A Brief Introduction to Hyperbolic Knots

This book gives an introduction to knots, links, and hyperbolic geometry. Before we begin, we need to carefully define what we mean by knots and links, and that is done in this chapter. We also introduce classical problems in knot theory, and problems motivated by geometry, especially hyperbolic geometry. This chapter is meant to motivate future chapters, and it has many references to content covered in more detail later in the book, where we address some of these problems. Many of the questions described in this chapter have partial answers, and many are still wide open.

0.1. An introduction to knot theory

The earliest study of knots seems to be by Gauss, Listing, and especially Tait, who published several papers on knot theory in the years 1876 through 1885. In a preface to his work on knot theory, republished in his 1898 Scientific papers [**Tai98**], Tait writes:

> "The subject [knot theory] is a very much more difficult and intricate one than at first sight one is inclined to think, and I feel that I have not succeeded in catching the key-note."

Since Tait's work, advances in knot theory have come through applications of topology, algebra, and invariants arising in quantum field theory, but no single mathematical field has led to simple tools that apply to all knots. In other words, perhaps mathematicians still have not succeeded in catching the "key-note." Perhaps there is no "key-note" in knot theory.

However, there are definitely mathematical techniques that work well when applied to particular problems or particular families. This book introduces techniques arising from geometry.

0.1.1. Basic terminology. To begin, we need careful definitions of the objects involved.

Definition 0.1. A *knot* $K \subset S^3$ is a subset of points homeomorphic to a circle S^1 under a piecewise linear (PL) homeomorphism. We may also think of a knot as a PL embedding $K\colon\ S^1 \to S^3$. We will use the same symbol K to refer to the map and its image $K(S^1)$.

More generally, a *link* is a subset of S^3 PL homeomorphic to a disjoint union of copies of S^1. Alternatively, we may think of a link as a PL embedding of a disjoint union of copies of S^1 into S^3.

A PL homeomorphism of S^1 is one that takes S^1 to a finite number of linear segments. Restricting to such homeomorphisms allows us to assume that a knot $K \subset S^3$ has a regular tubular neighborhood, that is, there is an embedding of a solid torus $S^1 \times D^2$ into S^3 such that $S^1 \times \{0\}$ maps to K. An embedding of S^1 into S^3 that cannot be made piecewise linear defines an object called a *wild knot*. Wild knots may have very interesting geometry, but we will only be concerned with the classical knots of Definition 0.1 here.

In fact, rather than working with PL embeddings and homeomorphisms, we obtain the same results working with smooth ones. That is, we could require instead in Definition 0.1 that a knot be a smooth embedding of S^1 into S^3, and we obtain an equivalent theory. We will assume this fact throughout the book, working with both PL and smooth maps, with very little mention of this fact.

Definition 0.2. We will say that two knots (or links) K_1 and K_2 are equivalent if they are *ambient isotopic*, that is, if there is a (PL or smooth) homotopy $h\colon\ S^3 \times [0,1] \to S^3$ such that $h(*,t) = h_t\colon\ S^3 \to S^3$ is a homeomorphism for each t, and

$$h(K_1,0) = h_0(K_1) = K_1 \quad \text{and} \quad h(K_1,1) = h_1(K_1) = K_2.$$

Such a map h is called an *ambient isotopy*.

A PL (or smooth) embedding of S^1 into S^3 defines two 3-manifolds, one open and one compact, as in the following definition.

Definition 0.3. For a knot K, let $N(K)$ denote an open regular neighborhood of K in S^3. The *knot exterior* is the manifold $S^3 - N(K)$. Notice that it is a compact 3-manifold with boundary homeomorphic to a torus.

The *knot complement* is the open manifold $S^3 - K$, homeomorphic to the interior of $S^3 - N(K)$.

Similarly, if L is a link the *link exterior* is $S^3 - N(L)$, and the *link complement* is $S^3 - L$.

It was an open question for many years as to whether two knots with homeomorphic complements must be equivalent (up to reflection). This was proved in the affirmative by Gordon and Luecke in 1989 [**GL89**].

