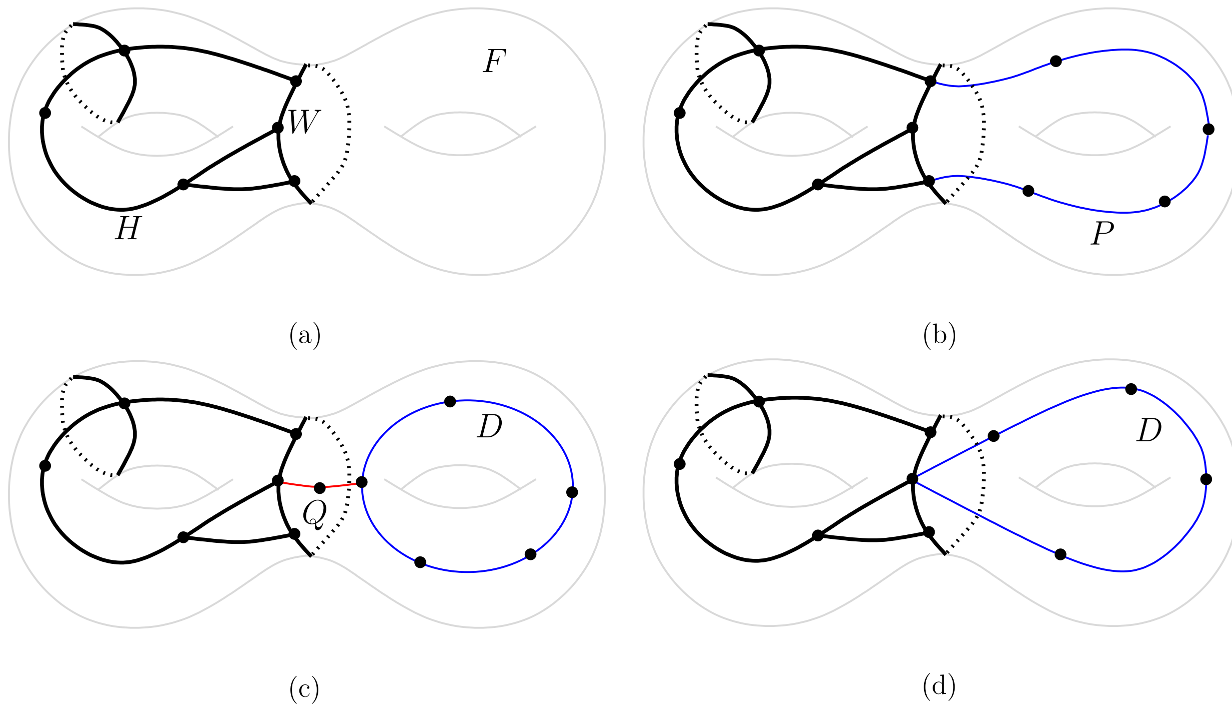


Figure 3.3: In (a) we depict the anchor H of a cluster of graphs $\mathcal{H}$ (the other graphs of $\mathcal{H}$ are not shown), where H has a face F with degeneracy zero and genus 1. The (unique) facial walk of H that bounds F is W. In (b), (c), and (d) we show the three structures involved in Claim A. In (b) we show an F-non-separating path. In (c) we have a non-separating cycle D completely contained in F; in particular, D is disjoint from W. (The path Q also shown in (c), with one endpoint in W and one endpoint in D, is used in the proof of Claim A). Finally, in (d) we show a non-separating cycle D that is contained in F, except for a single vertex that lies on W.

Possibility (i) in the statement of Claim A is illustrated in (b) of Figure 3.3. The two possible situations in (ii) (a non-separating cycle completely contained in F, and a non-separating cycle contained in F except for a single vertex that lies on W) are illustrated in (c) and (d) of Figure 3.3, respectively. (The path Q in (c) is used in the proof of Claim A).

Claim B. *There exists a subcollection $\mathcal{H}_\circ$ of $\mathcal{H}$ with the properties stated in Claim A.*

3.4.1 Proof of Claim A.

Proof. Consider the subcollection $\mathcal{H}_\circ$ whose existence is assumed in Claim A. Thus $\mathcal{H}_\circ$ consists either of one or two elements in $\mathcal{H}$. To simplify the discussion, it is valid to say that $\mathcal{H}_\circ = \{H_k, H_\ell\}$, where H_k and H_ℓ are *not necessarily distinct* elements of $\mathcal{H}$.

Our first step is to produce, from one of the structures in the statement of Claim A (an F-non-separating path or a non-separating cycle) a subgraph L of $H_k \cup H_\ell$ with certain properties. If there is an F-non-separating path P, then we let $L := P$. If there is a non-separating cycle D contained in F, except for a single vertex that lies on W, then we let $L := D$. In the alternative, there is a non-separating cycle D completely contained in F. In this case, since H_k and H_ℓ are connected, and each of them has at least one vertex in common with H, it follows that there is a path Q contained in $H_k \cup H_\ell$, with one endpoint in D and the other endpoint in W, that is otherwise disjoint from $D \cup W$. (See Figure 3.3(c)). Thus Q is contained in F, except for its endpoint in W. In this case we let $L := D \cup Q$.

We note that in every case L satisfies the following properties:

1. L is a subgraph of $H_k \cup H_\ell$.

2. L is contained in F, except for some vertices (either one or two) that lie on W.

3. $H \cup L$ has a degenerate face. (This is the important structural property).

Properties 1 and 2 follow immediately from the construction of L. To see that Property 3 holds, note that $F \setminus L$ is a face F_L of $H \cup L$. It is immediately verified that (regardless of whether L is a path, or a cycle, or a cycle plus a path) F_L bounds exactly two facial walks of $H \cup L$, and thus it is degenerate.

The argument to finish the proof of Claim A is heavily based on the following remark. In what follows, if e is an edge of a connected graph G that is incident with a degree 1 vertex, then we use $G - e$ to denote the graph that results by removing from G both e and this degree 1 vertex. If e is not incident with any degree 1 vertex, then $G - e$ is simply the graph that results by removing e from G.

Remark. Let G be a connected embedded graph, and let e be an edge of G such that $G - e$ is also connected. If $G - e$ has a degenerate face, then either (a) $\text{gen}(G) > \text{gen}(G - e)$; or (b) G has a degenerate face.

Proof of the Remark. Let G and e be as in the statement of the Remark, and let J be a degenerate face of $G - e$. If e is not inside the face J, then J is also a face of G, with facial walks a subdivision of the facial walks in $G - e$; thus in this case G has a degenerate face.

In the alternative, e is inside J. In this case J (which is a face of $G - e$) is not a face of G, but is contained in at most two faces of G. Now e must have at least one endpoint u in a facial walk U of J, and the other endpoint v of e is either inside J or lies on a facial walk of J. If v is inside J then it is a degree one vertex of G, and $J - e$ is a face of G with the same number of facial walks as J; in particular, in this case G has a degenerate face, namely $J - e$. Thus it only remains to analyse the case in which v lies on a facial walk U' of J.

If U and U' are distinct facial walks, then G has one more edge and one facial walk less than $G - e$ (the facial walks U, U' get collapsed into a single facial walk by the addition of e). Thus in this case an elementary counting argument shows that $\text{gen}(G) = \text{gen}(G - e) + 1 > \text{gen}(G - e)$, and so we are done.

Thus we are left with the case in which U and U' are the same facial walk. We note that then U together with e induces two facial walks U_1, U_2 in G.

Assume first that $J - e$ is connected. In this case $J - e$ is a face of G, which is bounded by (at least) the facial walks U_1 and U_2; in particular, G has a degenerate face, and so we are done. Finally, suppose that $J - e$ is disconnected. Since J is connected (it is a face of $G - e$) it follows that $J - e$ has exactly two components J_1, J_2, which are faces of G. One of these faces is bounded by U_1, and the other is bounded by U_2. Without loss of generality, U_1 bounds J_1 and U_2 bounds J_2. Now since J is degenerate, it follows that there is a facial walk U'', distinct from U, that also bounds J and has no vertex in comun with e. Then U'' must also bound either J_1 or J_2. In the former case, J_1 is degenerate, and in the latter case J_2 is degenerate. □

We are finally ready to finish the proof of Claim A, by showing that Lemma 4 follows by letting $\mathcal{H}'' := \mathcal{H}_\circ = \{H, H_k, H_\ell\}$. That is, we will show that either $\text{gen}(H) < \text{gen}(H \cup H_k \cup H_\ell)$, or $H \cup H_k \cup H_\ell$ has a degenerate face.

The connectedness of H_k and H_ℓ, and the fact that each of these graphs has at least one vertex in common with H, implies that there is a sequence $L_0, L_1, \ldots, L_m$ of subgraphs of $H \cup H_k \cup H_\ell$ such that the following hold: (i) $L_0 = H \cup L$; (ii) $L_m = H \cup H_k \cup H_\ell$; (iii) for $i = 0, 1, \ldots, m-1$ there is an edge e_i of $H_k \cup H_\ell$ such that $L_i = L_{i+1} - e_i$. Roughly speaking, starting from $H \cup H_k \cup H_\ell$ we can obtain $H \cup L$ by successively removing edges (if an edge is incident with a degree one vertex, we also remove that vertex, as we mentioned before the Remark above), so that at every step we have a connected graph.

If $L_m = H \cup H_k \cup H_\ell$ has a degenerate face then we are done (as then (ii) holds in the statement of Lemma 4). Thus we assume that L_m has no degenerate face. Let j be the smallest integer in $\{0, 1, \ldots, m\}$ such that L_j has no degenerate face. By assumption L_0 has a degenerate face, so $j \geq 1$. Thus L_{j-1} does have a degenerate face, and since $L_{j-1} = L_j - e_{j-1}$, we can apply the Remark above, obtaining that $\text{gen}(L_j) > \text{gen}(L_{j-1})$. Since $H \subseteq H \cup L \subseteq L_{j-1}$ and $L_j \subseteq H \cup H_k \cup H_\ell$, it follows that $\text{gen}(H) \leq \text{gen}(L_{j-1}) < \text{gen}(L_j) \leq \text{gen}(H \cup H_k \cup H_\ell)$, and so (i) in the statement of Lemma 4 holds. □

3.4.2 Proof of Claim B

Proof. Since $\bigcup \mathcal{H}$ is cellularly embedded, and F is a face of H with positive genus, it follows that there exist edges of $H_1 \cup H_2 \cup \cdots H_n$ contained in F. By relabelling if necessary, we can assume that for some $m \leq n$, $H_1, H_2, \ldots, H_m$ are the graphs in $\mathcal{H} \setminus \{H\}$ that contain at least one edge in F.

Now for each H_i with $i \in \{1, 2, \ldots, m\}$, let I_i be the subgraph of H_i induced by the edges of H_i inside the face F. Thus $\{I_1, \ldots, I_m\}$ is a collection whose union is contained in F, except for the *attachment* vertices, that is, those vertices in the graphs I_i that are in W (and thus in H). Note that each $I_i \in \{I_1, \ldots, I_m\}$ has at least one vertex of attachment: this follows since each $H_i \in \mathcal{H}$ is connected, and has at least one vertex in common with the anchor H: if some I_i had no vertices of attachment, then it (and hence H_i) would not intersect H. For convenience, although not all elements in the cluster necessarily have an edge inside F, moreover I_i is note necessarily connected, we choose to keep the labels of the vertices and edges in $I_1 \cup I_2 \cup \ldots \cup I_m$ as they are inherited from $H_1 \cup H_2 \cup \cdots H_m$, so that $I_i \subseteq H_i$.

Note that even though each I_i has at least one vertex in common with H, the collection $\{H, I_1, I_2, \ldots, I_m\}$ may not be a cluster of graphs, since the graphs I_i are not necessarily connected. However, this is not relevant to our purposes.

We now collapse W to a point u, obtaining the compact surface $\Sigma := F \cup \{u\}$ (we discard $\Sigma_g \setminus F$). Thus Σ has the same genus as F, and $I_1 \cup I_2 \cup \cdots \cup I_m$ naturally induces an embedded graph K in Σ: every edge and vertex of $I_1 \cup I_2 \cup \cdots \cup I_m$ is maintained, with the exception of the attachment vertices, which get identified to the single vertex u. For each $i = 1, 2, \ldots, m$, we let J_i be the subgraph of K naturally induced by I_i.

By letting J be the graph that consists solely of vertex u, the collection $\mathcal{J} := \{J, J_1, J_2 \ldots, J_m\}$ is then a cluster of graphs in Σ_h, as the property that each I_i had at least one attachment vertex implies that each J_i contains u, that is, intersects the anchor J. With the exception of u, each vertex or edge of J_i is inherited from a vertex or edge of I_i (and hence of H_i), and we choose to maintain their respective labels.

The cellularity of $\bigcup \mathcal{H}$ implies that $\bigcup \mathcal{J}$ is also cellularly embedded, and so we may apply Proposition 5 to $\mathcal{J}$, to obtain that there exists a subcollection $\mathcal{J}_\circ$ of $\mathcal{J}$, with $|\mathcal{J}_\circ| \leq 2$, such that $\bigcup \mathcal{J}_\circ$ contains a non-separating cycle. Thus there exist (not necessarily distinct) integers k, ℓ such that $J_k \cup J_\ell$ contains a non-separating cycle D.

To finish the proof it suffices to look at the subgraph D' of $\bigcup \mathcal{H}$ (back in Σ_g) induced by the edges of D. Note that D' is contained in $H_k \cup H_\ell$. If D contains u, then D' is either an F-non-separating path or a non-separating cycle that has a single vertex in common with W. If D does not contain u, then D' is a cycle completely contained in F. Thus the subcollection $\mathcal{H}_\circ := \{H_k, H_\ell\}$ of $\mathcal{H}$ has the required properties. □

3.5 Towards the proof of Proposition 5: short-circuiting non-separating cycles

In the context of the statement of Proposition 5 we have a cluster of graphs $\mathcal{G} = \{G, G_1, \ldots, G_n\}$, cellularly embedded in some surface Σ_g with $g \geq 1$. The graph G is an anchor of the cluster, and it consists of a single vertex a. Since G is edgeless, it follows that each edge e of $\bigcup \mathcal{G}$ belongs to G_i for exactly one $i \in \{1, \ldots, n\}$. To help comprehension, we say that e is *of colour* i. Thus each edge of $\bigcup \mathcal{G}$ has exactly one colour. A subgraph of $\bigcup \mathcal{G}$ is k*-coloured* if the number of distinct colours in its edge set is exactly k.

For brevity, we will refer to a non-separating cycle simply as an *ns-cycle*. Under this terminology, Proposition 5 claims the existence of a 1- or 2-coloured ns-cycle in $\bigcup \mathcal{G}$.

The existence of an ns-cycle follows since $\bigcup \mathcal{G}$ is cellularly embedded in Σ_g for some $g \geq 1$ (see [41, Lemma 11]). If such a cycle is k-colored for some $k \geq 3$, we need to find a way to "short-circuit" it to find an ns-cycle with fewer colours. (This short-circuiting idea is also central for the proof of the main theorem in [41]). The proof of Proposition 5 consists of iteratively applying this short-circuiting process, until we obtain a 1- or 2-coloured ns-cycle.

The central idea behind the short-circuiting process is that the set of ns-cycles in an embedded graph satisfies Thomassen's 3-*path-condition* [47, Proposition 3.5]: if R, S, T are pairwise internally disjoint paths with the same endpoints, and the cycle $R \cup S$ is non-separating, then one of the cycles $R \cup T$ and $S \cup T$ is non-separating. Thus if we start with an ns-cycle $R \cup S$ and a suitable path T internally disjoint from $R \cup S$ (where both endpoints of T are in $R \cup S$), we can apply the 3-path-condition to find an ns-cycle with fewer colours than $R \cup S$.

In the standard graph theory terminology, a *trail* is a walk in which no edge appears more than once. If the startpoint u and the endpoint v of a trail T are distinct, then T is a *uv-trail*. A *circuit* is a trail whose startpoint and endpoint are the same. If $W = v_0 e_1 v_1 \dots e_n v_n$ is a walk on a graph, then W^{-1} is the *reverse* walk of W, namely $v_n e_n v_{n-1} \dots v_1 e_1 v_0$. If W' is a walk $v_n e_{n+1} v_{n+1}, \dots, e_m v_m$ (the endpoint of W is the startpoint of W'), then WW' is the *concatenation* $v_0 e_1 v_1 \dots e_n v_n e_{n+1} v_{n+1} \dots e_m v_m$ of W and W'.

If T is not internally disjoint from $R \cup S$ (or even if T is not a path, but a trail, and/or $R \cup S$ is not a cycle but a circuit), we cannot apply the 3-path-condition. Thus we need a version of the 3-path-condition that applies to trails (instead of paths) and circuits (instead of cycles). As we shall see shortly, a property totally analogous to the 3-path-condition holds in the context of trails and circuits, by considering in this more general context (instead of non-separating cycles) the collection of non-null-homologous circuits in an embedded graph. Before moving on to this generalized version of the 3-path-condition, we recall the concepts of a trail and a circuit.

We adopt the (usual) point of view that a circuit is regarded as a cyclic sequence of vertices and edges, so that if $C = v_0 e_1 v_1 \dots e_n v_0$ is a circuit, then C is identical to the circuit $v_i e_{i+1} v_{i+1} \dots v_0 e_1 \dots e_{i-1} v_i$, for all $i = 1, \dots, n-1$. If $C = v_0 e_1 v_1 \dots e_i v_i e_{i+1} \dots e_n v_0$ is a circuit such that $v_0 = v_i$ for some $i \neq 0$, then $v_0 e_1 v_1 \dots v_i = v_0$ is a *subcircuit* of C.

We extend the notion of an ns-cycle to circuits, by means of simplicial homology over $\mathbb{Z}_2$. From this viewpoint, a cycle is an ns-cycle if (and only if) it is non-null-homologous. Thus we say that a circuit is an *ns-circuit* if it is non-null-homologous. The following trivial observation from elementary homology theory will be repeatedly invoked in the short-circuiting iterative process in the proof of Proposition 5. Nonetheless we only invoke these results, we refer the interested reader to [25] for further information on this subject.

Remark. Every ns-circuit contains an ns-cycle as a subcircuit.

We are now ready to state the extension (to trails and circuits) of the fact that the set of ns-cycles in an embedded graph satisfies Thomassen's 3-path-condition.

Observation 6 (3-trail condition for ns-circuits)**.** *Let T_1, T_2, T_3 be edge-disjoint trails in an embedded graph, with the same startpoint and the same endpoint. If $T_1 T_2^{-1}$ is an ns-circuit, then at least one of $T_1 T_3^{-1}$ and $T_3 T_2^{-1}$ is also an ns-circuit.*

Some variants of this observation are usually stated without proof (see for instance [9, Section 3.1]), as it is a trivial exercise in homology theory. We give the proof for completeness.

Proof. Regarding the circuits as 1-chains, we have that $T_1T_3^{-1} + T_3T_2^{-1} = T_1T_2^{-1}$, since T_3 and T_3^{-1} cancel each other. Since by assumption $T_1T_2^{-1}$ is non-null-homologous, it follows that at least one of $T_1T_3^{-1}$ and $T_3T_2^{-1}$ must also be non-null-homologous. □

With Observation 6 in our toolkit, we are finally ready to prove Proposition 5.

3.6 Proof of Proposition 5

Proof. Let $G_1, \ldots, G_n$ be the elements in $\mathcal{G} \setminus \{G\}$. As we mentioned in the previous section, to help comprehension we say that the edges of G_i are *of colour* i, for $i = 1, 2, \ldots, n$. Since G consists of a single vertex a, and $G_1, G_2, \ldots, G_n$ are pairwise edge-disjoint, it follows that each edge of $\bigcup \mathcal{G}$ has exactly one colour. If T is a trail in $\bigcup \mathcal{G}$ with all the edges of the same colour i for some $i \in \{1, 2, \ldots, n\}$, then T is *monochromatic*, and we say that i *is* the colour of T.

If C is a circuit in $\bigcup \mathcal{G}$, then C can be written as a concatenation $T_0T_1 \ldots T_{r-1}$ of maximal monochromatic trails. That is, for $i = 0, 1, \ldots, r-1$, T_i is monochromatic, and the colour of T_i is distinct from the colour of T_{i+1} (indices are read modulo r, and so T_{r-1} and T_0 are of different colours). We remark that T_i and T_j may be of the same colour for some $i \neq j$, as long as $j \notin \{i-1, i+1\}$. This decomposition of C as a concatenation of maximal monochromatic trails is unique, up to a cyclic permutation of the trails. This uniqueness allows us to call $T_0T_1 \ldots T_{r-1}$ the *canonical* decomposition of C; we call r the *rank* of the circuit C.

To prove Proposition 5 we show that there exists an ns-cycle of rank at most 2 (see Statement (4) below), and therefore a subset of $\{G_1, \ldots, G_n\}$ of size at most 2, whose union contains an ns-cycle, as required in the statement of the proposition.

Thus the final goal is to prove Statement (4) below. To help comprehension, we break the proof into several steps. As we will see, showing the existence of an ns-cycle of rank at most 3 is fairly easy (see Statement (2) below). Most of the work is involved with bringing the rank down to at most 2.

(1) *If there exists an ns-circuit of rank r, then there exists an ns-cycle of rank at most r.*

Proof. This follows immediately from the definition of an ns-circuit. Indeed, if C is an ns-circuit of rank r, then it contains an ns-cycle D as a subcircuit (see Remark before Observation 6); it is readily checked that the rank of D is at most r. It is worth noting that the property that D is a subcircuit of C (and not just an arbitrary ns-cycle contained in C) is essential in order to guarantee that the rank of D is at most the rank of C. □

(2) *There exists an ns-cycle of rank at most* 3.

Proof. The existence of an ns-cycle in $\bigcup \mathcal{G}$ follows from [41, Lemma 11], since $\bigcup \mathcal{G}$ is cellularly embedded in a surface of positive genus. In order to prove (2) it suffices to show that if D is an ns-cycle with canonical decomposition $P_0P_1 \ldots P_{r-1}$, where $r \geq 4$, then there exists an ns-cycle whose rank is smaller than r; an iterative application of this fact, starting with an arbitrary ns-cycle, yields the existence of an ns-cycle with rank at most 3.

Suppose first that there exists an $i \in \{1, \dots, n\}$ such that there are at least two paths in $\{P_0, P_1, \dots, P_{r-1}\}$ that are of colour i (note that since D is a cycle, the elements in its canonical decomposition $P_0 P_1 \dots P_{r-1}$ are paths). Since G_i is connected, it follows that there exist distinct P_j, P_k in the decomposition, both of colour i, and a path R of colour i whose startpoint u is in P_j and whose endpoint v is in P_k, and such that R does not contain any edge of D. Now let P, Q be the two uv-paths contained in D (thus $D = PQ^{-1}$). Then P, Q, and R are pairwise edge-disjoint uv-paths. It is readily verified that since R is of colour i, then the rank of each of PR^{-1} and RQ^{-1} is strictly smaller than r. Now since $D = PQ^{-1}$ is an ns-cycle, and in particular an ns-circuit, it follows from Observation 6 that one of PR^{-1} and RQ^{-1} is also an ns-circuit. Thus, one of these is an ns-circuit (and by (1), an ns-cycle) whose rank is smaller than r.

Suppose finally that all the paths $P_0, P_1, \dots, P_{r-1}$ are of distinct colours. By relabelling if necessary, we may assume that P_0 is of colour 0, and P_2 is of colour 2. Since G_0 and G_2 are connected and have at least one vertex in common (namely the vertex a in the anchor G), it follows that there exists a path U with the following properties:

(i) one endpoint v_0 of U is in P_0, and its other endpoint v_2 is in P_2;

(ii) U is the concatenation of a path of colour 0 with a path of colour 2 (one of these two paths may consist of a single vertex); and

(iii) U is edge-disjoint from D.

Now let S, T be the $v_0 v_2$-paths contained in D. It is easily seen that the rank of both SU^{-1} and UT^{-1} is smaller than r. Now since $D = ST^{-1}$ is an ns-cycle (and in particular an ns-circuit), it follows from Observation 6 that one of SU^{-1} and UT^{-1} is also an ns-circuit. Thus there exists an ns-circuit (and by (1), an ns-cycle) whose rank is smaller than r. □

The following statement gets us to the final goal (the existence of an ns-cycle with rank at most 2) in a particular case. Since a reduction to this case appears several times in the proof (4), it is convenient to deal with it before moving on to (4).

(3) *Let C be an ns-circuit with canonical decomposition $T_1 T_2 T_3$, where the colour of T_i is i, for $i = 1, 2, 3$. Suppose that T_2 does not contain the startpoint v of T_1 (which is the endpoint of T_3), but there is some edge of colour 2 incident with v. Then there is an ns-cycle with rank at most 2.*

Proof. We start by noting that the connectedness of G_2 and the assumption that there is an edge of colour 2 incident with v, imply that there is a path U of colour 2, that starts in v and ends in v_2 in T_2, which is edge-disjoint from C. Now let S be the subtrail of C obtained by starting at v, traversing T_1 completely, and then continuing along T_2 until we reach v_2 (it might be that v_2 is the endpoint of T_1, in which case S does not contain edges of T_2, but this is irrelevant). Now let T be the trail from v to v_2 such that $C = ST^{-1}$; thus

T^{-1} is obtained by continuing the traversal of C after we reached v_2, and in particular T_3 is a subtrail of T^{-1}.

The circuit SU^{-1} is the concatenation of T_1 (which has colour 1) with a trail of colour 2, and the circuit UT^{-1} is the concatenation of a trail of colour 2 with T_3 (which has colour 3). Thus both circuits have rank exactly 2. Since $ST^{-1} = C$ is an ns-circuit, it follows by Observation 6 that at least one of SU^{-1} and UT^{-1} is an ns-circuit. Thus there exists an ns-circuit of rank 2, and by (1) it follows that there is an ns-cycle of rank at most 2. □

(4) *There exists an ns-cycle with rank at most* 2.

Proof. By (2), there exists an ns-cycle D with rank at most 3. If the rank of D is 1 or 2 we are obviously done, so we assume that the rank of D is exactly 3. Thus D has a canonical decomposition $P_1P_2P_3$, where P_i is a path for $i = 1, 2, 3$. By relabelling the subgraphs $G_1, G_2, \dots, G_n$ if necessary, we may assume that P_i is in G_i (that is, P_i has colour i), for $i = 1, 2, 3$. We recall that G_1, G_2, and G_3 have at least one common vertex, namely the vertex a in the anchor G.

Suppose first that a is in D. By relabelling if necessary, we may assume that a is in P_3, and possibly also in P_1 but not in P_2. If a is the endpoint of P_3 (and thus the startpoint of P_1) then we are done by applying (3) with $T_i := P_i$ for $i = 1, 2, 3$. So we may assume that a is an internal vertex of P_3. Since G_1 is connected, there is a path R of colour 1 from a to a vertex v_1 in P_1, such that R has no edges in common with D. Let P, Q be the two av_1-paths contained in D, labelled so that every edge of P is of colour 1 or 3 (hence Q contains P_2). It is easily checked that PR^{-1} has rank 2, and RQ^{-1} has rank at most 3. Now $PQ^{-1} = D$ is an ns-cycle (in particular an ns-circuit), and so by Observation 6 one of PR^{-1} and RQ^{-1} is an ns-circuit. If PR^{-1} is an ns-circuit we are done, since it has rank 2. Thus we may assume that RQ^{-1} is an ns-circuit, and that its rank is exactly 3. Then RQ^{-1} is the concatenation of three trails: (i) a trail T_1 of colour 1, which is the concatenation of R with the subpath of P_1 from v_1 to the endpoint of P_1 (this last subpath may consists of the single vertex v_1); (ii) the trail $T_2 = P_2$, of colour 2; and (iii) the subpath of P_3 that starts at the startpoint of P_3 and ends at a; this last trail is of colour 3, and cannot consist of a single vertex, since a is an interior vertex of P_3. Since T_1, T_2, T_3 satisfy the conditions in (3), it follows that there exists an ns-cycle with rank at most 2, as required.

Finally suppose that a is not in D. We may assume that no vertex in D is in $G_1 \cap G_2 \cap G_3$ (a vertex with this property need not be unique), for if such a vertex exists, we let it play the role of a and we are done by the discussion above.

Since each of G_1, G_2 and G_3 is connected, it follows that for $i = 1, 2, 3$ there exists a path Q_i of colour i with startpoint v_i in P_i and endpoint a, where Q_i is edge-disjoint from D. Note that v_1, v_2, v_3 cannot all be the same vertex, since $D = P_1P_2P_3$ is a cycle. By relabelling if necessary, we may assume that $v_1 \neq v_3$. Let $U := Q_3Q_1^{-1}$, and let S, T be the two paths from v_3 to v_1 contained in D, where every edge of S is of colour 1 or 3 (thus T contains P_2). We note that the circuit SU^{-1} has rank 2 (its canonical decomposition consists of a trail of colour 1 followed by a trail of colour 3), and $U^{-1}T$ has rank 3 (its canonical decomposition consists of a trail T_1 of colour 1, followed by $T_2 = P_2^{-1}$ of colour 2, followed by a trail T_3 of colour 3). Since $ST^{-1} = D$ is an ns-cycle (and thus an ns-circuit) it follows from Observation 6 that at least one of SU^{-1} and UT^{-1} is an ns-circuit. In the former case we are done, since SU^{-1} is then an ns-circuit of rank 2, and by (1) there exists an ns-cycle of rank at most 2. In the latter case, UT^{-1} is an ns-circuit of rank 3 whose canonical decomposition $T_1T_2T_3$ described above satisfies the conditions in (3). Therefore also in this case there exists an ns-circuit (and by (1), an ns-cycle) of rank at most 2. □

As we observed before (1), Statement (4) completes the proof of the lemma. □

Chapter 4

Concluding remarks

It is natural to ask if the condition that there is a pseudocircle that intersects all other pseudocircles in the collection is absolutely necessary. To answer this question we note that it is necessary to require some sort of condition along these lines. Indeed, as observed by Ortner in [41, Figure 16], there exist arbitrarily large collections of pseudocircles (whose union is connected) that cannot be embedded into a sphere, and yet the removal of any pseudocircle leaves an arrangement that can be embedded into a sphere.

On the other hand, in order to have some version of Theorem 1 it is not strictly necessary to have a single pseudocircle intersecting all the others; our techniques and arguments are readily adapted under the assumption that there is a subcollection of bounded size that gets intersected by all other pseudocircles. More precisely, if we define an *m-arrangement* of pseudocircles as a collection in which there is a subcollection of size (at most) m such that every pseudocircle intersects at least one pseudocircle in this subcollection, then it is easy to show that the corresponding version of Theorem 1 reads as follows.

Theorem 7. *An m-arrangement of pseudocircles is embeddable into Σ_g if and only if all of its subarrangements of size at most $4g + (m + 5)$ are embeddable into Σ_g.*

We have proved that a strong arrangement of pseudocircles is embeddable into Σ_g if and only if all of its subarrangements of size at most $4g + 4$ are embeddable into Σ_g. As Ortner showed in [41, Figure 3], there are strong arrangements of size 4 that are not embeddable into a sphere, and yet all its subarrangements are embeddable into a sphere; thus this result cannot be improved for $g = 0$. Similarly, for arrangements of pseudocircles (with the more general definition we used throughout this work) the size bound $4g+5$ cannot be improved for the case $g = 0$. Indeed, the toroidal arrangement shown in Figure 4.1 has 5 pseudocircles, it cannot be embedded into the sphere, and yet all its subarrangements of size 4 can be embedded into the sphere.

Working under the framework of clusters of graphs, one can prove similar results to Theorem 1 for collections of other objects, such as *arcs*, which are homeomorphic images of the interval $[0, 1]$. Arrangements of arcs (and, in general, arrangements of curves) are investigated in [20] and [21]. (A different notion of an arrangement of arcs is used, for

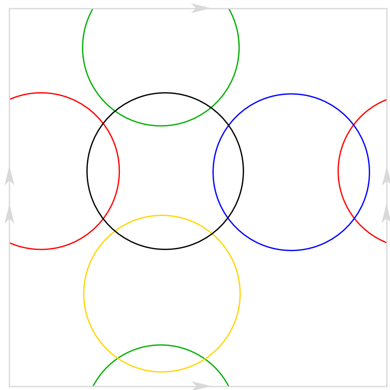


Figure 4.1: An arrangement of 5 pseudocircles in the torus (given in its polygonal representation). This arrangement is not embeddable into the sphere, but every subarrangement of size 4 is embeddable into the sphere.

instance, in [37]). In order to obtain a result along the lines of Theorem 1, some anchorness condition for an arrangement of arcs is required (a discussion analogous to the one given at the beginning of this section applies to these objects as well). If we consider an arrangement of arcs as a collection of arcs that pairwise intersect a finite number of times, and in which there is an arc that intersects all the other arcs in the collection, then the following analogue of Theorem 1 is a consequence of our Main Theorem.

Theorem 8. *An arrangement of arcs is embeddable into Σ_g if and only if all of its subarrangements of size at most $4g + 5$ are embeddable into Σ_g.*

We finally note that in [41], Ortner wrote that one could conjecture that embeddability (of a strong arrangement) into the surface Σ_g of genus g holds if and only if all $(4 + g)$-subarrangements are embeddable into Σ_g. We have proved that, for strong arrangements, embeddability into Σ_g holds if and only if all $(4+4g)$-subarrangements are embeddable into Σ. The question of whether or not this can be improved to $4 + g$, as conjectured in [41], remains open.

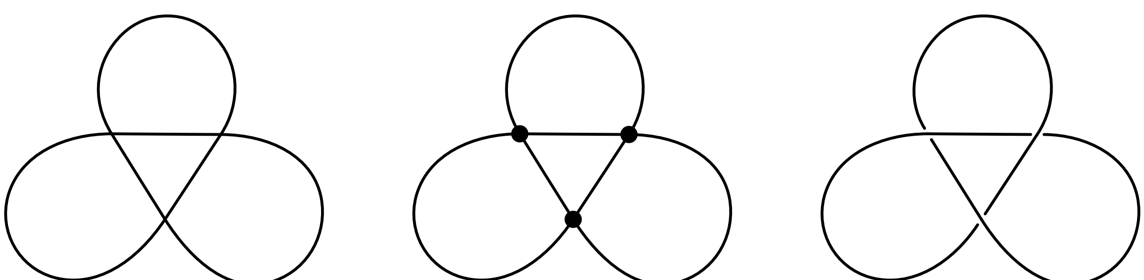
Chapter 5

Knots, shadows, and diagrams

We start the second part of this thesis with a review of some basic notions in knot theory, such as knots, projections, shadows, and diagrams. For a comprehensive introduction to knot theory we refer the reader to the standard reference [3].

5.1 Knots

A *knot* is a simple closed curve in $\mathbb{R}^3$, that is, the image of an injective continuous mapping $f : S^1 \to \mathbb{R}^3$. The usual way to visualize a knot K is by projecting K onto a plane.



(a) (b) (c)

Figure 5.1: In (a) we illustrate a *projection* of the trefoil knot; this is a (non-simple) closed curve, with three *crossing points*. In (b) we turn the closed curve into a combinatorial-topological object (a 4-regular graph) by regarding the crossing points as vertices; the result is a *shadow*. In (c) we have the familiar *diagram* of the trefoil knot, that is, a projection with over/under crossing information provided. The over/under crossing information in a diagram suffices to recreate the knot, in the sense that from this we can obtain a knot that is equivalent to the original knot.

As an example, in Figure 5.1(a) we illustrate a projection of the trefoil knot. For some purposes it is useful to regard a projection as a topological-combinatorial object, that is, as a plane graph. Such an object is a *shadow* of the knot. In Figure 5.1(b) we illustrate the shadow corresponding to the projection in Figure 5.1(a). These are essentially the same 2-dimensional object, but in the shadow we emphasize that the crossing points of the projection are regarded as vertices of a plane graph.

Some essential information about a knot (a 3-dimensional object) is inevitably lost if all we have available is a projection (or, for that matter, a shadow) of the knot. In order to recreate the knot, one provides the over/under crossing information at each crossing point of the projection; the resulting object is a knot *diagram*. In Figure 5.1(c) we show a diagram of the trefoil knot.

This informal description of shadows and diagrams is enough for many discussions. However, for our purposes we need to go a bit further, and introduce and discuss several concepts related to shadows and diagrams.

5.2 Projections

Let P be a plane in $\mathbb{R}^3$. The *orthogonal projection* $\pi_P : \mathbb{R}^3 \to P$ *onto* P is the mapping that takes each point q in $\mathbb{R}^3$ to the point p in P such that the line orthogonal to P that passes through p contains q. If K is a knot, then the *orthogonal projection* (or simply the *projection*) *of* K *onto* P is $\pi_P(K)$.

The most useful and natural projections have the properties that (i) no three points on the knot project to the same point; and (ii) the number of pairs of points in the knot that project to the same point is finite. Formally, for these projections we have that (i) for every $p \in P$, $|(\pi_P)^{-1}(p)|$ is either 0, 1, or 2; and (ii) the number of points in P such that $|\pi_p^{-1}(p)| = 2$ is finite. Such a projection is called *regular*. For instance, the projection on Figure 5.1 (a) is regular.

There exist knots that do not have a regular projection (such as wild knots). However as stated in [14], the main invariants of knot type, such as knot polynomials are not always defined for wild knots but for tame knots ($\mathcal{C}^1$-knots) and as they also proved, for every tame knot K there exists an homeomorphism of $\mathbb{R}^3$ onto itself which maps K onto a polygonal knot K' which has a regular projection. This allows us to adopt the usual point of view of considering only knots that admit regular projections. From now on, we work under the usual assumption that all knot projections under consideration are regular.

A knot projection is thus a closed curve, that is, the image of a continuous mapping $g : S^1 \to \mathbb{R}^2$ (we assume without loss of generality that the plane under consideration is the xy-plane). If a projection of a knot K is a simple closed curve then K is obviously the unknot. If the projection is not a simple closed curve, then (by the regularity assumption) there is a finite number of points p in the plane such that $|g^{-1}(p)| = 2$. These are the *crossing* points of the projection.

5.3 Shadows

As we mentioned above, it is often convenient to regard a projection as a topological-combinatorial object, namely as a plane graph. For this, we simply regard the crossing points as vertices. The result is then a 4-regular plane graph, which is a *shadow* of the knot under consideration. For convenience, we admit the possibility that a shadow is a vertex-less graph, which is what we obtain if the projection is a simple closed curve (and so the corresponding knot is trivial). We refer to a vertex-less shadow as a *trivial* shadow.

Thus every shadow is a 4-regular plane graph, but not every 4-regular plane graph is a shadow of a knot. We now describe the combinatorial characterization of which 4-regular plane graphs do arise as knot shadows.

If G is a plane graph (that is, a graph with a given planar embedding), then the *rotation* at a vertex v of G is a cyclic permutation of the edges incident with v; this cyclic rotation records the clockwise order in which these edges leave v in the embedding. We refer the reader to Figure 5.2.

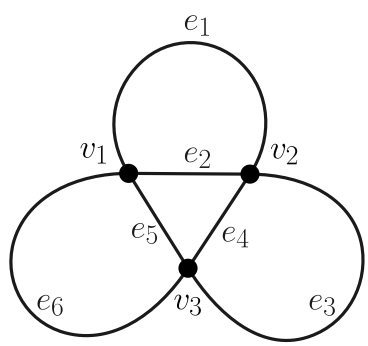


Figure 5.2: The rotation of vertex v_1 in this plane graph is the cyclic permutation $e_1e_2e_5e_6$; the rotation of v_2 is the cyclic rotation $e_2e_1e_3e_4$; and the rotation of v_3 is $e_4e_3e_6e_5$.

We recall that a *walk* in a graph G is a sequence $v_0e_1v_1 \ldots e_nv_n$, where v_i is a vertex of G for $i = 0, 1, \ldots, n$, and for $j = 1, 2, \ldots, n$, e_j is an edge of G whose endpoints are v_{j-1} are v_j (if e_j is a loop-edge, then its incident vertex is $v_{j-1} = v_j$). A walk is *closed* if its initial and final vertex are the same. A walk is *Eulerian* if it traverses each edge of the graph exactly once.

A *straight-ahead* walk in a plane graph G is a walk that always passes from an edge to the opposite edge in the rotation at each vertex. That is, if a rotation at vertex v is $e_1e_2e_3e_4$, then in a straight-ahead walk W that goes from e_1 towards v, the next edge in the walk (if any) is necessarily e_3; if W goes from e_2 towards v, then the next edge in the walk (if any) is necessarily e_4, etc. If W is a straight-ahead Eulerian closed walk of G, then v appears exactly twice in W: in one occurrence either e_1ve_3 or e_3ve_1 will appear, and in the other occurrence one of e_2ve_4 and e_4ve_2 will appear. In the example of Figure 5.2, the walk

$v_1e_1v_2e_4v_3e_6v_1e_2v_2e_3v_3e_5v_1$ is a straight-ahead closed Eulerian walk. We note that if G has an Eulerian straight-ahead closed walk, then for every vertex v of G there is an Eulerian straight-ahead closed walk that has v as its startpoint (and as its endpoint).

The characterization of when a 4-regular plane graph is a shadow of a knot is totally straightforward, but it is so important that it is worth highlighting:

Remark. *Let G be a 4-regular plane graph. Then G is a shadow of a knot if and only if G has a straight-ahead Eulerian closed walk.*

We also recall that if C is a connected subgraph of a graph G, and every vertex of C has degree two (in C), then C is a *cycle* of G. For our purposes, a vertex-less graph (which, we recall, consists of a single edge, homeomorphic to a simple closed curve) will be considered a cycle.

A *straight-ahead* cycle in a shadow S is a cycle C of S with the following property. There exists a vertex c (the *root*) of C such that there is a straight-ahead closed walk in S that starts (and ends) in c, and traverses the edges of C, and no other edges of S. In Figure 5.3 we illustrate two straight-ahead cycles in a shadow of the trefoil knot.

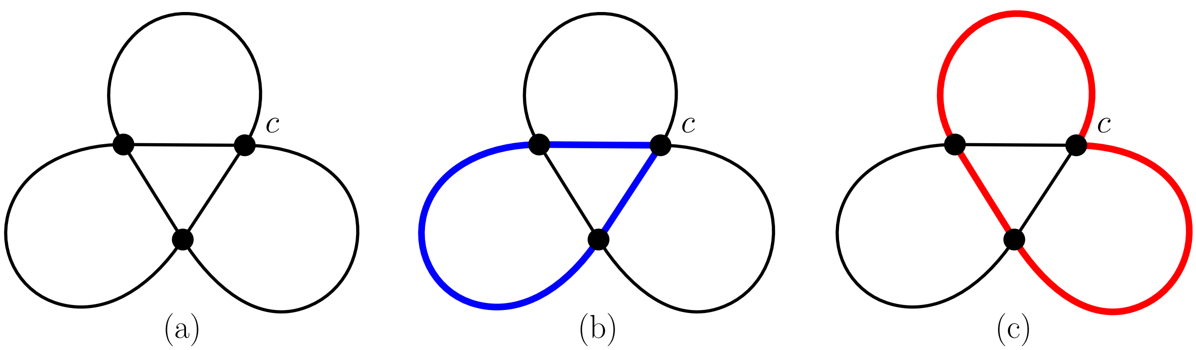


Figure 5.3: In (a) we depict a shadow of the trefoil knot. There are two straight-ahead cycles whose root is the vertex c. These cycles are highlighted in parts (b) and (c) of this figure.