Theorem 0.4 (Gordon–Luecke Theorem)**.** *If two knots have complements that are homeomorphic by an orientation-preserving homeomorphism, then the knots are equivalent.*

The complement of a knot and the complement of its reflection are homeomorphic, by the orientation-reversing reflection homeomorphism. However, the knot itself may not be equivalent to its reflection. In fact, hyperbolic geometry tools do not distinguish knots and their reflections, and so we often only consider knots up to reflection in this book. If we disregard reflections, the Gordon–Luecke theorem states that knots are determined by their complements.

The same is not true for links. There are infinitely many inequivalent links whose complements are homeomorphic. However, the ways in which such links can be constructed are relatively well-understood; see, for example [**Gor02**].

Definition 0.5. A *knot diagram* (or *link diagram*) is a 4-valent graph with over/under crossing information at each vertex. The diagram is embedded in a plane $S^2 \subset S^3$ called the *projection plane*, or *plane of projection*.

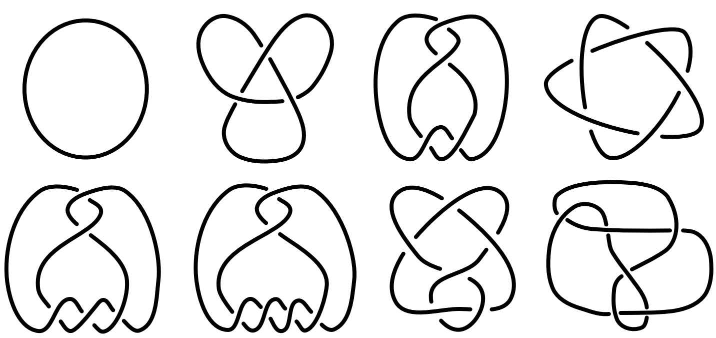


Figure 0.1. Knots with at most six crossings.

Figure 0.1 shows diagrams of the eight knots with at most six crossings. Classically, a knot has been described by a diagram. Tait's works give many diagrams. In modern work, knots also appear without diagrams, for example when they arise as periodic orbits of a dynamical system [**BW83**], or from a gluing of polyhedra [**CDW99**, **CKP04**, **CKM14**].

However, many open problems in knot theory still concern knot diagrams. One goal of Chapter 1, and then the next few chapters, is to give a method to pass from a knot or link *diagram* to a topological and then geometric description of the knot or link *complement*. That is, we start with a 4-valent graph describing a knot or link K, and obtain a mathematically rigorous decomposition of the 3-manifold $S^3 - K$ into simple 3-dimensional pieces, which will be useful for applying tools from geometry and 3-manifold topology.

0.2. Problems in knot theory

There are many open problems in knot theory, and as new mathematical fields are brought to bear upon these problems, new questions and problems arise. This section gives a few highlights of the most classical problems, and also problems that seem most amenable to geometric techniques. Probably the most long-standing problem, and also one of the most broad, is the following.

0.2.1. The classification problem. When do two different descriptions of knots yield equivalent knots? When do they have homeomorphic complements?

When the description of a knot is given by a diagram, this is the problem that Tait encountered while trying to list all knots with a fixed number of crossings. See Figure 0.2, which is modified from the 1884 paper [**Tai 4**].

There are a few moves that can be performed on a diagram that do not change the equivalence class of the underlying knot. For example, if the diagram contains a single crossing that forms a loop, as shown on the left of Figure 0.3, that loop can be untwisted to simplify the diagram.

Definition 0.6. A single crossing forming a loop, as on the left of Figure 0.3, is called a *nugatory crossing*.

More generally, a *reducible crossing* is a crossing through which we may draw a circle γ on the plane of projection such that γ meets the diagram only at one point, at the crossing. See Figure 0.3, right.