We note that in the case of Figure 5.3 there is a vertex c with the property that there are two distinct straight-ahead cycles that have c as their root. (Actually, by symmetry, all three vertices have this property). In an arbitrary shadow, the vertex a (the root of two distinct straight-ahead cycles) needs not to exists. Moreover, for an arbitrary vertex v in an arbitrary shadow, there is no reason why there must exist a straight-ahead cycle with root v. This is illustrated in Figure 5.4. In this shadow, there are exactly two straight-ahead cycles, one with root v_1, and one with root v_3. In particular, there is no straight-ahead cycle that has v_2 as its root.

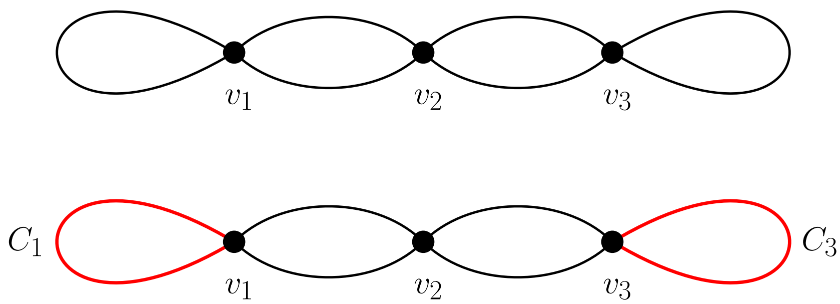


Figure 5.4: In the upper shadow of this figure, there are only two straight-ahead cycles C_1 and C_3, which are highlighted on the lower hand side.

We will make extensive use of the following elementary facts:

1. Let v be a vertex of a shadow S. Then there are exactly two straight-ahead closed walks W_1, W_2 that have v as their startpoint (and endpoint), and that do not contain v as an internal vertex. These walks are edge-disjoint, and if we regard them as subgraphs of S, then it is the union of the graphs $S = W_1 \cup W_2$.

2. Every nontrivial shadow has at least two distinct (necessarily edge-disjoint) straight-ahead cycles.

The first statement follows immediately from the fact that if v is a vertex of a shadow S, then there is an Eulerian straight-ahead closed walk W that starts and ends at v: this walk W is the concatenation of two walks W_1, W_2 with the given properties. In the smallest case, that is, when v is the only vertex of S, the two walks W_1 and W_2 are also straigh-ahead cycles of S. The second statement is easily proved by induction on the number of vertices in the shadow.

Later on we will investigate the role of cut vertices in a shadow; let us finish this discussion on shadows by recalling the concept of a cut-vertex in a graph. A 1-*separation* of a connected graph G is an ordered pair (H, K) of subgraphs of G, each having at least one edge, such that $H \cup K = G$ and $H \cap K$ is a graph that consists of a single vertex. The vertex of $H \cap K$ is the *cut-vertex* of the 1-separation.

5.4 Diagrams

A shadow contains only partial information of a knot, in the sense that if we are given only the shadow of a knot, then in general we cannot recreate the knot. The additional information required to reconstruct the knot is provided by a diagram, the concept we now proceed to further describe and explore.

A diagram D consists of a shadow S for which we provide the over/under-crossing information at each vertex. That is, for each vertex it is prescribed which strand goes locally above (or, equivalently, is the *overpass*) and which strand goes locally below (or, equivalently, is the *underpass*). We call this information the *prescription* at the vertex. At each vertex we have two possible prescriptions (corresponding to which of the strands is an overpass at the corresponding crossing). The collection of prescriptions for all the vertices yields an *assignment* for S, resulting in a diagram D that *has S as its underlying shadow*. If S has n vertices, then there are 2^n possible assignments for S, each one corresponding to exactly one diagram, and so there are 2^n distinct diagrams that have S as their underlying shadow.

Besides saying that D has S as its underlying shadow, we describe this relationship between S and D in other several natural ways. For instance, sometimes we simply say that *D is a diagram of S*. Often we write that S is *the shadow* associated to D, and we also say that D is a diagram *associated to S*.

In the description above (as in many knot theory papers and monographs) we have used the term strand without further explanation, since its meaning is intuitively obvious in the context. In general, in this work we will adopt the viewpoint that a *strand* is a part of a diagram that corresponds to (the projection of) a proper connected subset of its corresponding knot. If one imagines the knot as a rope, then a strand is a part of the diagram that corresponds to the projection of a connected piece of the rope.

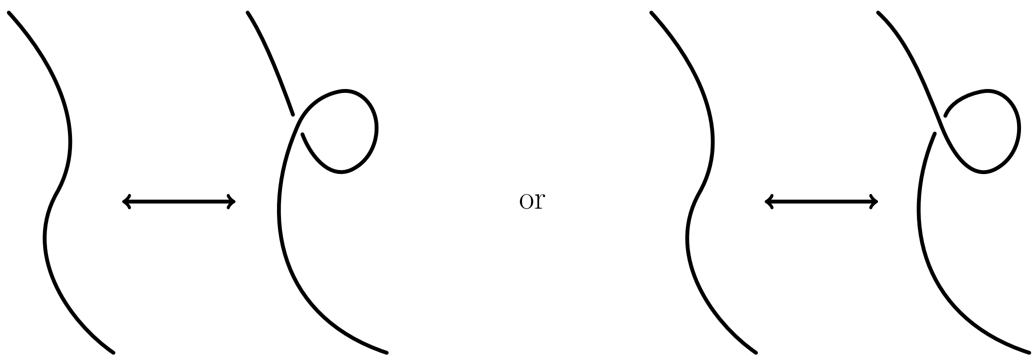
We note that the concept of a straight-ahead walk in a shadow S carries over naturally to a diagram D associated to S. Thus in a diagram we may speak of a *straight-ahead traversal* (we cannot use the term walk, since a diagram is not a graph). Moreover, there is no need to insist that the starting and ending point of a straight-ahead traversal be crossing points in D; these may be any two points in D. Thus we can equivalently say that a strand of a diagram D is simply a straight-ahead traversal in D, in which no non-crossing point occurs more than once. (If a non-crossing point occurred more than once, then we would be traversing the whole diagram, and then re-traversing some additional part of it). Note that we do not insist that the endpoints of a strand are distinct from each other.

Since a diagram consists of a shadow plus the prescription at each vertex, we may use the same labels of combinatorial objects of the underlying shadow (such as vertices and edges) for the diagram, keeping in mind that a vertex of the shadow corresponds to a crossing of the diagram. Thus, for instance, if v is a vertex of a shadow S, and D is a diagram of S, then, if no confusion arises, we may also use v to identify the corresponding crossing in D.

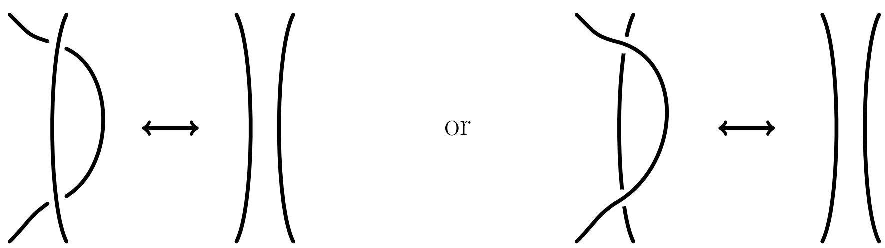
Two diagrams are *equivalent* if their corresponding knots are equivalent. If D is a diagram of the unknot, then we say that D is an *unknot diagram*. A diagram that contains no crossings is *trivial*. Evidently, every trivial diagram is an unknot diagram.

5.5 Reidemeister moves

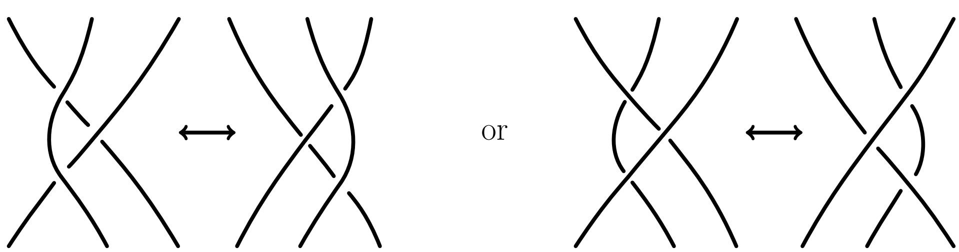
There are some important operations that transform a diagram into another diagram. One of these is a *planar isotopy*, which is simply a deformation of the diagram that can be obtained from a self-homeomorphism of the plane. Other very important operations are the *Reidemeister moves* of Types I, II, and III, which we illustrate in Figure 5.5.



(a) Type I Reidemeister move



(b) Type II Reidemeister move



(c) Type III Reidemeister move

Figure 5.5: The three types of Reidemeister moves.

A fundamental result in knot theory, proved independently by Reidemeister [44] and by Alexander and Briggs [6], is that if D and D' are equivalent diagrams, then one can take D to D' by a sequence of Reidemeister moves and planar isotopies.

5.6 Generalized Reidemeister moves

The three Reidemeister moves are particular instances of a more general operation that turns a diagram into an equivalent diagram. The following lively description of this operation is given in [22]: *"In practice it is often easier to redraw knots using the principle that any portion of a strand with only overcrossings may be replaced with another strand with the same endpoints and all new overcrossings, with the resulting breaks healing. Note that any such 'overpass move' can always be broken down into a sequence of Reidemeister moves and planar isotopies".*

This operation can be formalized as follows. Let x, y be (not necessarily distinct) points in a diagram D. Suppose that σ is a strand of D with endpoints x, y, with the property that as we traverse σ from x to y, all the crossings we find between x and y are overpasses (respectively, underpasses) for σ. Then we say that σ is an *over-strand* (respectively, *under-strand*) of D. Now suppose that D is a diagram, and σ is an over-strand (respectively, under-strand) of D with endpoints x, y. Let D' be a diagram obtained from D by removing σ, and replacing it by another over-strand (respectively, under-strand) with endpoints x, y. Then D and D' clearly are equivalent diagrams. We say that D' is obtained from D by a *generalized Reidemeister move.*

We note that we emphasized that x and y need not be distinct points of D. For instance, in Figure 5.6(a) we illustrate an overstrand (dotted) that has the crossing point x as both its initial and its final point. A valid generalized Reidemeister move is to replace this overstrand by the overstrand that consists only of x, thus obtaining the diagram in Figure 5.6(b). Thus the diagrams in Figure 5.6 are equivalent to each other.

It is easily seen that, in particular, each of the three classical Reidemeister moves (Figure 5.5) is a generalized Reidemeister move.

We note for later use that one can further extend this procedure to simplify a diagram as follows.

Remark. Let x, y be (not necessarily distinct) points in a diagram D. Suppose that σ is a strand of D with endpoints x, y, with the property that as we traverse σ from x to y, each crossing we find *for the first time* is an overpass (respectively, underpass) for the portion of σ we are traversing. Let D' be a diagram obtained from D by removing σ, and replacing it by any over-strand (respectively, under-strand) with endpoints x, y. Then D' can be obtained from D by a sequence of generalized Reidemeister moves and isotopies. In particular, D and D' are equivalent diagrams.

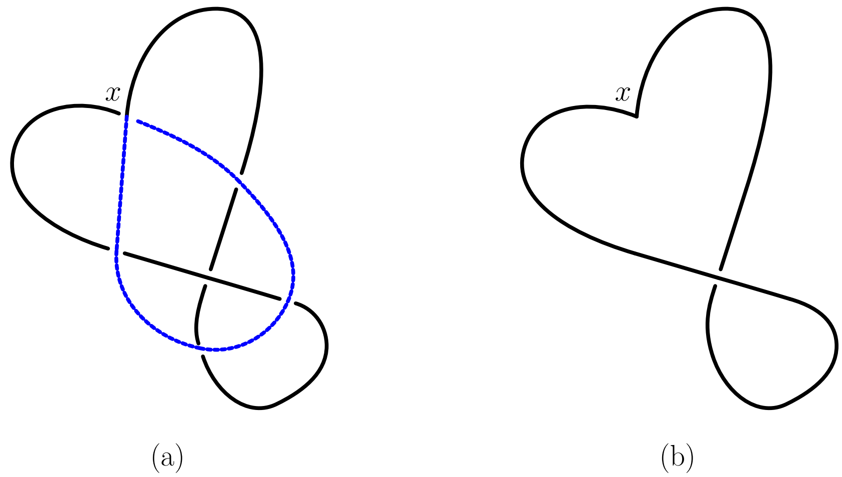


Figure 5.6: In the diagram in (a), the dotted strand that has the crossing point x as its initial point and endpoint is an overstrand. Thus we can apply a generalized Reidemeister move to this diagram, replacing this strand by the strand that consists solely of the point x, obtaining the diagram in (b). This is indeed a valid generalized Reidemeister move, since this one-point strand is also by definition an overstrand. Thus these two diagrams are equivalent.

We will make extensive use of generalized Reidemeister moves in this work. We note that the term "generalized Reidemeister move" is often used in the context of virtual knot theory [30]. Since we work exclusively with classic knots, there should be no confusion by this choice of terminology in our context.

Chapter 6

Unknot diagrams with the same underlying shadow

In this chapter we investigate the following question: given a shadow S, how many diagrams associated to S are unknot diagrams?

6.1 Introduction

6.1.1 The main question

Given a knot diagram D, it is natural to ask if it is always possible to change some over/under assignments so that the result is an unknot diagram. As we will recall shortly, this is indeed always possible. The minimum number of such operations is the *unknotting number* of D, a widely investigated knot theory parameter.

A closely related question is to take as input a shadow S, and find an over/under assignment that creates an unknot diagram. The fact that such an over/under assignment always exists (which implies that the unknotting number of a diagram is always well-defined) is a standar result in any introductory knot theory course, but it would be an overstatement to claim that it is trivial. The existence of such an over/under assignment is usually demonstrated with the following intuitive argument, which we reproduce from [31] (we do not reproduce Kauffman's figure; Figure 6.1 is our own example):

> *[Referring to a diagram such as the one on the right hand side of Figure 6.1, which is an unknot diagram]. Diagrams of this type are produced by tracing a curve and always producing an undercrossing at each return crossing. This type of knot is called a standard unknot. Of course we see clearly that a standard unknot is unknotted by just pulling on it, since it has the same structure as a coil of rope that is wound down onto a flat surface.*

This lively argument gives the following algorithm to create an unknot diagram, starting from any shadow S. Start at any point p of S that is not a vertex. We now traverse a straight-ahead Eulerian walk of S, starting (and ending) at p and each time we find a vertex for the first time, assign an overcrossing to the strand that is being traversed. In the example in Figure 6.1, the input is the shadow on the left hand side. By following the algorithm just described, starting from the marked point p in the direction shown, we obtain the unknot diagram on the right hand side. We note that the fact that this is an unknot diagram is a particular instance of the Remark at the end of the previous section.

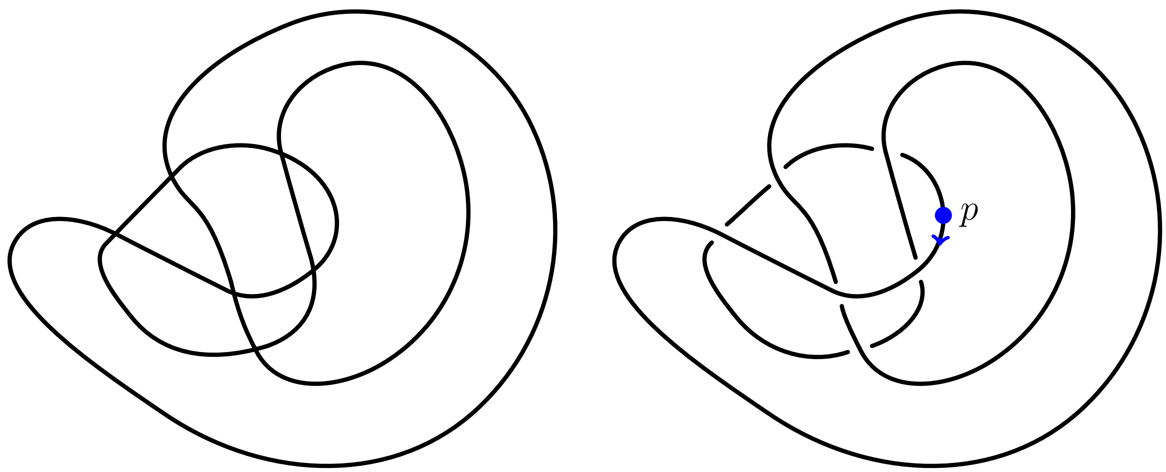


Figure 6.1: Creating an unknot diagram starting from any given shadow.

Thus for any given shadow there always exists a prescription for each of its vertices so that the resulting assignment gives an unknot diagram. If S is a shadow on n vertices, then there are 2^n diagrams associated to S. It seems natural to ask how many of these 2^n diagrams are unknot diagrams:

Main question. *Let S be a shadow with n vertices. Out of the 2^n possible assignments for S (that is, out of the 2^n diagrams associated to S), how many are unknot diagrams?*

The first rather obvious observation is that the answer to this question surely depends on the shadow under consideration. If a shadow S has at least one cut-vertex then every assignment of S will be a a diagram with at least one crossing that is reducible, meaning that it is removable via a simple twist. Thus the prescrition of a cut-vertex in S is somehow a degree of fredom for the whole assignment. A shadow without cut-vertices is associated to diagrams in which no crossing is reducible, these diagrams are called *reduced diagrams*.

For instance, consider a shadow of a reduced diagram of the trefoil knot (Figure 5.1 (b)). This shadow has 3 vertices, and so there are $2^3 = 8$ diagrams associated to this shadow. These 8 diagrams are depicted in Figure 6.2. It is an easy exercise to show that no crossing is removable via a simple twist and that 6 of these 8 diagrams correspond to the unknot, 1 corresponds to the trefoil knot, and 1 corresponds to the mirror of the trefoil knot.

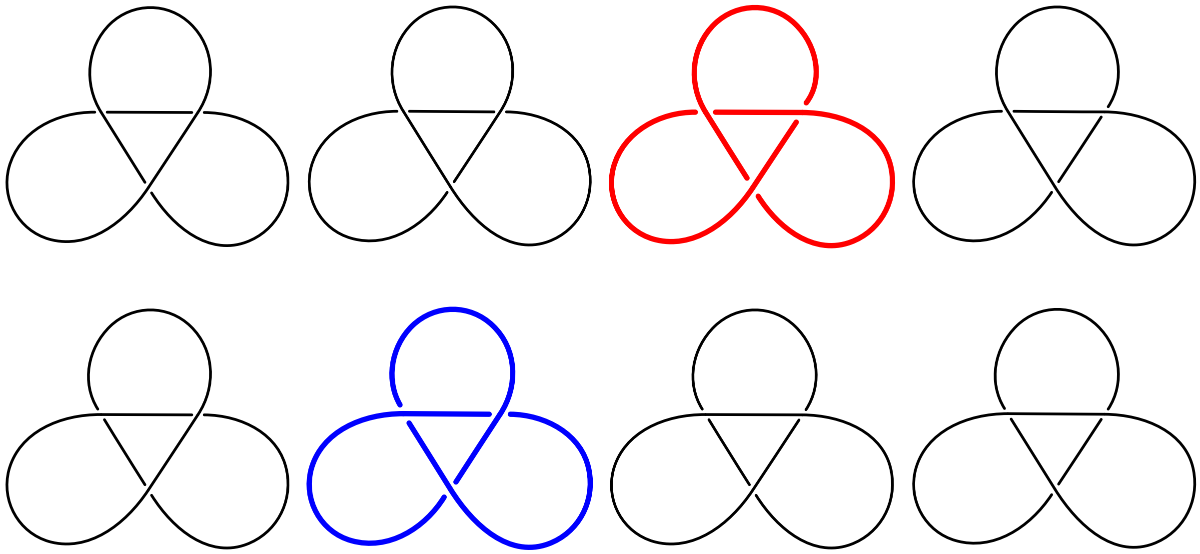


Figure 6.2: If we take a shadow of the trefoil knot with three crossings, then out the $2^3 = 8$ diagrams associated to the shadow, 6 are unknot, 1 (top row, third diagram from left to right) is equivalent to the trefoil knot, and 1 is equivalent to the mirror of the trefoil knot (bottom row, second diagram from left to right).

An extreme example in the opposite direction is illustrated in Figure 6.3. This shadow has 4 vertices each one being a cut-vertex, and it is easy to see that each of the $2^4 = 16$ diagrams associated to this shadow is an unknot diagram.

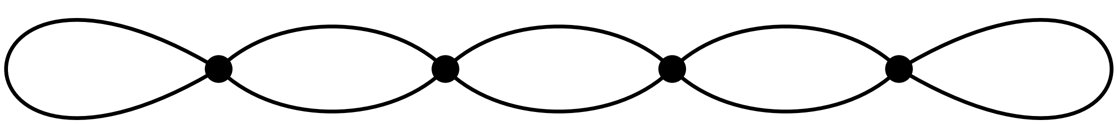


Figure 6.3: Every diagram associated to this shadow is an unknot diagram.