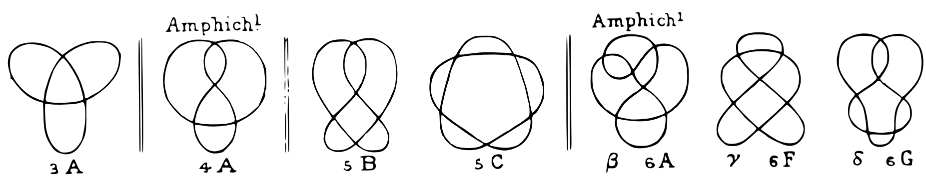


Figure 0.2. A very small portion of P. Tait's 1884 tables of knot diagrams, from [**Tai 4**]. The original contains a full page with such diagrams, with additional pages of diagrams in [**Tai 5**].

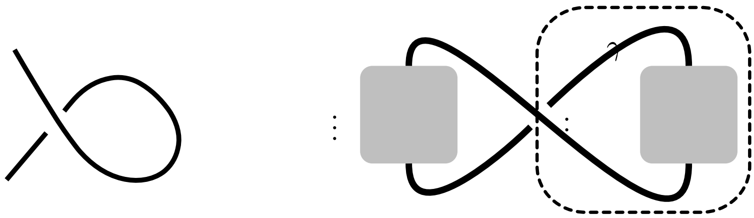


Figure 0.3. On the left, a nugatory crossing. On the right, a more general reducible crossing.

A diagram is *reduced* if it contains no reducible crossings.

Note that reducible crossings can be removed by an ambient isotopy of the diagram. We typically will assume that our knot diagrams are reduced.

There are other well-known moves to change a diagram into an equivalent diagram. These include the three moves shown in Figure 0.4, called Reidemeister moves.

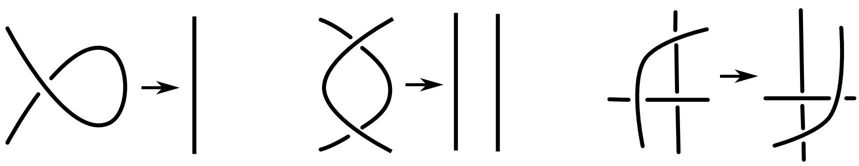


Figure 0.4. Three Reidemeister moves do not change knot equivalence.

The Reidemeister moves appear in work of Maxwell in the 1800s (see, for example, [**Epp98**]). In the 1920s, Reidemeister [**Rei27**] and Alexander and Briggs [**AB27**] independently gave rigorous proofs that two equivalent diagrams can always be related by a sequence of such moves.

The *crossing number* of a knot is the minimal number of crossings in all diagrams of the knot. A minimal crossing diagram will necessarily be reduced. However, a reduced diagram is not necessarily a minimal crossing diagram. For example, Figure 0.5 shows the reduced diagram of a knot that can, with a little work, be simplified to the *unknot*, i.e. the simple circle with no crossings. This diagram was discovered by Goeritz in 1934 [**Goe34**].

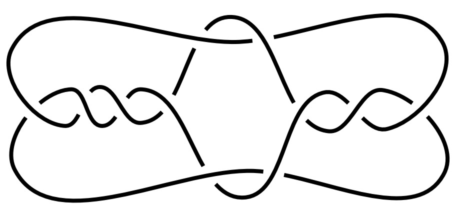


Figure 0.5. This diagram of the unknot was discovered in 1934 by Goeritz.

In fact, the diagram of Figure 0.5 is an example of a knot diagram that cannot be simplified by Reidemeister moves without first increasing the number of crossings of the diagram.

In addition to attempting to remove crossings, other moves can be performed on diagrams to simplify the classification problem. For example, there is a way of joining two simple diagrams into one more complicated diagram, shown in Figure 0.6.

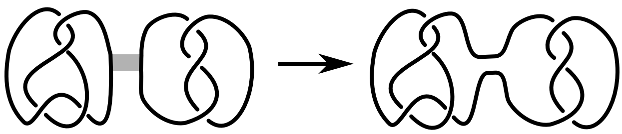


Figure 0.6. The knot sum of two knots.