These examples illustrate the observation that the answer to the Main Question depends on the shadow under consideration. The interesting question is whether or not, regardless of the shadow, one can give interesting, nontrivial general bounds for the proportion of diagrams (associated to a fixed shadow) that represent the unknot.

Suppose S is a shadow on n vertices. Using the framework of Gauss codes as in [30], it is easy to see that the procedure described above can be used to produce exactly $2n$ distinct unknot diagrams associated to S. Thus we have the following.

Remark. *If S is a shadow on n vertices, then there exist at least $2n$ unknot diagrams associated to S.*

Our work in this chapter was motivated by asking: how much better can one do? Is it true that there exists a superlinear function $f(n)$ such that every shadow with n vertices has at least $f(n)$ unknot diagrams associated to S? How about a superpolynomial $f(n)$? Or an exponential $f(n)$?

Before moving on to a summary of our results around these questions, let us briefly review some related work in the literature.

6.1.2 Related work

Recently, Cantarella et al. [10] carried out a complete investigation of all the knot types that arise in all shadows with 10 or fewer crossings. Their results show that, among all the diagrams associated to this huge collection of shadows, roughly 78% are unknot diagrams. Cantarella et al. explain that this large proportion is explained by the existence of "tree-like" shadows, such as the one we depict in Figure 6.3. These are shadows in which every vertex is a cut vertex, and so every associated diagram is unknot.

There is also a number of quite interesting results in the literature related to the complement of our Main Question from the previous subsection: out of the 2^n diagrams associated to S, how many are *not* unknot diagrams?

According to Sumners and Whittington [45], this was first asked in the context of long linear polymer chains, independently by Frisch and Wasserman [24] and Delbruck [15]. In our context, the Frisch-Wasserman-Delbruck Conjecture may be roughly paraphrased as follows. Let $p(n)$ be the probability that a random knot diagram with n crossings is an unknot diagram. Then $p(n)$ goes to zero as n goes to infinity.

The first issue to settle in an investigation of the Frisch-Wasserman-Delbruck Conjecture is: what is a random knot diagram? In [45], Sumners and Whittington investigated (and settled in the affirmative) this conjecture in the model of self-avoiding walks on the three-dimensional simple cubic lattice (see [43] for a closely related result). The conjecture has also been settled in other models of space curves, such as self-avoiding Gaussian polygons and self-avoiding equilateral polygons [7, 8, 17, 18, 27].

The problem of proposing suitable models of random knot diagrams is of interest by itself. The associated difficulties are explained and discussed in [16], where two different such models are presented and investigated. We also refer the reader to the preliminary report by Dunfield at al. in [19]. It is also worth mentioning the very recent work of Even-Zohar et al. [23], where several rigorous results are established on the distributions of knot and link invariants for the Petaluma model, which is based on the representation of knots and links as petal diagrams [4]. In a very interesting recent development, Chapman [11, 12] recently proved the Frisch-Wasserman-Delbruck Conjecture under a very general and natural model.

From the work of Chapman we know that if we take a random diagram on n vertices, then the probability that it is an unknot diagram decreases exponentially with n. With this knowledge in hand, our work in this chapter, revolving around the Main Question posed in the previous subsection, focuses on the abundance (or non-abundance) of unknot diagrams associated to an arbitrary (that is, not to a random) shadow.

6.1.3 Our results

Returning to the discussion at the end of Section 6.1.1, our goal is to investigate how many unknot diagrams are guaranteed to exist for an arbitrary shadow.

As we mentioned in Subsection 6.1.1, a folklore argument that is a standar result in all elementary knot theory courses can be used to show that the number of unknot diagrams associated to any given shadow is at least linear in the number of vertices in the shadow. Our main result in this chapter is the following statement, which claims that the number of unknot diagrams associated to S is actually superpolynomial.

Theorem 9 (Main Theorem)**.** *Let S be a shadow with n vertices. Then there exist at least $2^{\sqrt[3]{n}}$ unknot diagrams associated to S.*

In Section 6.2 we give an outline of the strategy behind the proof of Theorem 9. As we explain in Section 6.2, the main statements behind the proof are worked out in Sections 6.3, 6.4, and 6.5. From these results the proof of Theorem 9, given at the end of Section 6.2, is a mere formality.

It is natural to ask about the tightness of the bound given in Theorem 9. This bound is superpolynomial. Would it be possible to go to the next level, and prove an exponential lower bound? At this point we are unable to answer this question, but we have several remarks in this direction. This discussion is given in Section 7.1. We close Part II with some concluding remarks and open questions in Chapter 8.

6.2 Proof of Theorem 9

In what follows, if S is a shadow, then we let $\mathcal{U}(S)$ denote the set of unknot diagrams of S. Thus our ultimate goal is to show that if S has n vertices, then $|\mathcal{U}(S)| \geq 2^{\sqrt[3]{n}}$.

To prove Theorem 9 we need to show that any given shadow S has many unknot diagrams associated to it. The proof is of an inductive nature. Given a shadow S, the idea is to find a shadow S' smaller (with less vertices) than S, and then show that $|\mathcal{U}(S)|$ is greater than (typically twice) $|\mathcal{U}(S')|$. Applying a similar procedure to S', by an inductive reasoning we find the claimed lower bound for $\mathcal{U}(S)$.

The first technique we use to find from S a smaller shadow is the following. Let S be a shadow, and suppose that C is a straight-ahead cycle of S. Let S' be the shadow obtained from S by removing the edges of C, and suppressing any resulting degree 2 vertices. We use the notation $S' = S /\!\!/ C$ to denote that S' is obtained from S in this way.

The key claim is that if $S' = S /\!\!/ C$, then $|\mathcal{U}(S)| \geq 2|\mathcal{U}(S')|$, and that a stronger inequality holds if C has more than one vertex. Formally we have the following statement, whose proof is in Section 6.3.

Proposition 10. *Let S be a shadow, and let C be a straight-ahead cycle of S. Let $S' = S /\!\!/ C$. Then $|\mathcal{U}(S)| \geq 2|\mathcal{U}(S')|$. Moreover, if C has more than one vertex, then $|\mathcal{U}(S)| \geq 4|\mathcal{U}(S')|$.*

Continuing with the discussion, the idea to iteratively apply this idea until we end up with a trivial shadow. The formal setting makes use of the following definition.

Definition. (Cycle decomposition). Let S be a shadow. Suppose that $S = S_1, S_2, \ldots, S_p$ is a sequence of subshadows of S, with the following properties:

- For $i = 1, 2, \ldots, p-1$, there is a straight-ahead cycle C_i of S_i such that $S_{i+1} = S_i /\!\!/ C_i$.
- S_p is the trivial shadow.

Set $C_p := S_p$. Then the sequence $C_1, C_2, \ldots, C_p$ is a *cycle decomposition* of S *of size* p.

We refer the reader to Figure 6.4 for an example of a cycle decomposition.

An iterative application of Proposition 10 yields the following statement, which is the first main ingredient in the proof of Theorem 9.

Lemma 11. *Let S be a shadow with n vertices. Suppose that S has a cycle decomposition of size at least $\sqrt[3]{n}$. Then $|\mathcal{U}(S)| \geq 2^{\sqrt[3]{n}}$.*

This lemma, whose proof is also in Section 6.3 settles Theorem 9 as long as S has a cycle decomposition of the given size. If all cycle decompositions of S are smaller, then we need a different approach. The second tool we develop for dealing with this case is more involved, but the core ideas are easy to explain.

To illustrate this second technique, consider the shadow S in Figure 6.5. We highlight two straight-ahead cycles B (solid) and R (thick). Let us say that the edges of B are *blue*, the edges of R are *red*, and the edges that are neither blue nor red are *gray* (see Figure 6.5(a)).

In order to produce many unknot diagrams associated to this kind of shadow, we proceed as follows. First of all, we focus on diagrams of S for which every blue-gray crossing is an overpass for the blue strand, and every red-gray crossing is an overpass for the red strand. Throughout this informal discussion let us refer to these as *good diagrams*. Thus loosely

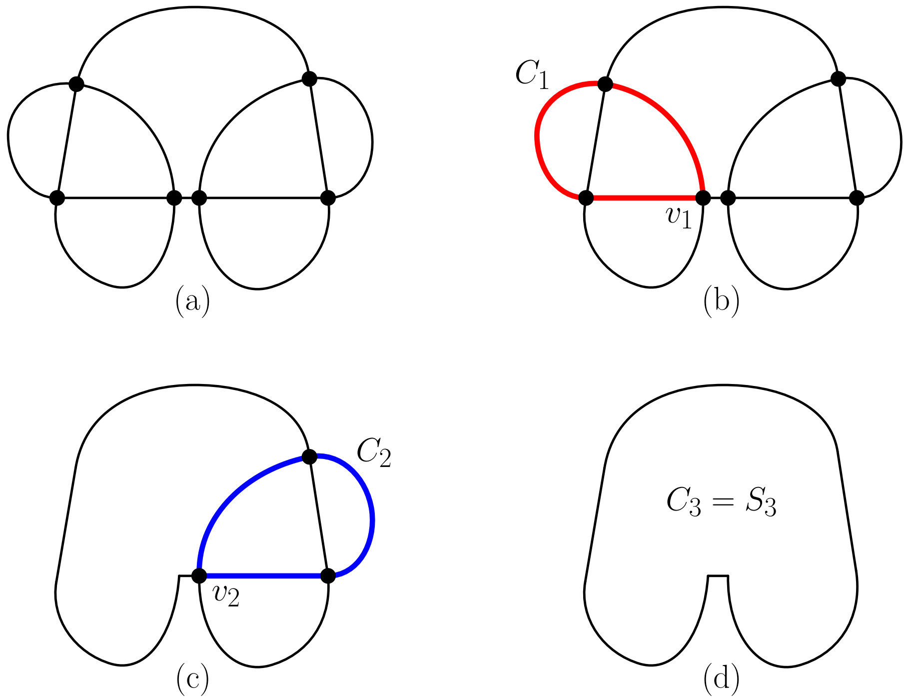


Figure 6.4: A cycle decomposition of a shadow. In (a) we depict a shadow $S = S_1$. In (b) we highlight a straight-ahead cycle C_1. After removing the edges of C_1, and suppressing the three resulting degree 2 vertices, we reach the shadow $S_2 = S_1 /\!\!/ C_1$ in (c). Also in (c) we highlight a straight-ahead cycle C_2 of S_2. After removing the edges of C_2, and suppressing the three resulting degree 2 vertices, we obtain the trivial shadow $S_3 = S_2 /\!\!/ C_2$, as shown in (d). We finally set $C_3 = S_3$ (thus C_3 is a vertex-free cycle). The sequence C_1, C_2, C_3 is then a cycle decomposition of $S = S_1$.

speaking a good diagram of S is one in which all the gray strands lie "below" the blue and the red strands. We refer the reader to Figure 6.5(b).

In one case, $S = B \cup R$, so there are no gray edges. Let us focus on this discussion in the more difficult case in which there do exist gray edges. In this case, if we ignore completely the gray edges, $B \cup R$ may be regarded as a shadow of a link with two components. It is not too difficult to show that if B and R have m vertices in common, then m is even, and $B \cup R$ has at least $2^{m/2}$ unknot diagrams associated to it.

Going back to S, which consists of $B \cup R$ plus some gray edges, the central idea is that an unlink diagram for $B \cup R$ can be extended to an unknot diagram of S. Roughly speaking, since the gray strands lie below the blue and red strands, we may perform Reidemeister moves on $B \cup R$ safely, essentially ignoring the gray part, since these moves take place "above" the gray part. An unlink diagram of $B \cup R$ can thus be extended to an unknot diagram of S by performing the moves on $B \cup R$ until we unlink them, then collapsing each of B and R to a single point, and then unknotting the gray part (for this last part, it suffices to take any unknot diagram of the gray part).

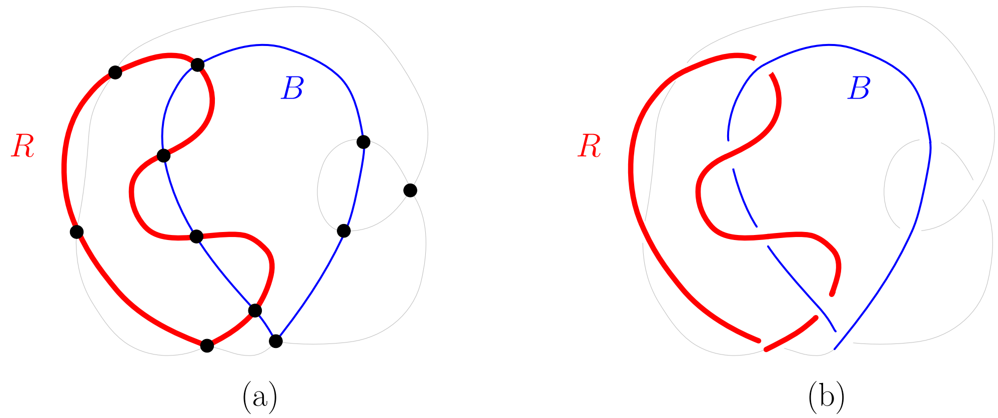


Figure 6.5: In (a) we depict a shadow S that has two straight-ahead cycles, the red cycle R (thick edges) and the blue cycle B. The edges that are neither red nor blue are gray. In (b) we show a diagram of S, in which every blue-gray (respectively, red-gray) crossing is an overpass for the blue strand (respectively, red strand).

As one would expect, there are some technical difficulties that need to be sorted out in the formal proof of this argument, but the simple core idea is the one we have just explained.

Our second main ingredient is summarized in the following statement.

Lemma 12. *Let S be a shadow, and let B and R be distinct straight-ahead cycles of S. Suppose that B and R have exactly m vertices in common. Then $|\mathcal{U}(S)| \geq 2^{m/2}$.*

This statement, whose proof is given in Section 6.4, settles Theorem 9 if S has two straight-ahead cycles that have at least $2\sqrt[3]{n}$ vertices in common. Not surprisingly, there exist shadows that satisfy neither the condition in Lemma 11 nor the condition in Lemma 12. On the other hand, one can show that if S does not satisfy the condition in Lemma 11, then there is a substructure of S that does satisfy the condition in Lemma 12. To make this precise, let us introduce the concept of a subshadow. This notion is defined by the following conditions:

1. Every shadow is a subshadow of itself.
2. If S is a shadow and C is a straight-ahead cycle of S, then $S /\!/ C$ is a subshadow of S.
3. (Transitivity). If S' is a subshadow of S, and S'' is a subshadow of S', then S'' is a subshadow of S.

Equivalently, we can say that T is a subshadow of S if there is a sequence $S = S_1, S_2, \ldots, S_r = T$ such that for $i = 1, 2, \ldots, r-1$, there is a straight-ahead cycle C_i of S_i such that $S_{i+1} = S_i /\!/ C_i$. Note that Condition 1 above admits the possibility that $r = 1$, that is, $S = T$.

An easy but crucial fact that follows from Proposition 10 is that if T is a subshadow of S, then S has at least as many unknot diagrams as T. Thus in order to apply Lemma 12 to obtain Theorem 9, we do not need the existence of two straight-ahead cycles with these properties in S itself; it suffices to guarantee the existence of such straight-ahead cycles in some subshadow T of S. This last ingredient is provided by the following statement, whose proof is in Section 6.5.

Lemma 13. *Let S be a shadow with n vertices. Suppose that every cycle decomposition of S has size at most $\sqrt[3]{n}$. Then there is a subshadow T of S with the following property. There exist straight-ahead cycles B and R of T, such that B and R have at least $2\sqrt[3]{n}$ vertices in common.*

With this last ingredient in hand, we can finally give the proof of Theorem 9.

Proof of Theorem 9. We start by noting that it immediately follows, from the definition of a subshadow and Proposition 10, that if T is a subshadow of S, then $|\mathcal{U}(S)| \geq |\mathcal{U}(T)|$.

Let S be a shadow with n vertices. If S has a cycle decomposition of size at least $\sqrt[3]{n}$, then we are done by Lemma 11. If every cycle decomposition of S has size at most $\sqrt[3]{n}$, then by Lemma 13 it follows that S has a subshadow T, such that T has straight-ahead cycles B and R such that B and R have at least $2\sqrt[3]{n}$ vertices in common. Applying Lemma 12 to T, we obtain that $|\mathcal{U}(T)| \geq 2^{\sqrt[3]{n}}$. Finally, from the observation in the previous paragraph we have that $|\mathcal{U}(S)| \geq |\mathcal{U}(T)|$, and so $|\mathcal{U}(S)| \geq 2^{\sqrt[3]{n}}$. □

6.3 Finding unknot diagrams using cycle decompositions: proof of Lemma 11

Our aim in this section is to prove Proposition 10 and Lemma 11. The latter follows from the former by an easy induction argument.

Proof of Proposition 10. Let S, S' be as in the statement of the proposition. Note that for every vertex v' in S' there is a vertex v in S that naturally corresponds to v'. Now let D and D' be diagrams of S and S', respectively. If for every vertex v in S its prescription in D coincides with the prescription of its corresponding vertex v' in D', we say that D is an *extension* of D'.

We illustrate this idea in Figure 6.6. In part (a) of this figure we have a shadow S, and a straight-ahead cycle C of S. In (b) we have the shadow $S' = S /\!/ C$. On the right-hand side we have an unknot diagram D' of S', and on the left-hand side we have four diagrams of S that are extensions of D'.

Let c be the root of C. Consider a fixed unknot diagram D' of S'. We let $\mathcal{D}_{D'}$ denote the set of those diagrams D of S such that (i) D is an extension of D'; and (ii) the strand σ of D that corresponds to C is either an overstrand or an understrand.

If c is not the only vertex of C , then $\mathcal{D}_{D'}$ consists of exactly four diagrams. Two of these diagrams are the extensions of D' in which σ is an overstrand (there are two such diagrams,

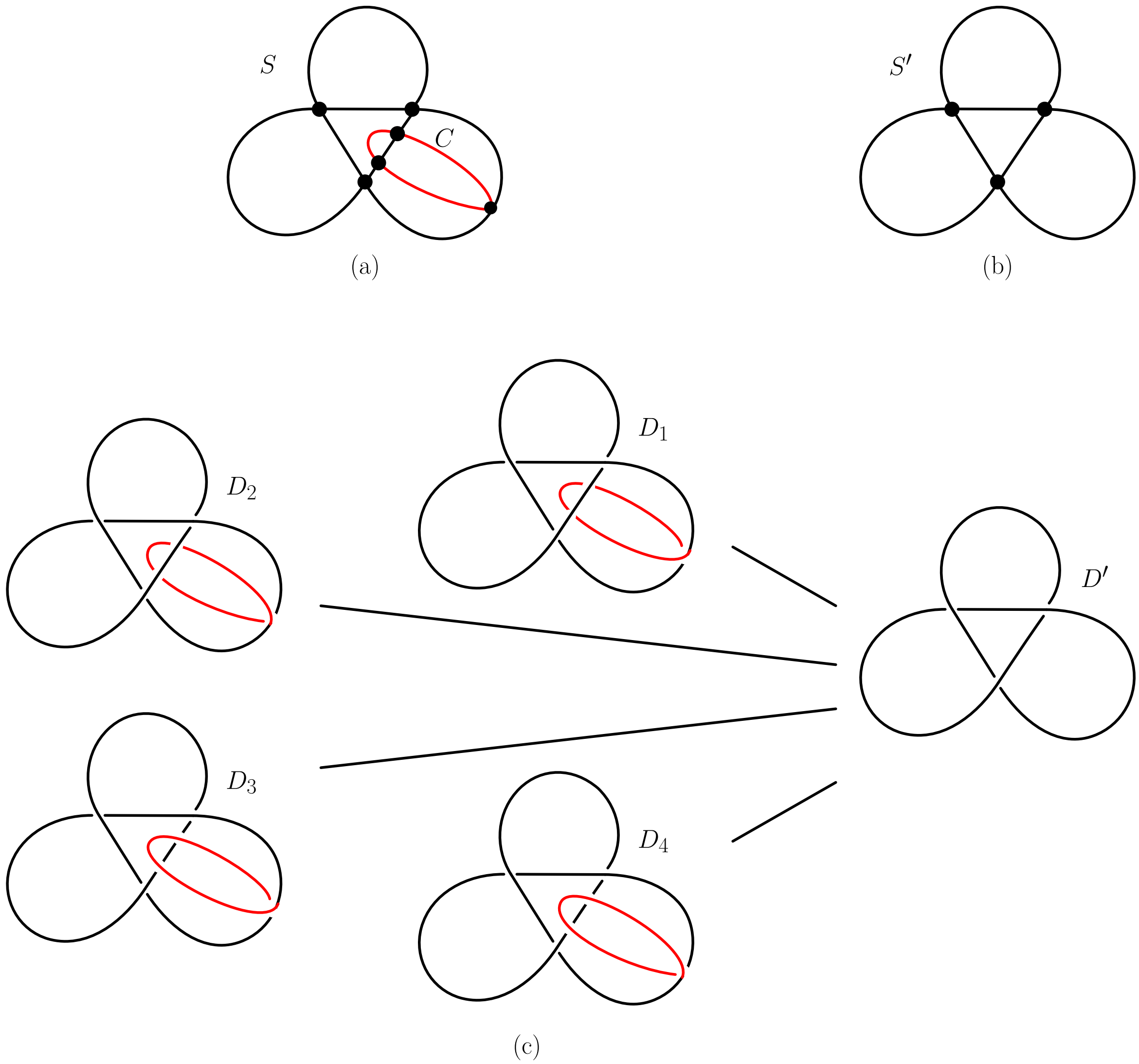


Figure 6.6: Illustration of the proof of Proposition 10. In (a) we depict a shadow S, and a straight-ahead cycle C of S. In (b) we illustrate the shadow $S' = S /\!/ C$. On the right hand side of (c) we depict an unknot diagram D' of S'. On the left hand side of (c) we illustrate the four unknot diagrams of S that we obtain as extensions of the unknot diagram D' of S'.

since the only freedom left is the prescription at c), and the other two diagrams are the extensions of D' in which σ is an understrand. We refer the reader again to Figure 6.6. For this example, $\mathcal{D}_{D'}$ consists precisely of the four diagrams of S shown on the left hand side of (c). The two diagrams on top are the extensions of D' in which the strand σ that corresponds to C is an understrand, and the two diagrams at the bottom are the extensions of D' in which σ is an overstrand.

On the other hand, if c is the only vertex of C, then $\mathcal{D}_{D'}$ consists of exactly two diagrams: condition (ii) is vacuous in this case, and so the two possible extensions of D' are given by the two possible prescriptions at c.

Now consider any $D \in \mathcal{D}_{D'}$. Since σ is an overstrand or an understrand in D, it follows that we can perform a generalized Reidemeister move on σ, replacing it with the strand that consists solely of the point c, thus obtaining diagram D'. Therefore D is equivalent to D', and since D' is an unknot diagram, it follows that D is also an unknot diagram.

We now note that if D', D'' are distinct unknot diagrams of S', for they have a different prescription of the vertices in S', then $\mathcal{D}_{D'}$ and $\mathcal{D}_{D''}$ are disjoint. From this fact and the observations in the previous two paragraphs it follows that if c is the only vertex of C, then $|\mathcal{U}(S)| \geq 2|\mathcal{U}(S')|$, and if c is not the only vertex of C, then $|\mathcal{U}(S)| \geq 4|\mathcal{U}(S')|$. □

Proof of Lemma 11. Let S be a shadow with n vertices, such that S has a cycle decomposition $S_1, S_2, \ldots, S_p$ with $p \geq \sqrt[3]{n}$. A repeated application of Proposition 10 shows that $|\mathcal{U}(S)| \geq 2^{p-1}|\mathcal{U}(S_p)|$. Since S_p is the trivial shadow, which has only one unknot diagram, we obtain that $|\mathcal{U}(S)| \geq 2^{p-1}$.

The bound $|\mathcal{U}(S)| \geq 2^{p-1}$ just obtained is the best we can obtain from Proposition 10 only if at each step in the cycle decomposition, we remove a cycle that has only one vertex. Otherwise we can apply at least once the second part of Proposition 10, and obtain that $|\mathcal{U}(S)| \geq 2^p$. In such a case we are then done, since $p \geq \sqrt[3]{n}$.

Thus we are done unless, at each step in the cycle decomposition, we remove a cycle that has only one vertex. It is straightforward to see that this can happen only if $p = n + 1$. In this case the bound $|\mathcal{U}(S)| \geq 2^{p-1}$ reads $|\mathcal{U}(S)| \geq 2^n$. Since $2^n > 2^{\sqrt[3]{n}}$ for every positive integer n, it follows that in this case we are also done. □

6.4 Finding unknot diagrams using a pair of straight-ahead cycles: proof of Lemma 12

The arguments for the proof of Lemma 12 depend on whether or not $S = B \cup R$. As we will see below, in each case the proposition follows from a sequence of claims. The statements of these claims could be easily merged to take into account both cases, and their proofs could be correspondingly adapted to have a single sequence of claims covering both cases. We feel that this would artificially shorten the proof of Lemma 12, and so we have decided against this possibility. We think it is more natural, and much easier to follow, if these cases are handle separately when discussing and proving these claims.

An additional advantage we see from dealing with these cases separately is that we can work out the easier case $S = B \cup R$ first, paving the way for the other, slightly more difficult case.

In both cases, a major ingredient in the proof is the analysis of digons in a system of two simple closed curves in the plane. With the purpose of not interrupting the discussion at a later point with this analysis, we start this section by establishing the respective results.

6.4.1 Digons in systems of closed curves

Let β be a simple closed curve in the plane. A *segment* of β is a subset of β that is homeomorphic to the closed interval $[0,1]$. Thus a segment of β is simply a non-closed curve (including its endpoints) contained in β.

Let β, ρ be simple closed curves that pairwise intersect a finite number of times. A *digon of* (β, ρ) is a pair of segments (α, γ), where α is in β and γ is in ρ, such that (i) α and γ have common endpoints x, y; (ii) the only points in ρ that are in α are x and y; and (iii) the only points in β that are in γ are x and y. In Figure 6.7 we illustrate this concept.

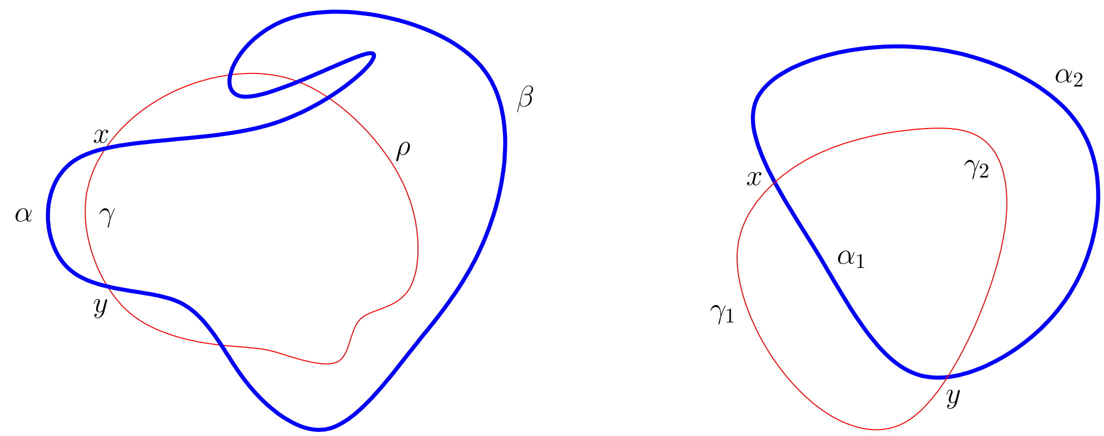


Figure 6.7: On the left hand side we have two simple closed curves β (thick) and ρ (thin). The common endpoints of the blue segment α and the red segment γ are x and y. The only points in ρ (respectively, β) that are in α (respectively, γ) are x and y, and so (α, γ) is a digon of (β, ρ). On the right hand side we have two simple closed curves $\beta = \alpha_1 \cup \alpha_2$ (thick) and $\rho = \gamma_1 \cup \gamma_2$ (thin). This example illustrates that a digon does not necessarily bound an empty disk (that is, a disk whose interior does not contain any point in the curves). Indeed, the pair (α_2, γ_1) is a digon of $(\alpha_1 \cup \alpha_2, \gamma_1 \cup \gamma_2)$, and $\alpha_2 \cup \gamma_1$ does not bound an empty disk.

If (α, γ) is a digon of (β, ρ), then $\alpha \cup \gamma$ is a simple closed curve. Properties (ii) and (iii) above imply that at least one of the two connected components of $\mathbb{R}^2 \setminus (\alpha \cup \gamma)$ does not have any point in $\beta \cup \rho$ (that is, at least one of these connected components is an *empty region*). We remark that it is not necessarily true that this empty region is the disk bounded by $\alpha \cup \gamma$ (see right hand side of Figure 6.7).

We now state and prove the results on digons that we will need in the proof of Lemma 12.

Claim A. *Let β, ρ be simple closed curves in the plane that intersect each other a finite, positive number of times. Suppose that every intersection point of β and ρ is a crossing (rather than tangential). Then there are at least 4 distinct digons of (β, ρ). Moreover, if b is a point of β that is not in ρ, and r is a point of ρ that is not in β, then there exists a digon (α, γ) of (β, ρ) such that $\alpha \cup \gamma$ contains neither b nor r.*

Proof. To help comprehension, we colour β blue and ρ red. Let Δ denote the closed disk bounded by β.

We can naturally regard the part of $\beta \cup \rho$ contained in Δ as a 3-regular plane graph: each vertex is incident with two blue edges and one red edge. It is easily seen that if we take the dual of this plane graph, and discard the vertex corresponding to the unbounded face, the result is a tree. Now each leaf of this tree corresponds to a digon of (β, ρ), and it is an elementary graph theoretical fact that every tree has at least two leaves. Thus there are at least two digons of (β, ρ) whose red segments are inside Δ. A totally analogous argument shows that there are at least two digons of (β, ρ) whose red segments are outside Δ. This shows the existence of at least four distinct digons of (β, ρ).

The existence of a digon (α, γ) of (β, ρ) such that $\alpha \cup \gamma$ contains neither b nor r is easily worked out by a simple case analysis if β and ρ have exactly two intersection points. We note that, in this case, there is exactly one digon of (β, ρ) with this property.

Suppose now that β and ρ have at least 4 points in common. Then if (α, γ) and (α', γ') are distinct digons, we must have that $\alpha \neq \alpha'$ and $\gamma \neq \gamma'$. It follows that only one digon of (β, ρ) can contain b, and only one digon of (β, ρ) can contain r. Since there are at least four digons of (β, ρ), it follows that there are at least two (and in particular at least one) digons (α, γ) of (β, ρ) such that $\alpha \cup \gamma$ contains neither b nor r. □

Claim B. *Let β, ρ be simple closed curves in the plane that intersect each other a finite number ℓ of times, where $\ell \geq 3$. Suppose that there is exactly one intersection point z of β and ρ that is tangential; all the other intersection points are crossings. Then there exists a digon (α, γ) of (β, ρ) such that $\alpha \cup \gamma$ does not contain z.*

Proof. As in the proof of Claim A, we colour β blue and ρ red, and let Δ denote the closed disk bounded by β.

Suppose that if we take a small disk δ centered at z (small enough so that z is the only point in $\beta \cap \rho$ that is in δ), then $\rho \cap \delta$ lies outside Δ (see Figure 6.8). The other possibility ($\rho \cap \delta$ lies inside Δ) is handled in a totally analogous manner.

Now we proceed as in Claim A, ignoring the part of $\beta \cup \rho$ outside Δ, and find at least two distinct digons $(\alpha, \gamma), (\alpha', \gamma')$ of (β, ρ), such that γ and γ'(not necessarily distinct) are inside Δ. Only one of α and α' can contain z. If α contains z, then (α', γ') is a digon of (β, ρ) that does not contain z. If α' contains z, then (α, γ) is a digon of (β, ρ) that does not contain z. □

6.4.2 Proof of Lemma 12 for the case $S = B \cup R$

In this case, the root of B is the same as the root of R. We denote this common root by r. To help comprehension, we say that the edges of B are *blue*, and the edges of R are *red*.

The case where S has exactly one vertex is handled in Claim C below. Thus, suppose for now that S has more than one vertex, it follows from Claim B, above, that there exist e

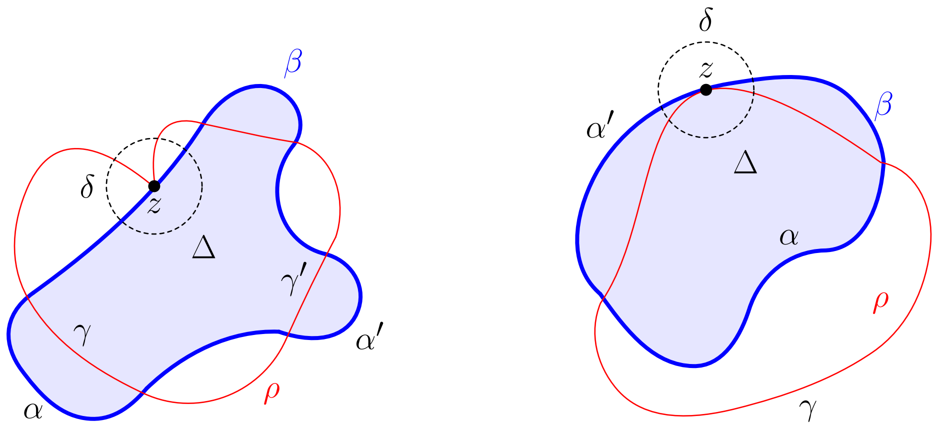


Figure 6.8: Illustration of the proof of Claim B. In each side of this figure we depict a pair of simple closed curves β and ρ with a finite number of intersections, one of which (namely z) is a tangential intersection point. On the left hand side we have the case in which $\delta \cap \rho$ lies outside the disk Δ bounded by β. Both γ and γ' are segments inside Δ, and both (α, γ) and (α', γ') are digons, neither of which contains z. In the example on the right hand side, $\delta \cap \rho$ lies inside Δ, and γ is the only red segment that lies outside Δ. By letting α and α' be the blue segments into which ρ naturally breaks β, we have that (α, γ) and (α', γ) are both digons. Of these two digons, only (α, γ) does not contain z.

and f edges of S, where e is blue and f is red, such that e and f have common endpoints diferent from r and neither of them contains r as an interior vertex. Then we call (e, f) a *valid pair* of S.

Let (e, f) be a valid pair of S, and let u, v be the common endpoints of e and f. Since r is not an endpoint of e and f, it follows that the rotation around each endpoint of e (and f) is blue-red-blue-red. Loosely speaking, we might say that the cycles B and R cross at both u and v.

We obtain from S a new shadow S' as follows. First we split u into two degree 2 vertices. There are two ways to do such a splitting, so that the result is still a plane graph. We choose to split u so that, in the resulting graph, one of the two resulting degree 2 vertices is incident with e, and the other is incident with f. We refer the reader to Figure 6.9. We proceed analogously with v: we split v so that one of its two resulting degree 2 vertices is incident with the remains of e, and the other one is incident with the remains of f. Finally, we suppress the four degree 2 vertices obtained in the process.

It is readily seen that S' inherits from S the property that it is the union of two straight-ahead cycles. These straight-ahead cycles can be naturally labelled B' and R', with the convention that B' is the straight-ahead cycle that contains the remains of the red edge f, and R' is the straight-ahead cycle that contains the remains of the edge e. We refer the

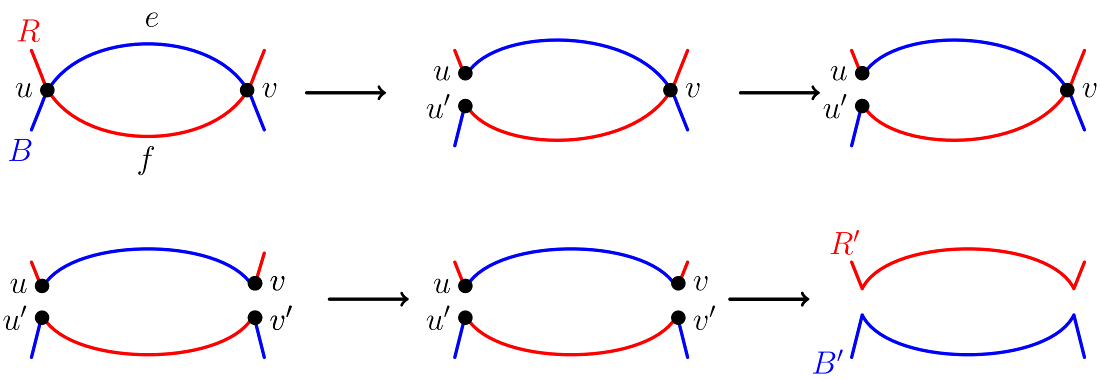


Figure 6.9: If (e, f) is a valid pair of edges whose common endpoints are u and v, then we obtain from S a shadow S' by splitting u and v as shown. The resulting shadow S' also has the property that it is the union of two straight-ahead cycles. We can naturally label these cycles B' and R', with the convention that B' is the cycle that contains the remains of the red edge f, and R' is the cycle that contains the remains of the blue edge e.

reader again to Figure 6.9, where (parts of) the cycles B' and R' are shown. With this convention, all the edges of B that were not affected by this operation are edges of B' (so we can naturally colour the edges of B' blue), and all the edges of R that were not affected by this operation are edges of R' (so we can naturally colour the edges of R' red). Moreover, it is immediately seen that r is the common root of B' and R'. We let $(S, B, R) \to_{e,f} (S', B', R')$ denote this oparation. Whenever it is not needed to specify that we obtain S', B', and R' from the particular pair (e, f), we simply write $(S, B, R) \to (S', B', R')$.

We also refer the reader to Figure 6.10. In part (a) of this figure we have a shadow S that is the union of two straight-ahead cycles B and R, and identify a valid pair (e, f). in part (b) we show the shadow S', and the cycles B' and R', such that $(S, B, R) \to_{e,f} (S', B', R')$.

Recall that our final aim (Lemma 12) is to show that S has at least $2^{m/2}$ unknot diagrams, where m is the number of vertices that B and R have in common. In the present case $S = B \cup R$, and so each vertex of S is in both B and R. Thus in this case m is simply the number n of vertices of S. We first deal with the case $m = 1$. Although this case is a triviality, we prefer to state it formally to maintain a paralelism with the case $S \neq B \cup R$.

Claim C. *Suppose that S consists of a single vertex. Then $|\mathcal{U}(S)| \geq 2$.*

Proof. In this case S consists of two loop-edges, both incident with the only vertex of S. Trivially, each of the two possible prescriptions at this vertex yields an unknot diagram. □

Claim D. *Suppose that $(S, B, R) \to_{e,f} (S', B', R')$. Then $|\mathcal{U}(S)| \geq 2|\mathcal{U}(S')|$.*

Proof. Let u, v be the common endpoints of e and f. Recall that u, v are the only two vertices of S that get removed in the process of getting S'. Thus every vertex of S' corresponds naturally to a vertex in S.

Let D, D' be diagrams of S and S', respectively. Borrowing the terminology from the proof of Lemma 11 (there should be no confusion, since these are totally separate statements), we say that D is an *extension* of D' if for each vertex in S', its prescription in D' coincides with the prescription of its corresponding vertex in D. For each diagram D' of S', there are exactly 4 diagrams of S that are extensions of D', as there are two ways to give prescriptions to each of u and v.

To prove Claim D we show that if D' is an unknot diagram of S', then out of the four diagrams of S that are extensions of D', at least two are unknot diagrams. We refer the reader to Figure 6.10 for an illustration of the argument. In (a) we have a shadow $S = B \cup R$, where B and R are straight-ahead cycles of S. We identify a valid pair (e, f), and in (b) we show the shadow S', and the straight-ahead cycles B' and R' of S', such that $(S, B, R) \to_{e,f} (S', B', R')$. In (e), we show an unknot diagram of S'.

Let D' be an unknot diagram of S'. Consider the diagram D_1 of S such that (i) D_1 is an extension of D'; and (ii) at both crossings u and v, the red strands are overpasses. Because of (ii), we can perform a Reidemeister Type II move on the strands that correspond to e and f, and then perform a series of isotopies so that we end up with the diagram D'. Since D' is an unknot diagram, it follows that D_1 is also an unknot diagram. In Figure 6.10 (c) we show the resulting diagram D_1 obtained in this way from the unknot diagram D' in (e).

Finally, let D_2 be the diagram of S such that (i) D_2 is an extension of D'; and (ii) at both crossings u and v, the blue strands are overpasses. A totally analogous argument shows that D_2 is also an unknot diagram. In Figure 6.10 (d) we show the resulting diagram D_2 obtained in this way from the unknot diagram D' in (e).

We have thus proved that if D' is a diagram in $\mathcal{U}(S')$, then there are two diagrams D of S, such that D is an extension of D', and D belongs to $\mathcal{U}(S)$. We finally observe that if D', F' are distinct diagrams of S', and D, F are diagrams of S such that D is an extension of D' and F is an extension of F', then D and F are distinct. These two facts imply that $|\mathcal{U}(S)| \geq 2|\mathcal{U}(S')|$. □

Claim E. *Let S be a shadow, and let B, R be straight-ahead cycles of S such that $S = B \cup R$. Set $S_1 := S$. Then there is a sequence $S_1, S_2, \ldots, S_p$ of shadows with the following properties. (i) For $i = 1, 2, \ldots, p$, S_i has two straight-ahead cycles B_i, R_i, such that $S_i = B_i \cup R_i$; (ii) $(S_i, B_i, R_i) \to (S_{i+1}, B_{i+1}, R_{i+1})$ for $i = 1, 2, \ldots, p-1$; and (iii) S_p has only one vertex.*

Proof. We prove this claim by induction on the number of vertices of S. In the base case S has exactly one vertex, and the statement is trivially true.

At this point we note that in the present case ($S = B \cup R$, where B and R are straight-ahead cycles of S), S must have an odd number of vertices. This is an immediate consequence of the Jordan curve theorem. Indeed, if we regard B and R as simple closed curves, then the common root of B and R is a tangential intersection of these curves, and all the other intersections between these curves are crossings.