Starting with two diagrams side-by-side, take a rectangle embedded in the plane of projection that has one side on one diagram, avoiding crossings, an opposite side on the other diagram, again avoiding crossings, and the final two sides disjoint from the two diagrams. Form the new diagram by removing the two edges of the rectangle that lie on the knots, and joining the knots along the two opposite sides of the rectangle. The resulting knot is called the *knot sum*. It is also sometimes called the *connected sum of the knots*.

Given a knot sum of two knot diagrams, consider the embedded curve γ in the plane of projection of the diagram that encircles exactly one of the original diagrams, cutting through the rectangle in the definition of the knot sum. This curve γ meets the diagram of the knot sum in exactly two points, it bounds disks on both sides (thinking of the projection plane as $S^2 \subset S^3$), and both discs contain crossings. We say that a diagram is *prime* if no such curve γ exists. That is, a knot or link diagram is *prime* if, for every simple closed curve γ in the plane of projection, if γ meets the knot exactly twice transversely away from crossings, then γ bounds a region of the diagram with no crossings.

Curves such as γ above detect knot sums. When knots are classified by a diagram, listed according to crossing number, typically only prime diagrams are included.

The problem of listing all knots by crossing number, without duplicates, is a difficult one. There are 1,701,936 prime knots with at most 16 crossings, classified by Hoste, Thistlethwaite, and Weeks in 1998 [**HTW98**]. More recently, Burton classified 352,152,252 prime knots up to 19 crossings [**Bur20**]. These knots can be downloaded with the 3-manifold software

Regina [**BBP$^+$19**]. In both instances, the knots are only classified up to reflection in the plane of projection.

Definition 0.7. A *knot invariant* is a function from the set of knots to some other set whose value depends only on the equivalence class of the knot. A *link invariant* is defined similarly.

The crossing number of a knot is an example of a knot invariant.

Knot and link invariants are used to prove that two knots or links are distinct, or to measure the complexity of the link in various ways. We will revisit examples of knot invariants below, particularly geometric ones.

Notice that the number of knots with a given crossing number grows very rapidly. There does not seem to be a natural way of enumerating knots within a fixed class of a crossing number. And while the crossing number was one of the first knot invariants to be studied by knot theorists, it does not seem to relate well to other knot invariants, particularly those that arise in geometry. For these reasons and others, other ways of classifying knots have arisen over the years, which we will discuss further below.

In this book we will apply geometry to the problem of the classification of knots. It has been known since the early 1980s, due to work of Thurston [**Thu82**], that the complement of a knot decomposes into pieces, each admitting a 3-dimensional geometry. By using geometric properties of knot complements, we can often distinguish knots. This brings us to the second problem in knot theory that we discuss here.

0.2.2. The problem of determining geometry of the complement. Briefly, the complement of a knot is *hyperbolic* if and only if it admits a complete metric with all sectional curvatures equal to -1. We will give other equivalent definitions of hyperbolic knots in later chapters, which will often be more useful for calculations, computations, and examples.

For now, it is known that when a knot complement is hyperbolic, its hyperbolic metric is unique. That is, hyperbolic knot complements that are homeomorphic must also be isometric under any hyperbolic metrics placed upon their complements. Moreover, a large number of knots are hyperbolic, and many that are not hyperbolic decompose into hyperbolic pieces.

More precisely, consider the following families of knots.

Definition 0.8. A *torus knot* is a knot that can be embedded on the surface of an unknotted torus in S^3 (without crossings on the surface of the torus). See Figure 0.7.

By an unknotted torus, we mean the boundary of a regular neighborhood of an unknot in S^3, with no crossings.

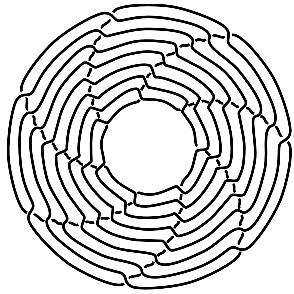


Figure 0.7. A torus knot.

Definition 0.9. A *satellite knot* is a knot that can be embedded (nontrivially) in a regular neighborhood of a different, nontrivial knot in S^3. See Figure 0.8.