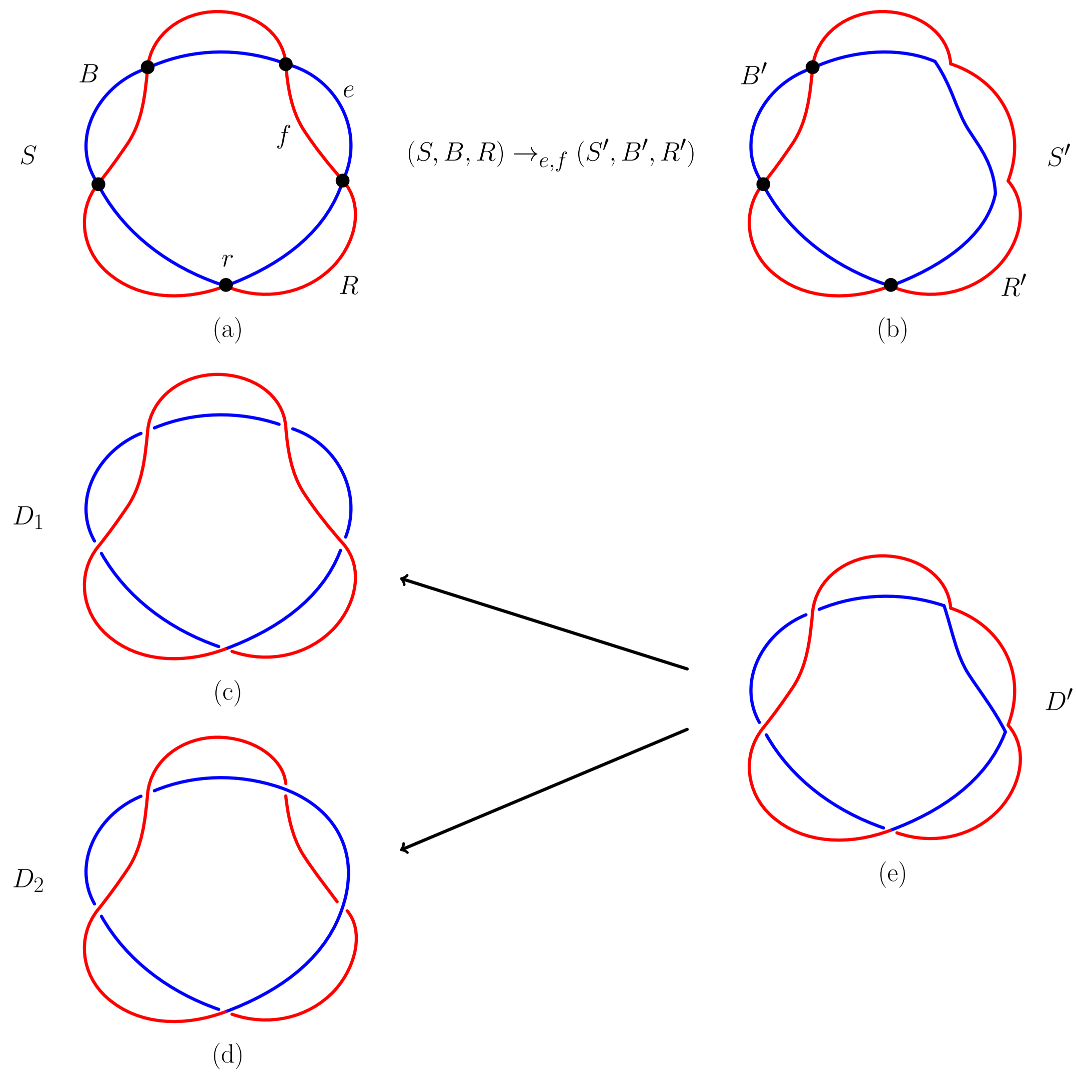


Figure 6.10: In part (a) we have a shadow S with a valid pair (e, f), on which we perform the splitting process. The resulting shadow S' is depicted in (b). In (e) we have an unknot diagram D' of S', and in (c) and (d) we show the unknot diagrams D_1, D_2 of S that are extensions of D'.

For the inductive step, we assume that Claim E is true whenever S has $2k - 1$ vertices, for some integer $k \geq 1$, and consider a shadow S with $2k + 1$ vertices. Recall that $S_1 = S$, and set $B_1 := B$ and $R_1 := R$.

We claim that it suffices to show that there exists a valid pair of edges (e_1, f_1) of (B_1, R_1). Indeed, the existence of this valid pair yields the existence of a shadow S_2, with straight-ahead cycles B_2, R_2, such that $S_2 = B_2 \cup R_2$ and $(S_1, B_1, R_1) \to (S_2, B_2, R_2)$. The inductive hypothesis, applied on S_2, B_2, and R_2, implies the existence of a sequence $S_2, S_3, S_4, \ldots, S_p$ of shadows with the required properties, and so $S_1, S_2, \ldots, S_p$ is the desired sequence. Thus we finish the proof by showing the existence of a valid pair of edges (e_1, f_1) of (B_1, R_1).

Let β_1, ρ_1 be the simple closed curves underlying the straight-ahead cycles B_1 and R_1, respectively. Let r_1 be the common root of B_1 and R_1. By Claim B, (β_1, ρ_1) has a digon (α_1, γ_1) such that r_1 (regarded as a point in $\beta_1 \cup \rho_1$) is not one of the common endpoints of α_1 and γ_1. Let e_1 be the edge of B_1 that corresponds to α_1, and let f_1 be the edge in R_1 that corresponds to γ_1. Then (e_1, f_1) is a valid pair of (B_1, R_1). □

Conclusion of the proof of Lemma 12 for the case $S = B \cup R$. We recall from the proof of Claim E that the number m of vertices of S is odd. Let $k := (m-1)/2$. We will show that $|\mathcal{U}(S)| \geq 2^{k+1}$, which implies Lemma 12, since $k + 1 > m/2$.

Consider the sequence $S = S_1, S_2, \ldots, S_p$ guaranteed by Claim E. We note that S_{i+1} has exactly two fewer vertices than S_i for $i = 1, 2, \ldots, p-1$. Since S_p has one vertex, and S_1 has $m = 2k + 1$ vertices, it follows that $(2k + 1) - 1 = 2(p - 1)$. Thus $p = k + 1$.

By Claim D, we have that $|\mathcal{U}(S_i)| \geq 2|\mathcal{U}(S_{i+1})|$ for $i = 1, 2, \ldots, p-1$. Thus $|\mathcal{U}(S)| = |\mathcal{U}(S_1)| \geq 2^{p-1}|\mathcal{U}(S_p)|$. Since S_p has only one vertex, from Claim A we have that $|\mathcal{U}(S_p)| \geq 2$, and so $|\mathcal{U}(S)| \geq 2^{p-1} \cdot 2 = 2^p = 2^{k+1} > 2^{m/2}$. □

6.4.3 Proof of Lemma 12 for the case $S \neq B \cup R$

As in the case $S = B \cup R$, we say that the edges of B are *blue*, and the edges of R are *red*. We let b denote the root of B, and we let r denote the root of R. The assumption $S \neq (B \cup R)$ implies that, in this case, b and r are distinct vertices. The scenario we face in this case is illustrated in Figure 6.11(a).

Since $S \neq B \cup R$, it follows that S has edges that are neither blue nor red. We colour these edges *gray*, and let Y be the subgraph of S induced by the gray edges.

We say that a vertex v of S is *blue-gray* if the rotation around v consists of a blue edge, followed by a gray edge, followed by a blue edge, and then followed by a gray edge. We extend this naturally to define red-gray, blue-red, and gray-gray vertices. In the example in Figure 6.11(a), vertex w_2 is blue-gray; vertex w_1 is red-gray; and vertices u, v, w_3 and w_4 are blue-red.

We note that b and r are the only vertices of S that are neither blue-gray, nor red-gray, nor blue-red, nor gray-gray. Indeed, b is incident with two blue edges and two gray edges, but since B is a straight-ahead cycle of S, then its rotation is blue-blue-gray-gray, so it is not a blue-gray vertex according to our convention. An analogous argument shows that r is not red-gray. We also note that, since B and R are straight-ahead cycles, then every vertex that is incident with two blue edges and two red edges is necessarily blue-red.

We say that a crossing of a diagram of S is *blue-gray* if its corresponding vertex in S is blue-gray. Thus at each blue-gray crossing, a blue strand crosses (overpasses or underpasses) a gray strand. We define red-gray, blue-red, and gray-gray crossings analogously.

Let D be any diagram of S. Since B and R are straight-ahead cycles of S, it follows that then D is the union of four strands: (i) one blue strand σ_B, whose startpoint and endpoint is b; (ii) one red strand σ_R, whose startpoint and endpoint is r; and (iii) two gray strands, each of which has b and r as its endpoints.

Now let $\mathcal{D}_{B,R}$ be the set of diagrams of S with the following properties:

(P1) Each blue-gray crossing is an overpass for the blue strand σ_B.

(P2) Each red-gray crossing is an overpass for the red strand σ_R.

We remark that (P1) does not fix a prescription for b, since b (even though it is incident with two blue and two gray edges) is not a blue-gray crossing. Analogously, (P2) does not fix a prescription for r, since r is not a red-gray crossing. For instance, the diagrams of S illustrated in Figure 6.11(c) and (d) belong to $\mathcal{D}_{B,R}$.

We let $\mathcal{U}_{B,R}$ denote the set of diagrams of $\mathcal{D}_{B,R}$ that are unknot diagrams. We will show Lemma 12 for the current case ($S \neq B \cup R$) by proving that $|\mathcal{U}_{B,R}| \geq 2^{m/2}$ (we recall that in the stetement of Lemma 12, m is the number of common vertices of B and R).

The concept of a valid pair of edges was fundamental in the proof for the case $S = B \cup R$. Our next step here is to extend this concept in the current context, to valid pairs of paths.

Suppose that P, Q are internally disjoint paths of S with the same endpoints u, v, such that P (respectively Q) is contained in B (respectively, R). We remark that neither u nor v can be equal to b or r, since b is incident with two blue edges and two gray edges, and v is incident with two red edges and two gray edges. Further suppose that P does not contain b (as we have just remarked, b cannot be an endpoint of P, but it could be an internal vertex of P), and Q does not contain r. Finally, suppose that every internal vertex of P (respectively, Q), if any, is blue-gray (respectively, red-gray). That is, the only blue-red vertices of P and Q are precisely u and v. Then we say that (P, Q) is a *valid pair of paths* of (B, R).

For instance, in the shadow S of Figure 6.11 if we let P be the blue path uw_2v, and let Q be the red path uw_1v, then (P, Q) is a valid pair of (B, R). Also if we let P be the blue (single-edge) path uw_4 and we let Q be the red (single-edge) path uw_4, then (P, Q) is also a valid pair. In this same figure, if we let P be the blue path w_3bv and let Q be the red path w_3v, then (P, Q) is not a valid pair, since P contains the root b of B.

In analogy with the way we proceeded in the case $S = B \cup R$, here we will obtain from S a new shadow S' by means of a splitting process. Let (P, Q) be a valid pair of (B, R), where u, v are the common endpoints of P and Q. First we split u into two degree 2 vertices. There are two ways to do such a splitting, so that the result is a plane graph. We choose to split u so that, in the resulting graph, one of the resulting two degree 2 vertices is incident with P, and the other is incident with Q. We then proceed analogously with v: we split v so that one of its two resulting degree 2 vertices is incident with the remains of P, and the other one is incident with the remains of Q. Finally, we suppress the four degree 2 vertices obtained in the process. In Figure 6.11(b) we illustrate the shadow S' that results by applying this operation to the shadow S in part (a), using the blue path $P = uw_2v$ and the red path $Q = uw_1v$.

It is readily seen that S' inherits from S the property that it is the union of two straight-ahead cycles. These straight-ahead cycles can be naturally labelled B' and R', with the convention that B' is the straight-ahead cycle that contains the remains of the red path Q, and R' is the straight-ahead cycle that contains the remains of the blue path P. With this

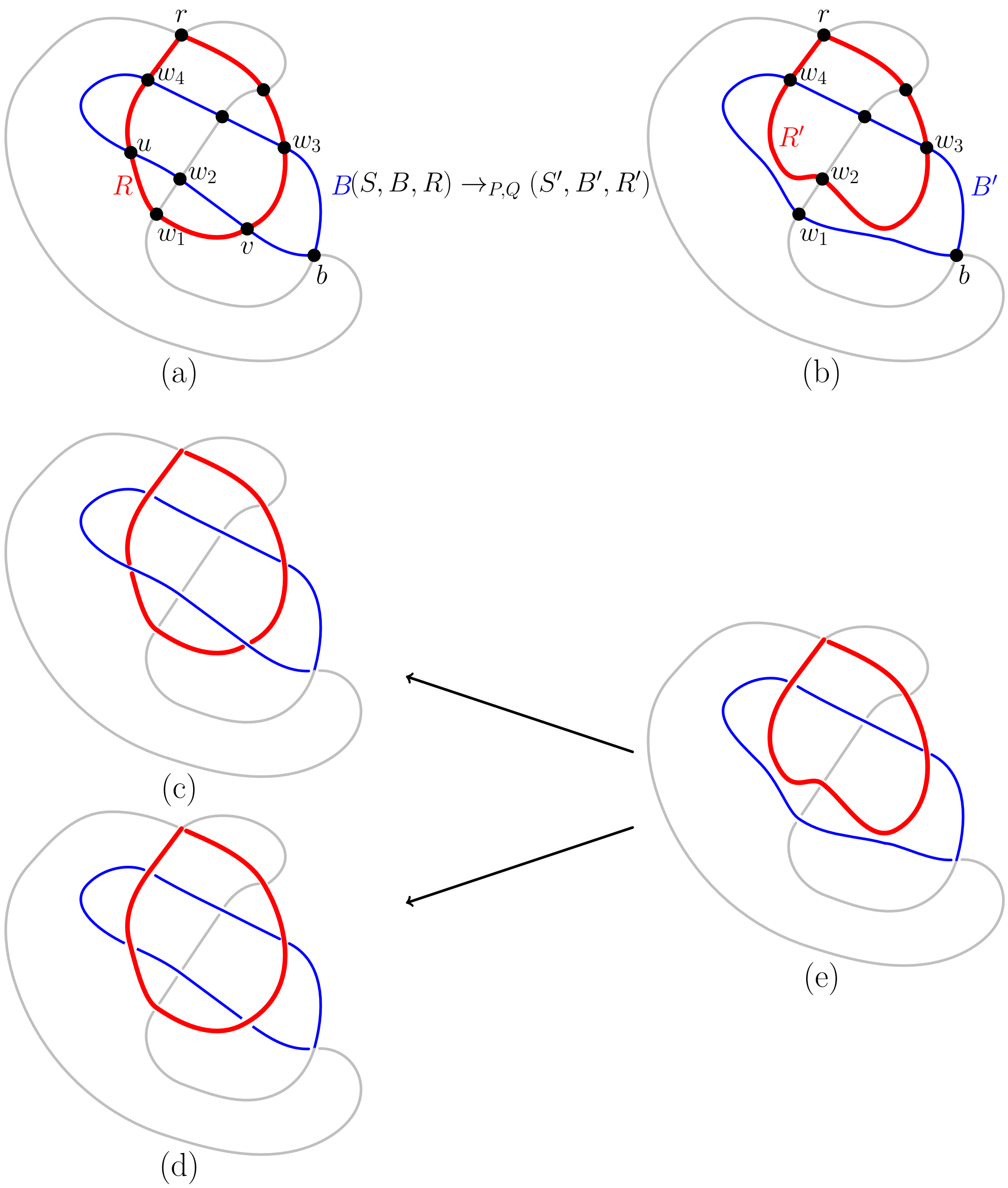


Figure 6.11: In (a) we have a shadow S, with two straight-ahead cycles B and R. By letting $P = uw_2v$ and $Q = uw_1v$, we have that (P, Q) is a valid pair of paths. We apply the splitting process to this pair (thus vertices u and v disappear), and obtain the shadow S' shown in (b). In (e) we have an unknot diagram D' of S', and in (c) and (d) we illustrate the unknot diagrams of S that are extensions of D'.

convention, all the edges of B that were not affected by this operation are edges of B' (so we can naturally colour the edges of B' blue), and all the edges of R that were not affected by this operation are edges of R' (so we can naturally colour the edges of R' red). Moreover, it is immediately seen that b is the root of B' and r is the root of R'. For this operation we write $(S, B, R) \to_{P,Q} (S', B', R')$. Whenever it is not needed to specify that we obtain S', B', and R' from the particular pair (P, Q), we simply write $(S, B, R) \to (S', B', R')$.

We now establish the following analogue of Claim C.

Claim F. *Suppose that B and R are disjoint. Then $|\mathcal{U}_{B,R}(S)| \geq 4$.*

Proof. Let $S_1 := S /\!\!/ B$, and let $S_2 := S_1 /\!\!/ R$. Thus S_2 is the shadow obtained by removing all the blue and red edges, and suppressing all the resulting degree 2 vertices. We note that every vertex in S_2 naturally corresponds to a vertex in S.

Let D, D_2 be diagrams of S and S_2, respectively. We say that D is an *extension* of D_2 if for every vertex in D_2, its prescription is the same as the prescription of its corresponding vertex in D.

Now let D_2 be an unknot diagram of S_2. There are exactly four diagrams D of S that satisfy the following: (i) D is an extension of D_2; and (ii) D satisfies (P1) and (P2). Indeed, (i) means that the prescription at every crossing that is gray-gray is induced from its prescription in D_2, and (ii) means that the prescription at every blue-gray or red-gray crossing is also determined (there are no blue-red crossings, since $B \cap R = \emptyset$). Thus the prescriptions at b and r are the only ones that have not been fixed for a diagram of S that satisfies these properties. The only freedom left are the prescription at b (two possibilities), and the prescription at r (two possibilities). Thus there are indeed exactly four diagrams of S that satisfy (i) and (ii). We finish the proof of Claim F by arguing that these four diagrams of S are unknot diagrams.

Let D be any of these four diagrams. Since D satisfies (P1), and B is disjoint from R, it follows that the strand σ_B in D that corresponds to B is an overstrand. Thus we can perform a generalized Reidemeister move to this strand, replacing it with the strand that consists solely of the point b. Let D' be the diagram obtained. Since D satisfies (P2), it follows that the strand σ_R in D' that corresponds to R is an overstrand. Thus we can perform a generalized Reidemeister move to this strand, replacing it with the strand that consists solely of the point r. As a result we obtain D_2. Since D_2 can be obtained from D by performing a sequence of two generalized Reidemeister moves, and D_2 is an unknot diagram, it follows that D is an unknot diagram of S.

Thus there are at least four unknot diagrams of S that are extensions of D_2. Each of these diagrams satisfies (P1) and (P2), and so it belongs to $\mathcal{U}_{B,R}$. Thus $|\mathcal{U}_{B,R}(S)| \geq 4$. □

Now we have the following analogue of Claim D.

Claim G. *Suppose that $(S, B, R) \to_{P,Q} (S', B', R')$. Then $|\mathcal{U}_{B,R}(S)| \geq 2|\mathcal{U}_{B',R'}(S')|$.*

Proof. Let u, v be the common endpoints of P and Q. We recall that u and v are the only vertices of S that get removed in the process of getting S'. Thus every vertex of S' corresponds naturally to a vertex of S.

Let D, D' be diagrams of S and S', respectively. We say that D is an *extension* of D' if for each vertex in S', its prescription in D' coincides with the prescription of its corresponding vertex in D.

Now let D' be a diagram in $\mathcal{U}_{B',R'}(S')$. There are exactly four diagrams D of S such that D is an extension of D', as there are two ways to give prescriptions to each of u and v. Clearly, these four diagrams are in $\mathcal{D}_{B,R}(S)$. To prove Claim G we show that at least two of these four diagrams of S are actually in $\mathcal{U}_{B,R}(S)$, that is, they are unknot diagrams of S.

Consider the diagram D_1 of S, such that D_1 is an extension of D', and such that at the crossings that correspond to u and v, the blue strands are overpasses. It is readily verified that then D' can be obtained from D_1 by a series of generalized Reidemeister moves and isotopies. Since D' is an unknot diagram, it follows that D_1 is also an unknot diagram. For instance, let S be the diagram in Figure 6.11(a), and let S' is the diagram in Figure 6.11(b) (so that $(S, B, R) \to_{P,Q} (S', B', R')$). If we let D' be the unknot diagram of S' shown in Figure 6.11(e) (it is readily verified that this is indeed unknot), then the diagram D_1 just described is the one shown in Figure 6.11(c).

Finally, let D_2 be the diagram of S, such that D_2 is an extension of D', and such that at the crossings that correspond to u and v, the red strands are overpasses. A totally analogous argument to the one used for D_1 shows that D_2 is also an unknot diagram. In our running example in Figure 6.11, the diagram D_2 is the one shown in Figure 6.11(d).

We have thus proved that if D' is a diagram in $\mathcal{U}_{B',R'}(S')$, then there are two diagrams D of S such that D is an extension of D', and D is in $\mathcal{U}_{B,R}(S)$. We finally observe that if D', F' are distinct diagrams in $\mathcal{U}_{B',R'}(S')$, and D, F are diagrams in $\mathcal{U}_{B,R}(S)$ such that D is an extension of D' and F is an extension of F', then D is distinct from F. These two facts imply that $|\mathcal{U}_{B,R}(S)| \geq 2|\mathcal{U}_{B',R'}(S')|$. □

We now state and prove the following analogue of Claim E.

Claim H. *Let S be a shadow and let B, R be straight-ahead cycles of S, such that $S \neq B \cup R$. Set $S_1 := S$. Then there is a sequence $S_1, S_2, \ldots, S_p$ of shadows with the following properties. (i) For $i = 1, 2, \ldots, p$, S_i has two straight-ahead cycles B_i, R_i, such that $S_i \neq B_i \cup R_i$; (ii) $(S_i, B_i, R_i) \to (S_{i+1}, B_{i+1}, R_{i+1})$ for $i = 1, 2, \ldots, p-1$; and (iii) in the shadow S_p, the straight-ahead cycles B_p and R_p are disjoint.*

Proof. At this point we note that in the present case (there are straight-ahead cycles B and R of S, and $S \neq B \cup R$), B and R must have an even number of vertices in common. This is an immediate consequence of the Jordan curve theorem.

We prove the claim by induction on the number of common vertices of B and R. In the base case B and R are disjoint, and so the sequence that consists of the single shadow $S_1 = S$ obviously satisfies the required conditions.