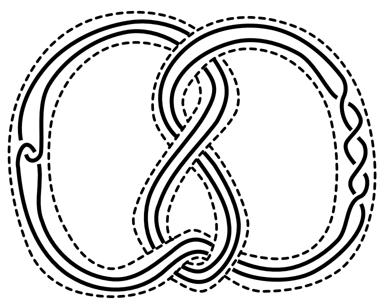


Figure 0.8. An example of a satellite knot. The dotted line forms the boundary of a neighborhood of a different knot, and the satellite lives inside that neighborhood.

The complement of a torus knot admits a 3-dimensional geometry that is not hyperbolic, due to work of Thurston [**Thu82**]. He also showed that the complement of a satellite knot cannot be hyperbolic, but can be cut along a torus to decompose into pieces that admit 3-dimensional geometry, which could possibly be hyperbolic. For example, the knot complement in Figure 0.8 can be cut along the dashed solid torus into two hyperbolic pieces, as we will see later in this book.

Thurston showed that every knot in S^3 that is neither a torus knot nor a satellite knot must have a hyperbolic complement [**Thu82**].

Thus hyperbolic geometry can be a useful tool in the classification problem of knots — in theory.

In practice, we need tools and techniques to determine when a knot complement is hyperbolic. For example, if a knot is given by a messy diagram, how does one determine whether or not it is equivalent to a torus or satellite knot? How can we determine whether its complement is hyperbolic? And if it is hyperbolic, how can we find a hyperbolic metric?

Thurston outlined a procedure for finding a hyperbolic metric using the diagram of the figure-8 knot in his 1979 lecture notes [**Thu79**]. This process was generalized by others, for example [**Men83**], and even made algorithmic, in Weeks' 1985 PhD thesis [**Wee85**]. There is now software that determines, given a knot diagram, whether or not the knot complement is hyperbolic. This is the computer program SnapPy, which is freely available [**CDGW16**].

Indeed, using computational tools, Burton has determined that of all prime knots with up to 19 crossings, $352,151,858$ are hyperbolic, and only 395 are not hyperbolic [**Bur20**]. These are split into 14 torus knots and 380 satellite knots.

The next four chapters of this book concern the problem of determining a hyperbolic metric on a knot complement. We will step carefully through the necessary definitions and procedures, using Thurston's decomposition of the figure-8 knot complement as an example. This will give our first potential method to find a hyperbolic metric.

Chapters 5 and 6 give additional methods and tools from hyperbolic geometry to find or deform a hyperbolic metric. These first six chapters form the foundation required to discuss hyperbolic geometry and knots in more detail.

Of course, these chapters require some work. The fact that software exists that can compute hyperbolic geometry of knots begs the question, why work through such computations by hand at all? Why not just work with the computer? There are many reasons, related to additional open problems. One reason is the next problem.

0.2.3. The problem of determining geometry for families of knots. A computer program computes hyperbolic geometry for one knot at a time, or for a finite number of knots. But what can be said about infinite families of knots? For example, how does one determine the hyperbolic geometry of knots with descriptions given by infinite classes of diagrams? If two knots in a family are "similar", is their geometry also similar?

Potential answers to such questions seem to depend very heavily on the family of knots given. For example, for fixed c, it does not seem to be the case that the (finite) family of knots with crossing number c have very similar hyperbolic geometry.

On the other hand, certain infinite families of knots do exist with very similar hyperbolic geometry, and others at least seem to have geometry that reflects properties of the diagrams. We will discuss such knots and their properties, for example in chapters 7, 10, and 11, with careful proofs. For now, we will present a definition of one such family.

Definition 0.10. A *bigon* is a region of a graph bounded by exactly two edges and exactly two vertices.

For example, Figure 0.9 shows several bigons connected end-to-end in a portion of a diagram graph of a knot.

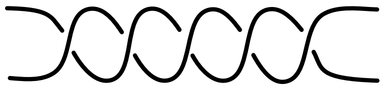


Figure 0.9. A twist region of a diagram.

Definition 0.11. A *twist region* of a diagram of a knot is a maximal portion of the knot diagram where two strands twist around each other, as in Figure 0.9.

More precisely, recall that a diagram of a knot is a 4-valent graph with over/under crossing information at each vertex. A twist region is a string of bigon regions in the diagram graph, arranged end-to-end at their vertices, which is maximal in the sense that there are no additional bigon regions meeting the vertices on either end. A single crossing adjacent to no bigons is also a twist region. We will further restrict so that all twist regions are alternating, meaning crossings alternate over and under while following a strand of the twist region. If not, the second Reidemeister move applied to the diagram removes two crossings from the twist region.

The condition that twist regions be maximal ensures that there is only one way to put together exactly two twist regions in a diagram.

Definition 0.12. The *twist knot* $J(2, n)$ is the knot with a diagram consisting of exactly two twist regions, one of which contains two crossings. The other twist region contains $n \in \mathbb{Z}$ crossings. The direction of crossing depends on the sign of n.

Twist knots $J(2, 2)$, $J(2, 3)$, $J(2, 4)$, and $J(2, 5)$ are shown in Figure 0.10.

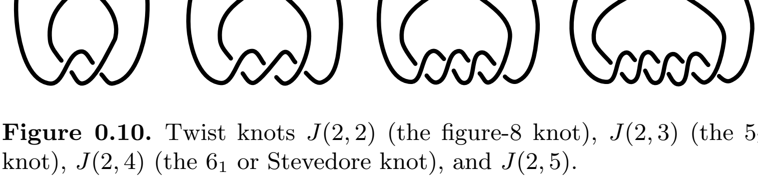


Figure 0.10. Twist knots $J(2, 2)$ (the figure-8 knot), $J(2, 3)$ (the 5_2 knot), $J(2, 4)$ (the 6_1 or Stevedore knot), and $J(2, 5)$.

The family of twist knots $J(2, n)$ has very nice hyperbolic geometry, which we discuss in Chapter 7. In particular, as n approaches infinity, we will see that the hyperbolic geometry of twist knot complements limits, in a precise sense, to the hyperbolic geometry of the Whitehead link complement; the Whitehead link is shown in Figure 0.11.

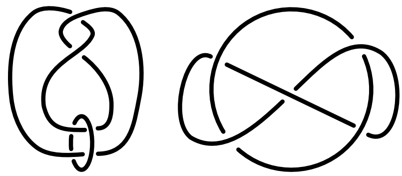


Figure 0.11. Two diagrams of the Whitehead link.

More generally, any family of knots containing higher and higher numbers of crossings in a twist region will have complements converging to a link with a simple circle encircling that twist region. Knots with high numbers of crossings in twist regions are called *highly twisted*. Again these are discussed in Chapter 7.

Given a diagram of a link, we can combine twist regions by performing a sequence of moves on the diagram called *flypes*.

Definition 0.13. Let γ be a simple closed curve meeting the diagram of K transversely exactly four times away from crossings, with two intersections adjacent to a crossing on the outside of γ. A *flype* is a move on the diagram that rotates the region inside γ by 180°, moving the crossing adjacent to γ to become a crossing adjacent to γ but between the opposite two strands. See Figure 0.12.

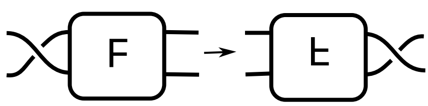


Figure 0.12. A flype.

Now, suppose a simple closed curve γ in the plane of projection meets a diagram transversely exactly four times away from crossings, and suppose also that the curve is adjacent to crossings on both sides. Then we can perform a flype to move one of the crossings to the opposite side of the curve, to form a bigon. If the bigon is not alternating, remove both crossings, producing a diagram with fewer crossings. Otherwise, there are two cases. Either the curve γ encloses only bigons on one side to begin with, and the flype produces a diagram that is unchanged, or the flype has moved a crossing out of one twist region, on one side of γ, into a distinct twist region on the other side of γ. Performing the same flype a finite number of times will move all crossings in the twist region on one side of γ into the twist region on the other side, thus reducing the number of twist regions of the diagram. Thus by performing a finite number of flypes, we obtain a diagram with a minimal number of twist regions. Such a diagram is called *twist-reduced*.