For the inductive step, we assume that Claim H is true whenever B and R have exactly $2k-2$ vertices in common, for some integer $k \geq 1$, and consider a shadow S in which B and R have exactly $2k$ vertices in common. Recall that $S_1 := S$, and let $B_1 := B$ and $R_1 := R$.

We claim that it suffices to show that there exists a valid pair of paths (P_1, Q_1) of (B_1, R_1). Indeed, the existence of this valid pair yields the existence of a shadow S_2, with straight-ahead cycles B_2, R_2, such that $S_2 \neq B_2 \cup R_2$ and $(S_1, B_1, R_1) \to (S_2, B_2, R_2)$. The inductive hypothesis, applied to S_2, B_2, and R_2, then implies the existence of a sequence $S_2, S_3, S_4, \ldots, S_p$ of shadows with the required properties, and so $S_1, S_2, \ldots, S_p$ is the desired sequence. Thus we finish the proof by showing the existence of a valid pair of paths (P_1, Q_1) of (B_1, R_1).

Let β_1, ρ_1 be the simple closed curves underlying the straight-ahead cycles B_1 and R_1, respectively. Recall that b is the root of $B_1 = B$ and r is the root of $R_1 = R$. Note that since $S \neq B \cup R$, neither b nor r (regarded as points in β_1 and ρ_1, respectively) are in $\beta_1 \cap \rho_1$. By Claim A, (β_1, ρ_1) has a digon (α_1, γ_1) such that neither b nor r is in $\alpha_1 \cup \gamma_1$. Let P_1 be the path in B_1 that corresponds to α_1, and let Q_1 be the path of R_1 that corresponds to γ_1. Then (P_1, Q_1) is a valid pair of (B_1, R_1). □

We finally show how Lemma 12 in the case $S \neq B \cup R$ follows from Claims F, G, and H.

Conclusion of the proof of Lemma 12 for the case $S \neq B \cup R$. Recall that m denotes the number of common vertices of B and R. We will show that $|\mathcal{U}_{B,R}| \geq 2^{m/2}$. This obviously implies Lemma 12, since $\mathcal{U}_{B,R}(S) \subseteq \mathcal{U}(S)$.

We recall from the proof of Claim H that the number m is even. Let $k := m/2$. The aim is then to show that $|\mathcal{U}_{B,R}(S)| \geq 2^k$.

Consider the sequence $S = S_1, S_2, \ldots, S_p$ guaranteed by Claim H. We note that the number of common vertices of B_{i+1} and R_{i+1} is exactly two less than the number of common vertices of B_i and R_i, for $i = 1, 2, \ldots, p-1$. Thus $m = 2k = 2(p-1)$, and so $p = k+1$.

By Claim G, we have that $|\mathcal{U}_{B_i,R_i}(S_i)| \geq 2|\mathcal{U}_{B_{i+1},R_{i+1}}(S_{i+1})|$ for $i = 1, 2, \ldots, p-1$. Thus $|\mathcal{U}_{B,R}(S)| = |\mathcal{U}_{B_1,R_1}(S_1)| \geq 2^{p-1}|\mathcal{U}_{B_p,R_p}(S_p)|$. Since B_p and R_p are disjoint, from Claim F we have that $|\mathcal{U}_{B_p,R_p}(S_p)| \geq 4$, and so $|\mathcal{U}_{B,R}(S)| \geq 2^{p-1} \cdot 4 = 2^{p+1} = 2^{k+2} > 2^k$. □

6.5 Finding a subshadow with two suitable straight-ahead cycles: proof of Lemma 13

Let S be a shadow with n vertices. Let $C_1, C_2, \ldots, C_p$ be a cycle decomposition of S. We recall that the cycles $C_2, \ldots, C_p$ are not necessarily cycles in S. (The vertex-free cycle C_p is certainly not a cycle of S). Indeed, for $i \geq 2$, C_i is a cycle of S_i, but since S_i is not a subgraph of S (due to the suppression of degree 2 vertices), C_i is not necessarily a cycle back in S. On the other hand, for $i = 1, 2, \ldots, p$ there is a cycle D_i in S that is a subdivision of C_i. We say that $D_1, D_2, \ldots, D_p$ is the *primary sequence* associated to the cycle decomposition $C_1, C_2, \ldots, C_p$. We note that the cycles $D_1, D_2, \ldots, D_p$ are pairwise edge-disjoint, and that $S = D_1 \cup D_2 \cup \cdots \cup D_p$ is the union of these graphs.

The workhorses for the proof of Lemma 13 are the following two statements:

Claim A. *Let S be a shadow with n vertices. Let $C_1, C_2, \ldots, C_p$ be a cycle decomposition of S, and let $D_1, D_2, \ldots, D_p$ be the associated primary sequence. Then there exist j, k, $1 \leq j < k \leq p$, such that D_j and D_k have at least $2n/p^2$ common vertices.*

Claim B. *Let S be a shadow with n vertices. Let $C_1, C_2, \ldots, C_p$ be a cycle decomposition of S, and let $D_1, D_2, \ldots, D_p$ be the associated primary sequence. For any j, k, $1 \leq j < k \leq p$, there exists a subshadow T of S that has two straight-ahead cycles B and R such that D_j is a subdivision of B, and D_k is a subdivision of R.*

Deferring the proofs of these claims for the moment, we show they easily imply Lemma 13.

Proof of Lemma 13. Let S be a shadow with n vertices. Let $C_1, C_2, \ldots, C_p$ be a cycle decomposition of S, and let $D_1, D_2, \ldots, D_p$ be the associated primary sequence. By assumption, $p \leq \sqrt[3]{n}$. Thus it follows from Claim A that there exist j, k, $1 \leq j < k \leq p$, such that D_j and D_k have at least $2n/p^2 = 2\sqrt[3]{n}$ common vertices.

By Claim B, there exists a subshadow T of S that has two straight-ahead cycles B and R such that D_j is a subdivision of B, and D_k is a subdivision of R. Now it suffices to note that, since D_j is a subdivision of B, and D_k is a subdivision of R, then every vertex that is in both D_j and D_k is also a vertex in both B and R. Thus T is a subshadow of S that has two straight-ahead cycles B and R, such that B and R have at least $2\sqrt[3]{n}$ common vertices. □

We conclude the section by proving Claims A and B.

Proof of Claim A. Let n be the number of vertices of S. We recall that each edge in S belongs to D_i for exactly one $i \in \{1, 2, \ldots, p\}$. Moreover, if v is a vertex in S, then v is incident with two edges in D_i and two edges in D_ℓ, for exactly two distinct $i, \ell \in \{1, 2, \ldots, p\}$.

For each two elements D_i and D_ℓ in the primary sequence of S, we denote $\mathrm{n}(i, \ell)$ to be the number of vertices in $D_i \cup D_\ell$, then the sum over all pairs of elements in the primary sequence $\sum_{i=1}^{p-1}(\sum_{\ell=i+1}^{p} \mathrm{n}(i, \ell))$ must be exactly n.

Seeking a contradiction, suppose that for any pair i, ℓ of distinct elements of $\{1, 2, \ldots, p\}$, the number $\mathrm{n}(i, \ell)$ is less than $2n/p^2$. There are $\binom{p}{2}$ distinct pairs in $\{1, 2, \ldots, p\}$, and so the toatl sum is less than $2n/p^2 \cdot \binom{p}{2} = n(p-1)/p$. But this is impossible since $n(p-1)/p < n$ as we have observed.

Thus there are distinct $i, \ell \in \{D_1, D_2, \ldots, D_p\}$ such that D_i and D_ℓ have at least $2n/p^2$ common vertices in S. □

Proof of Claim B. We prove first this claim for the case $j = 1$, reading as follows *Let S be a shadow with n vertices. Let $C_1, C_2, \ldots, C_p$ be a cycle decomposition of S, and let $D_1, D_2, \ldots, D_p$ be the associated primary sequence. For any k, $1 < k \leq p$, there exists a subshadow T of S that has two straight-ahead cycles B and R such that D_1 is a subdivision of B, and D_k is a subdivision of R.*

We prove this statement by induction on p. In the base case $p = 2$. Thus $k = 2$, and both D_1 and D_2 are straight-ahead cycles of S. The statement then trivially follows by setting $T = S$, $B = D_1$ and $R = D_2$.

For the inductive step we work with a shadow S as given in the statement, and assume that the statement holds for every cycle decomposition (of every shadow) of size at most $p - 1$, for some $p \geq 3$.

We make essential use of the fact that every nontrivial shadow has at least two straight-ahead cycles. From this it immediately follows that there is an $\ell \in \{2, 3, \ldots, p\}$ such that

D_ℓ is a straight-ahead cycle of S. If $\ell = k$ then we are done by setting $T = S, B = D_1$, and $R = D_k$. Thus we may assume that $\ell \neq k$.

Let $U = S /\!\!/ D_\ell$. In U we have a collection $F_1, F_2, \ldots, F_{\ell-1}, F_{\ell+1}, \ldots, F_p$ such that D_i is a subdivision of F_i for each $i \in \{1, 2, \ldots, \ell - 1, \ell + 1, \ldots, p\}$. Moreover, the cycle decomposition of S induces a cycle decomposition $I_1, I_2, \ldots, I_{\ell-1}, I_{\ell+1}, \ldots, I_p$ of U, which has $F_1, F_2, \ldots, F_{\ell-1}, F_{\ell+1}, \ldots, F_p$ as its primary sequence. By the inductive hypothesis, there is a subshadow T of U that has two straight-ahead cycles B and R such that F_1 is a subdivision of B, and F_k is a subdivision of R.

Since U is a subshadow of S, then T is a subshadow of S. Since D_1 (respectively, D_k) is a subdivision of F_1 (respectively, F_k), then D_1 (respectively, D_k) is a subdivision of B (respectively, R). Thus T is a subshadow of S with the required properties.

Now we go on to the case when $j > 1$ for which we will make use of the statment above. Let $S_{i+1} = S_i /\!\!/ C_i$, for $i = 1, 2, \ldots, p - 1$. Then $C_j, C_{j+1}, \ldots, C_p$ is a cycle decomposition of S_j. If we let $F_i := C_{j+i-1}$ for $i = 1, 2, \ldots, p - j + 1$, we have that $F_1, F_2, \ldots, F_{p-j+1}$ is a cycle decomposition of S_j. Let $H_1, H_2, \ldots, H_{p-j+1}$ be the primary sequence of this cycle decomposition. By assumption, we have that there is a subshadow T of S_j that has two straight-ahead cycles B, R such that H_1 is a subdivision of B, and H_{k-j+1} is a subdivision of R. We note that D_j is a subdivision of H_1, and H_{k-j+1} is a subdivision of D_k. Thus T is a subshadow of S (since T is a subshadow of S_j, and S_j is a subshadow of S) that has two straight-ahead cycles B, R such that D_j is a subdivision of B, and D_k is a subdivision of R, as required. $\square$

Chapter 7

Knotted diagrams obtained from a given shadow

7.1 Introduction

From our work in the previous chapter we know that every shadow has many unknot diagrams associated to it. The driving questions behind the work reported in the present chapter are the following. Let us say that a shadow S is *simple* if every assignment on S yields an unknot diagram. Which shadows are simple? If S is not simple, then is it necessarily true that there is an assignment on S that yields a diagram of the trefoil knot? We focus as a first step on the trefoil knot, since it is the nontrivial knot with the smallest *crossing number* (three). As we will see shortly, we also investigated this question for the figure-eight knot, unveiling a totally different scenario than for the trefoil knot.

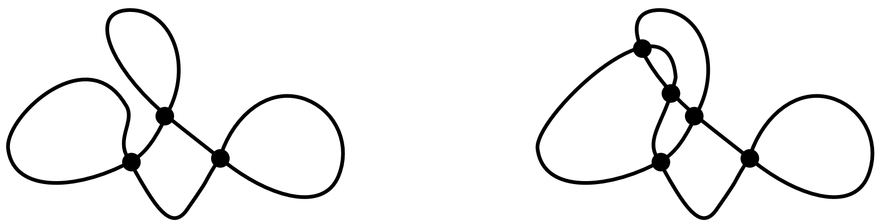
The following statement characterizes which shadows are simple.

Theorem 14. *Let S be a shadow. Then S is simple if and only if every vertex of S is a cut-vertex.*

We answered the second question posed above positively:

Theorem 15. *If S is a non-simple shadow, then there is a diagram associated to S that is a diagram of the trefoil knot.*

It is natural to ask how far we can get along the lines of Theorem 15. The next obvious step is to investigate the figure-eight knot (see Figure 7.2), which is the only knot with *crossing number* four, namely, the minimum number of crossings among all diagrams of the figure-eight knot is exactly four. One might think that for the figure-eight knot, a result analogous to Theorem 15 could hold, for all sufficiently large non-simple shadows, or at least perhaps for all sufficiently large shadows that do not have any cut-vertex. However, this is not the case, not only for the figure-eight knot, but for every knot with even crossing number:



(a) (b)

Figure 7.1: Every vertex in the shadow in (a) is a cut-vertex, and so by Theorem 14 this shadow is simple. The shadow in (b) has only one cut-vertex, and so by Theorem 14 it is not simple. Moreover, by Theorem 15 this last shadow has an assignment that yields a diagram of the trefoil knot.

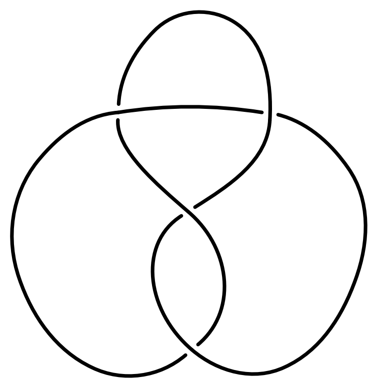


Figure 7.2: The figure-eight knot.

Observation 16. *Let K be any knot with even crossing number, and let m be any positive integer. Then there exists a shadow S with more than m vertices, none of which is a cut-vertex, with the following property. For every diagram D of S, the knot corresponding to D is not equivalent to K. In particular, there exist arbitrarily large shadows, with no cut-vertices, that have no diagram that corresponds to the figure-eight knot.*

These statements are proved in Sections 7.2, 7.3, and 7.4, respectively.

7.2 Characterization of simple shadows: proof of Theorem 14

First we show that if S is a simple shadow, then each of its vertices must be a cut-vertex. We prove this by induction on the number of vertices of S. In the base case S consists of exactly one vertex v, and two loop-edges based on v. Thus this only vertex of S is a cut-vertex, so the statement holds for the base case. For the inductive step, we assume that the statement holds for every shadow with fewer than n vertices, for some $n \geq 2$, and let

S be a simple shadow with n vertices. We will show that then every vertex of S must be a cut-vertex. We first derive a contradiction by assuming that S has no cut-vertex. For suppose that none of the vertices of S is a cut-vertex. Consider any *alternating diagram* associated to S, meaning that the prescriptions of the crossings alternate between under and over as one travels along any eulerian walk. Since S has no cut vertices, then D is a reduced diagram. Since the number of crossings of D is n (this follows by the validity of the first Tait Conjecture [29, 40, 46]), then D is not an unknot diagram, contradicting the assumption that S is simple. Thus we may assume that S has at least one cut-vertex. Let v be a cut-vertex of S. Then there are subgraphs H, K of G, each having at least one edge, such that $H \cup K = G$ and $H \cap K$ (the subgraph of G that is a subgraph of both H and K) is the graph that consists solely of v. Let H' (respectively, K') be the graph obtained by suppressing (the degree 2 vertex) v in H(respectively, K). Then H' and K' are shadows on their own right with less than n crossings. We claim that both H' and K' are simple. For suppose that H' is not simple, and let $D_{H'}$ be a knotted diagram of H'. Let K'' be a graph that is a mirror image of K' with respect to some line ℓ traversing K' trough v and consider $D_{K''}$ to be an unknot diagram of K''. If we suitably sum $D_{H'}$ and $D_{K''}$, and then perform a simple twist over K'' corresponding to a flip (mirror image of K'') with respect to ℓ, then the result is a diagram D of S. Since $D_{H'}$ is not an unknot diagram, it follows that D is not an unknot diagram. The existence of an unknot diagram associated to S implies that S is not simple, contradicting our hypothesis. Thus both H' and K' are simple, as claimed. The inductive hypothesis then implies that every vertex of H' and K' is a cut-vertex, and from this it immediately follows that every vertex of S is a cut-vertex.

Now we show, also by induction on the number of vertices, that if every vertex of S is a cut-vertex, then S is a simple shadow. In the base case S consists of exactly one vertex v, and two loop-edges based on v. The two possible assignments of S evidently yield unknot diagrams, and so S is simple. For the inductive step, we assume that the statement holds for every simple shadow with fewer than n vertices, for some $n \geq 2$, and let S be a shadow with n vertices, such that every vertex of S is a cut-vertex. We consider an arbitrary diagram D associated to S, and show that D is an unknot diagram.

Consider any vertex v of S. By assumption, v is a cut-vertex, and so its corresponding crossing in D is a reducible (also called nugatory) crossing. We can thus reduce, the diagram D to a diagram D' with one fewer crossing by performing a Reidemeister move of type I, then D and D' are equivalent. It is readily verified that the shadow associated to D' also has the property that every vertex is a cut-vertex, and so by the induction hypothesis D' is an unknot diagram. By the previous observation, D is then also an unknot diagram. □

7.3 Every non-simple shadow yields the trefoil knot: proof of Theorem 15

We first observe that it suffices to prove Theorem 15 for shadows that do not contain any cut-vertex. This follows easily by an inductive argument. For suppose Theorem 15 holds for every non-simple shadow on $n-1$ vertices. Then $n-1 \geq 3$, as a quick analysis shows that every shadow with fewer than 3 vertices is simple. The base case of this induction is precisely a shadow of a trefoil knot, for which the theorem obviously holds. Now consider a non-simple shadow S with n vertices. If S has a cut-vertex v, then at least one of the shadows we get by splitting at v is non-simple, and so by the inductive hypothesis it has a trefoil diagram. By giving an unknot assignment to the other shadow, the result is a trefoil diagram for S.

Thus let S be a shadow that does not contain any cut-vertex, and consider a straight-ahead cycle C of S. Let c be the root of C. Let $d_1 d_2 e_1 e_2$ be the rotation at c, where d_1, d_2 are the edges that are part of C. Now let W be the straight-ahead walk whose first edge is e_1 and whose last edge is e_2. Thus C and W are edge-disjoint, and $S = C \cup W$.

We assume that both e_1 and e_2 are outside the closed disc Δ bounded by C. This is the case shown in Figure 7.3 where the cycle C is represented by thick black edges. The alternative case (e_1 and e_2 are inside Δ) is handled in a totally analogous manner.

As we walk along W starting at c, at some point we must encounter some vertex of C other than c. This follows since (i) c is not a cut-vertex, and so C must have some vertex other than c; and (ii) since every vertex of S has degree four, then every vertex in C is also a vertex of W. Let u be the first vertex of C, other than c, that we find as we traverse W starting at c. Let W_1 be the part of W that we have traversed until (and including) we reach u, and let f_1 be the ending edge of W_1; thus f_1 is incident with u.

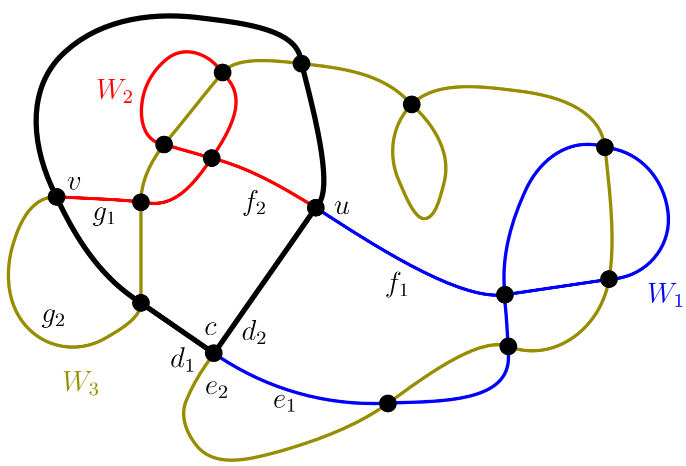


Figure 7.3: A (non-simple) shadow S containing a straight-ahead cycle C (thick edges) with root vertex c. The complement of C in S is the union of three straight-ahead walks W_1, W_2 and W_3. The walk W_2 is entirely contained in the disk bounded by the simple closed curve representing C, and the walk W_1 is completely contained in the complement of this disk.

We now continue the traversal of W at the point we interrupted it. Thus we start at u, and the first edge f_2 of this second part of W is inside Δ; this follows since all the edges of W_1, and in particular f_1, are necessarily outside Δ, and W is a straight-ahead walk. We continue this second part of the traversal of W until we reach a vertex v of C. Since this part of the traversal is contained in Δ, it follows that v cannot be c. We let g_1 be the last edge of W traversed in this second stage (thus g_1 is incident with v), and let W_2 be this subwalk of W whose first edge is f_2 and whose last edge is g_1. We emphasize that W_2 is contained inside Δ.

We now complete the traversal of W, starting at the other edge g_2 of W incident with v, and ending at e_2; let W_3 be this subwalk of W. Note that since g_1 is inside Δ, then g_2 is outside Δ, but there may be parts of W_3 that lie inside Δ. Thus W is the concatenation of W_1, W_2, and W_3.

We claim that it suffices to consider the case in which each of W_1, W_2, and W_3 is actually a path. To see this, suppose that W_1 is not a path. Then W_1 necessarily contains a straight-ahead cycle F. Let $S' := S /\!/ F$. It is straightforward to see that if S' has a trefoil diagram, then S also has a trefoil diagram: a trefoil diagram of S' can be extended to a trefoil diagram of S by making the strand corresponding to F an overstrand. Let T be the subshadow of S that results by recursively applying the same procedure to any remaining straight-ahead cycles of W_1 (or of W_2, or of W_3). Then, in T, the walks W_1, W_2, and W_3 have been reduced to paths P_1, P_2, and P_3, respectively, and it suffices to show that T has a trefoil diagram.