Every knot has a twist-reduced diagram with some number of twist regions. On the other hand, for a fixed positive integer T, there are only finitely many ways of combining twist regions to form a twist-reduced diagram with T twist regions. The collection of twist-reduced diagrams with T twist regions forms an infinite family of diagrams. Two highly twisted diagrams with the same pattern of twist regions will have similar hyperbolic geometry, in ways that can be quantified. Thus rather than classifying knots by crossing number, from a geometric perspective it may make more sense to classify knots by number of twist regions in a twist-reduced diagram, or *twist-number*. This brings us to another (broad and vaguely worded) problem.

0.2.4. The problem of enumerating knots by geometry. Enumerating knots by twist region may make more geometric sense than enumerating by crossing number, because highly twisted knots have diagrams that relate well to their geometry, in a sense that will be made precise in Chapter 7. Given any knot, is there always a diagram that encodes hyperbolic geometry?

Schubert considered a family of knots in 1956 [**Sch56**]. He called the knots *2-bridge knots*. They can be described diagrammatically by taking four parallel strands, and twisting pairs of the strands into sequences of twist regions, then capping off either end with two "bridges." A general form of such a diagram is shown in Figure 0.13; see also Chapter 10.

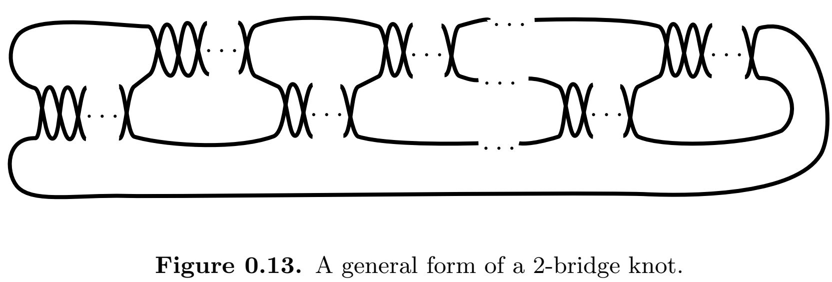


Figure 0.13. A general form of a 2-bridge knot.

Although Schubert's work pre-dates the first work on the hyperbolic geometry of knots by nearly two decades, his 2-bridge knots turn out to be very amenable to hyperbolic geometry techniques. We will see early on in this book that any knot exterior $S^3 - N(K)$ can be decomposed into a collection of truncated tetrahedra. Equivalently, $S^3 - K$ is formed by gluing tetrahedra whose vertices have been removed. This is called an *ideal triangulation* of the knot exterior, or sometimes simply a *triangulation*.

In the case of 2-bridge knots, we will see that a triangulation of the knot complement can be read easily off the diagram. Not only that, we will see in Chapter 10 that the edges and faces of the triangulation can be made totally geodesic under the hyperbolic metric, and the tetrahedra can be straightened simultaneously to be convex, with piecewise geodesic boundaries. Thus the combinatorics of the diagram of a 2-bridge knot gives a combinatorial method of describing the geometry of the 2-bridge knot. This is very powerful.

It would be great to be able to extend these techniques to all knots, and some progress has been made with applications to other families, such as n-bridge knots for higher n. However, few families seem to be quite as nice as 2-bridge knots.

There is still much ongoing work on triangulating knot exteriors and determining geometric properties of triangulations. We will discuss some of the techniques and applications in Chapter 9.

We have mentioned above that any knot exterior can be triangulated. In fact, any 3-manifold with torus boundary components can be decomposed into truncated tetrahedra. When the tetrahedra are convex hyperbolic tetrahedra, we say the triangulation is *geometric*. The software SnapPy has a census of orientable manifolds built up of at most nine geometric tetrahedra [**CDGW16**]. Some of these are knot complements.

This leads to a new way of classifying hyperbolic knots: by the number of geometric tetrahedra required to triangulate their exterior. This method of enumerating knots has been employed in [**CDW99**, **CKP04**, **CKM14**].