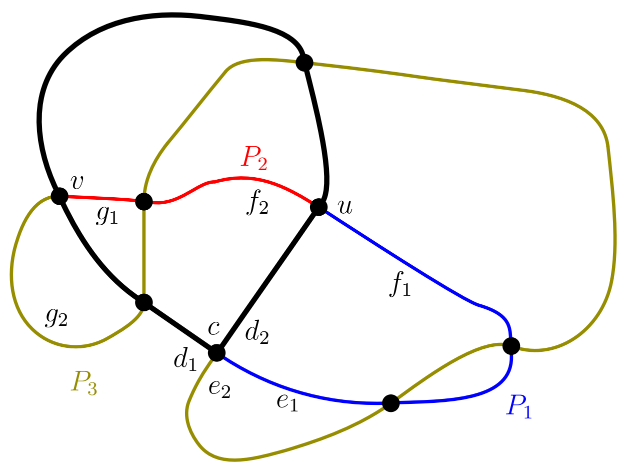


Figure 7.4: The shadow T obtained from the shadow S in Figure 7.3 after removing all straight-ahead cycles from W_1, W_2 and W_3. In T these walks have been reduced to paths P_1, P_2, and P_3. The shadow T has an assignment that yields a trefoil knot diagram. Since T is a subshadow of S, it follows that S has an assignment with the same property.

Let us pause for a moment to review where we stand. We have a shadow T, with a straight-ahead cycle J (obtained from C) with root is c. Indeed, in the process of eliminating straight-ahead cycles from W_1, W_2, and W_3, some of these straight-ahead cycles have had vertices in common with C: these vertices became degree 2 vertices in the process, and were suppressed. We let J denote the cycle that is induced from C, so that C is a subdivision of J. Now the remains of W (in case we eliminated straight-ahead cycles) form a walk U that starts and ends at c, and that is the concatenation of three paths P_1, P_2, and P_3. The path P_1 is contained outside Δ, the path P_2 is contained inside Δ, and the path P_3 has its first edge and its last edge outside Δ, but it may have edges inside Δ. To finish the proof we need to show that T has a trefoil diagram.

Consider now a diagram D in which the strand corresponding to P_3 is an overstrand. We note that this fixes the prescriptions at all the vertices of T, with the exception of c, u, and v, whose prescriptions we leave open for the moment. The fact that this fixes the prescription at every vertex $w \notin \{c, u, v\}$ follows simply because the structure of T is such that every vertex $w \notin \{c, u, v\}$ is necessarily an internal vertex of P_3.

We can now perform a generalized Reidemeister move on the strand corresponding to P_3, transforming it into a strand that is contained outside Δ and does not have any internal crossings. The result is a diagram D' with only three crossings (those corresponding to c, u, and v), whose shadow consists of a 3-cycle plus one parallel edge added to each edge. Thus if the prescriptions at c, u, and v are suitably chosen in D, then D' is a trefoil diagram, see Figure 7.5. Since D is equivalent to D', it follows that there exist suitable prescriptions at c, u, and v such that D is a trefoil diagram.

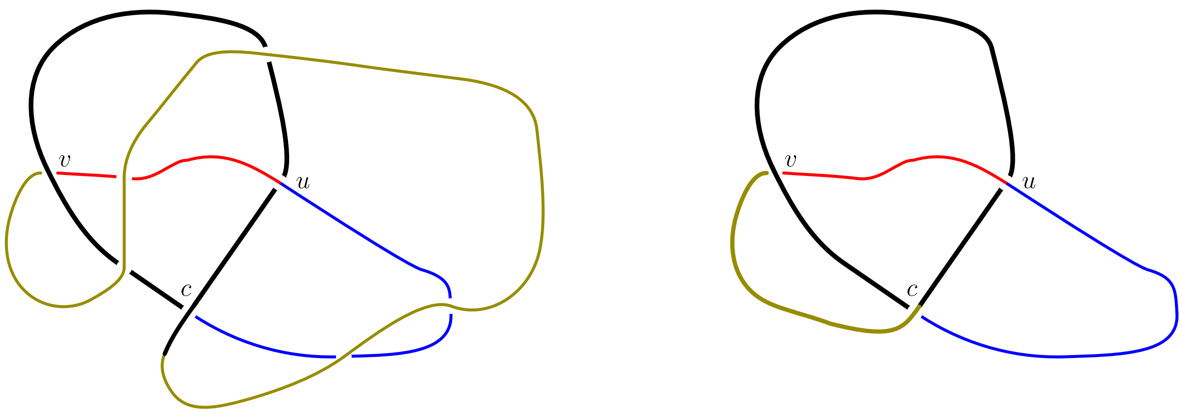


Figure 7.5: On the left hand side of this figure we have a diagram associated to the shadow in Figure 7.4. Note that the strand corresponding to P_3 is an overstrand. Thus we can perform a generalized Reidemeister move on this strand, obtaining the trefoil diagram on the right hand side.

7.4 Shadows that do not induce any knot with even crossing number: proof of Observation 16

For each odd integer $n \geq 5$, let C_n be a shadow obtained from a cycle with n vertices by adding a parallel edge to each edge. We call C_n the *n-circulant*. In Figure 7.6 we illustrate the 7-circulant C_7.

To prove the observation it suffices to show the following: for every odd integer n, if D is a diagram of C_n, then the knot corresponding to D has odd crossing number.

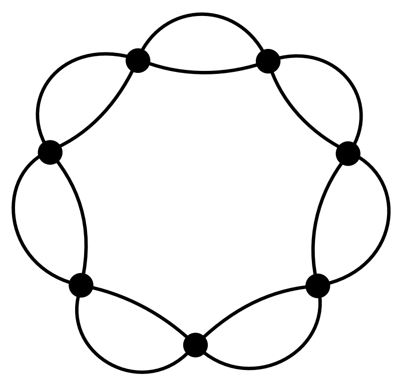


Figure 7.6: The 7-circulant C_7. No diagram associated to C_7 is equivalent to a diagram of a knot with even crossing number (and in particular to the figure-eight knot).

A straightforward inductive argument shows that every diagram of C_n is equivalent to an alternating diagram of C_t, for some odd $t \leq n$. We recall that if D is an alternating reduced diagram of a knot, then the number of crossings of D is the crossing number of this knot. Thus if D is a diagram of C_n, then the knot corresponding to D has odd crossing number, as claimed. $\square$

Chapter 8

Concluding remarks and open questions

It seems natural to ask if one could hope to prove a much better bound than the one given by Theorem 9. To explore this question, we let $\mathcal{U}_{\min}(n)$ denote the best possible bound. That is, for each positive integer n,

$$\mathcal{U}_{\min}(n) := \min_{|S|=n} \mathcal{U}(S),$$

where the minimum is taken over the number of unknot diagrams $\mathcal{U}(S)$ of all shadows S with n vertices.

Under this terminology, Theorem 9 gives the superpolynomial bound $\mathcal{U}_{\min}(n) \geq 2^{\sqrt[3]{n}}$. Leaving aside relatively marginal possible improvements, such as showing that $\mathcal{U}_{\min}(n) \geq 2^{\sqrt{n}}$, the important question seems to be the following.

Question 17. *Is $\mathcal{U}_{\min}(n)$ an exponential, or a subexponential function?*

With this question in mind, we start by noting that [11, Proposition 10] implies the existence of a constant $d < 1$ such that, for all sufficiently large n, the probability that a diagram is unknot is at most d^n. From this it immediately follows that $\mathcal{U}_{\min}(n) < b^n$, for some constant b, $1 < b < 2$. We also recall the following result of Chapman:

Theorem 18 ([11, Theorem 1])**.** *As the number of crossings n of a randomly sampled knot diagram grows large, the probability that the diagram is knotted tends to* 1 *exponentially quickly.*

Although it must be possible to extract, from the proof of Theorem 1 in [11], an explicit constant $d \in (1, 2)$ such that $\mathcal{U}_{\min}(n) < d^n$, we have not been able to estimate it.

The task of finding such an explicit constant d is obviously complicated by the difficulty of estimating $|\mathcal{U}(S)|$ for a shadow S with a large number of vertices. This is a hopeless problem except for particularly simple families of shadows. For instance, this calculation can be carried out exactly for the family of circulant shadows (see Section 7.4). A tedious but elementary exercise shows that C_n has exactly $\binom{n+1}{\frac{n+1}{2}} \approx \sqrt{\frac{8}{\pi(n+1)}} \cdot 2^n$ unknot diagrams associated to it.

An approach to derive upper bounds for $\mathcal{U}_{\min}(n)$ is to consider any fixed shadow T, and take the connected sum T^k of k copies of T. (One can easily extend to shadows the definition of the connected sum of diagrams). Suppose that T has m vertices, and $|\mathcal{U}(T)| = t$. Since the connected sum of two diagrams is unknot if and only if each of the diagrams is unknot, it follows that T^k has exactly t^k unknot diagrams. Thus T^k is a shadow with $n := km$ vertices and $|\mathcal{U}(T^k)| = t^k = (t^{k/n})^n = (t^{1/m})^n$.

For instance, we can apply this idea to a shadow of the trefoil knot with 3 vertices, which has 6 unknot diagrams. With this construction we obtain that for every positive integer n divisible by 3 there is a shadow S_n with n vertices such that $|\mathcal{U}(S_n)| = (6^{1/3})^n \approx (1.861)^n$. This shows that, for every positive integer n divisible by 3, $\mathcal{U}_{\min}(n) \leq (6^{1/3})^n$.

To improve this upper bound we need to exhibit a shadow T with m vertices and t unknot diagrams, such that $t^{1/m} < 6^{1/3}$. The best result we have found so far in this direction is using the shadow depicted in Figure 8.1. This shadow has 16 vertices, and using SnapPy [13] we found that it has at most 6416 unknot diagrams. Thus it follows that, for every positive integer n divisible by 16, $\mathcal{U}_{\min}(n) \leq (6416^{1/16})^n \approx (1.729)^n$.

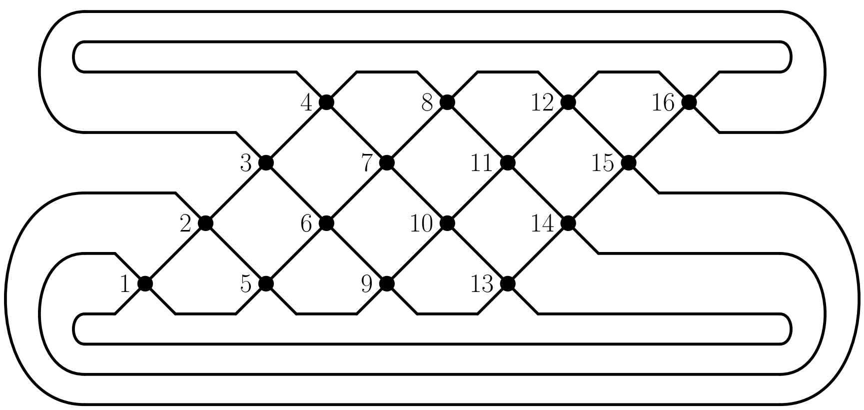


Figure 8.1: This shadow has 16 vertices and at most 6416 unknot diagrams. By taking the connected sum of k copies of this shadow, for every $k \geq 1$, it follows that for every positive integer n divisible by 16 there is a shadow with n vertices and $(6416^{1/16})^n \approx (1.729)^n$ unknot diagrams.

Returning to the discussion in Section 7.4, we note that the arbitrarily large shadows that do not produce any knot with even crossing number have a very special structure: they are "long and thin". In structural graph theory terminology, this infinite family of shadows has bounded (actually, very small) path-width [39]. It seems natural to wonder if there exist infinite families of shadows with similar properties, but with unbounded path-width. We suspect that this is not the case:

Conjecture 19. *For each fixed knot K, there is a constant $c := c(K)$ with the following property. If S is a shadow with path-width at least c, then there is a diagram associated to S whose corresponding knot is equivalent to K.*

An alternative, perhaps more natural version of this conjecture involves the notion of tree-width, rather than path-width. We refer the reader to the standard reference [39] for a comprehensive discussion on this standard graph theoretical parameter. For this discussion we recall that a graph with large tree-width contains a large planar grid as a minor. Working with a fixed diagram D, by this property it seems reasonable to expect that D can be obtained from every shadow with sufficiently large tree-width. Thus we put forward the following variant of the previous conjecture. This is actually a weaker version of Conjecture 19, as graphs with large tree-width have large path-width as well.

Conjecture 20. *For each fixed knot K, there is a constant $d := d(K)$ with the following property. If S is a shadow with tree-width at least d, then there is a diagram associated to S whose corresponding knot is equivalent to K.*

Bibliography

[1] Bernardo M. Ábrego and Silvia Fernández-Merchant, *A lower bound for the rectilinear crossing number*, Graphs Combin. **21** (2005), no. 3, 293–300.

[2] Bernardo M. Ábrego, Silvia Fernández-Merchant, and Gelasio Salazar, *The rectilinear crossing number of K_n: closing in (or are we?)*, Thirty essays on geometric graph theory, Springer, New York, 2013, pp. 5–18.

[3] Colin C. Adams, *The knot book. An elementary introduction to the mathematical theory of knots*, American Mathematical Society, Providence, RI, 2004.

[4] Colin Adams, Thomas Crawford, Benjamin DeMeo, Michael Landry, Alex Tong Lin, MurphyKate Montee, Seojung Park, Saraswathi Venkatesh, and Farrah Yhee, *Knot projections with a single multi-crossing*, J. Knot Theory Ramifications **24** (2015), no. 3, 1550011, 30.

[5] Pankaj K. Agarwal, Eran Nevo, János Pach, Rom Pinchasi, Micha Sharir, and Shakhar Smorodinsky, *Lenses in arrangements of pseudo-circles and their applications*, J. ACM **51** (2004), no. 2, 139–186.

[6] J. W. Alexander and G. B. Briggs, *On types of knotted curves*, Ann. of Math. (2) **28** (1926/27), no. 1-4, 562–586.

[7] J. Arsuaga, T. Blackstone, Y. Diao, E. Karadayi, and M. Saito, *Linking of uniform random polygons in confined spaces*, J. Phys. A **40** (2007), no. 9, 1925–1936.

[8] Gregory R. Buck, *Random knots and energy: elementary considerations*, J. Knot Theory Ramifications **3** (1994), no. 3, 355–363. Random knotting and linking (Vancouver, BC, 1993).

[9] Sergio Cabello, Éric Colin de Verdière, and Francis Lazarus, *Finding shortest non-trivial cycles in directed graphs on surfaces*, J. Comput. Geom. **7** (2016), no. 1, 123–148.

[10] Jason Cantarella, Harrison Chapman, and Matt Mastin, *Knot probabilities in random diagrams*, J. Phys. A **49** (2016), no. 40, 405001, 28.

[11] Harrison Chapman, *Asymptotic laws for knot diagrams*, Disc. Math. and Theoret. Comp. Sci. proc. **BC** (2016), 323–334. FPSAC 2016 (Vancouver, Canada).

[12] ———, *Asymptotic laws for random knot diagrams.* arXiv:1608.02638v2 [math.GT]. Available at `https://arxiv.org/abs/1608.02638`.

[13] 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` (17/02/2017).

[14] R.H. and Fox Crowell R.H., *Introduction to knot theory*, Graduate texts in mathematics, vol. 57, Springer-Verlag, 1977.

[15] M. Delbruck, *Knotting problems in biology*, Richard Bellman, editor. Mathematical Problems in the Biological Sciences. Proceedings of Symposia in Applied Mathematics. 14, American Mathematical Society, 1962, pp. 55–63.

[16] Yuanan Diao, Claus Ernst, and Uta Ziegler, *Generating large random knot projections*, Physical and numerical models in knot theory, Ser. Knots Everything, vol. 36, World Sci. Publ., Singapore, 2005, pp. 473–494.

[17] Y. Diao, J. C. Nardo, and Y. Sun, *Global knotting in equilateral random polygons*, J. Knot Theory Ramifications **10** (2001), no. 4, 597–607.

[18] Yuanan Diao, Nicholas Pippenger, and De Witt Sumners, *On random knots*, J. Knot Theory Ramifications **3** (1994), no. 3, 419–429. Random knotting and linking (Vancouver, BC, 1993).

[19] N. Dunfield, A. Hirani, M. Obeidin, A. Ehrenberg, S. Bhattacharyya, and D. Lei, *Random Knots: A preliminary report, 2014.* (2014). `http://www.math.uiuc.edu/~nmd/preprints/slides/random_knots.pdf`.

[20] Herbert Edelsbrunner, Leonidas Guibas, John Hershberger, János Pach, Richard Pollack, Raimund Seidel, Micha Sharir, and Jack Snoeyink, *On arrangements of Jordan arcs with three intersections per pair*, Discrete Comput. Geom. **4** (1989), no. 5, 523–539.

[21] Herbert Edelsbrunner, Leonidas Guibas, János Pach, Richard Pollack, Raimund Seidel, and Micha Sharir, *Arrangements of curves in the plane—topology, combinatorics, and algorithms*, Theoret. Comput. Sci. **92** (1992), no. 2, 319–336.

[22] Mohamed Elhamdadi and Sam Nelson, *Quandles—an introduction to the algebra of knots*, Student Mathematical Library, vol. 74, American Mathematical Society, Providence, RI, 2015.

[23] Chaim Even-Zohar, Joel Hass, Nati Linial, and Tahl Nowik, *Invariants of random knots and links*, Discrete Comput. Geom. **56** (2016), no. 2, 274–314.

[24] H.L. Frisch and E. Wasserman, *Chemical topology 1*, J. Am. Chem. Soc. **83(18)** (1961), 3789–3795.

[25] P. J. Giblin, *Graphs, surfaces and homology*, Chapman and Hall, London; Halsted Press [John Wiley & Sons], New York, 1977. An introduction to algebraic topology; Chapman and Hall Mathematics Series.

[26] Branko Grünbaum, *Arrangements and spreads*, American Mathematical Society Providence, R.I., 1972. Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, No. 10.

[27] Douglas Jungreis, *Gaussian random polygons are globally knotted*, J. Knot Theory Ramifications **3** (1994), no. 4, 455–464.

[28] Ross J. Kang and Tobias Müller, *Arrangements of pseudocircles and circles*, Discrete Comput. Geom. **51** (2014), no. 4, 896–925.

[29] Louis H. Kauffman, *State models and the Jones polynomial*, Topology **26** (1987), no. 3, 395–407.

[30] ———, *Virtual knot theory*, European J. Combin. **20** (1999), no. 7, 663–690.

[31] ———, *Knot diagrammatics*, Handbook of knot theory, Elsevier B. V., Amsterdam, 2005, pp. 233–318.

[32] Louis H. Kauffman and Sofia Lambropoulou, *Hard unknots and collapsing tangles*, Introductory lectures on knot theory, Ser. Knots Everything, vol. 46, World Sci. Publ., Hackensack, NJ, 2012, pp. 187–247.

[33] Jan Kynčl (https://mathoverflow.net/users/24076/jan-kyncl), *Drawings of complete graphs with $Z(n)$ crossings*, MathOverflow. URL: https://mathoverflow.net/q/128878 (version: 2013-05-06).

[34] Johann Linhart and Ronald Ortner, *On the combinatorial structure of arrangements of oriented pseudocircles*, Electron. J. Combin. **11** (2004), no. 1, Research Paper 30, 13.

[35] ———, *An arrangement of pseudocircles not realizable with circles*, Beiträge Algebra Geom. **46** (2005), no. 2, 351–356.

[36] ———, *A note on convex realizability of arrangements of pseudocircles*, Geombinatorics **18** (2008), no. 2, 66–71.

[37] Johann Linhart and Yanling Yang, *Arrangements of arcs and pseudocircles*, Beiträge Algebra Geom. **37** (1996), no. 2, 391–398.

[38] László Lovász, Katalin Vesztergombi, Uli Wagner, and Emo Welzl, *Convex quadrilaterals and k-sets*, Towards a theory of geometric graphs, Contemp. Math., vol. 342, Amer. Math. Soc., Providence, RI, 2004, pp. 139–148.

[39] Bojan Mohar and Carsten Thomassen, *Graphs on surfaces*, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, 2001.

[40] Kunio Murasugi, *Jones polynomials and classical conjectures in knot theory*, Topology **26** (1987), no. 2, 187–194.

[41] Ronald Ortner, *Embeddability of arrangements of pseudocircles into the sphere*, European J. Combin. **29** (2008), no. 2, 457–469.

[42] ———, *Forcing subarrangements in complete arrangements of pseudocircles*, J. Comput. Geom. **6** (2015), no. 1, 235–248.

[43] Nicholas Pippenger, *Knots in random walks*, Discrete Appl. Math. **25** (1989), no. 3, 273–278.

[44] Kurt Reidemeister, *Elementare Begründung der Knotentheorie*, Abh. Math. Sem. Univ. Hamburg **5** (1927), no. 1, 24–32.

[45] D. W. Sumners and S. G. Whittington, *Knots in self-avoiding walks*, J. Phys. A **21** (1988), no. 7, 1689–1694.

[46] Morwen B. Thistlethwaite, *A spanning tree expansion of the Jones polynomial*, Topology **26** (1987), no. 3, 297–309.

[47] Carsten Thomassen, *Embeddings of graphs with no short noncontractible cycles*, J. Comb. Theory Ser. B **48** (1990), no. 2, 155–177.