To date, 502 hyperbolic knots, built of at most eight geometric tetrahedra, have been classified. The diagrams of these knots often have large numbers of crossings. The knots built of at most four tetrahedra are shown in Figure 0.14.

Classifying knots by triangulations of their exteriors seems to be more difficult than classifying them by diagrams. This is because, given a triangulation of a 3-manifold with torus boundary, it is not obvious that the underlying space is a knot complement for a knot K in S^3. We will discuss some techniques to detect whether such a manifold is a knot complement in Chapter 8.

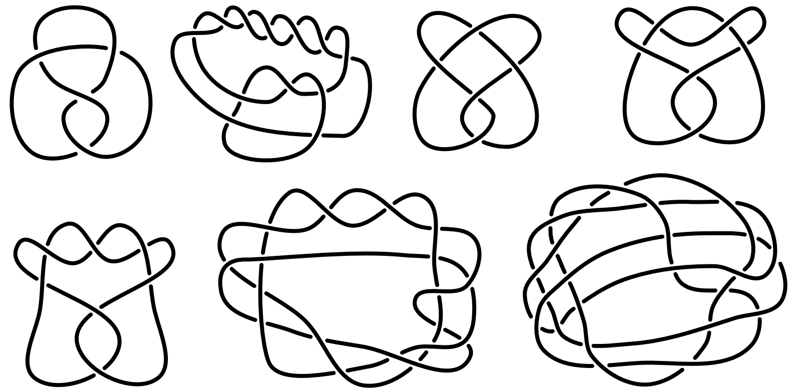


Figure 0.14. The seven simplest hyperbolic knots, built of at most four geometric tetrahedra.

0.2.5. The problem of finding geometric diagrams. Twist knots and 2-bridge knots have standard diagrams that encode a great deal of information about the geometry of the knot. Does every knot have such a diagram? (Probably not.) Does every knot have a diagram from which we may read some geometric information?

Alternating knots are another family of knots that seem to be amenable to hyperbolic geometric techniques.

Definition 0.14. An *alternating diagram* is a diagram of a knot or link that has an orientation such that, when following the knot in the direction of the orientation, the crossings alternate between over and under. An *alternating knot or link* is a knot or link that has an alternating diagram.

We will see in the exercises in Chapter 1 that alternating knot complements decompose into pieces with the same combinatorics of the diagram. In chapters 11, 12, and 13 we will use this decomposition to determine some geometric information on the knot complement.

How useful is this work broadly? All knots with at most seven crossings have alternating diagrams. Tait began his work [**Tai98**] by assuming diagrams were alternating (although he did publish diagrams of eight- and ten-crossing nonalternating examples in 1877). However, the proportion of alternating knots in diagrams enumerated by crossing number rapidly drops to zero [**ST98**, **Thi98**]. As for knots enumerated by geometric triangulations, nonalternating examples seem to be even more common; a nonalternating example appears as the second knot on the list in Figure 0.14. Thus, unfortunately, alternating knots and links are not very common.

An open research question is, how many of the techniques presented in these chapters for determining geometry of alternating links generalize to other knots and links? There has been much work in recent years in extending this work to other families of knots, and some success. We are far from using such techniques to find hyperbolic geometry of all knots, though.

0.2.6. The problem of determining geometric invariants. One way of distinguishing knots is to compute invariants for each of them. If the invariants disagree, then the knots cannot be equivalent.

Several knot invariants arise classically, such as the crossing number that we encountered above. Many additional knot invariants arise through geometry. One aim of this book is to discuss such invariants, and give tools to calculate them.

One of the most straightforward knot invariants that arises in geometry is the volume of a knot. We will show in Chapter 5 that any knot complement that admits a hyperbolic structure has finite volume. Thus volumes of knots give knot invariants.

For those knots whose diagrams are particularly amenable to geometric techniques, such as twist knots, 2-bridge knots, and alternating knots, there are known methods to estimate volume using the combinatorics of the diagram. This is discussed along the way, but especially in Chapter 13, where we bring to bear several tools in geometry to give two-sided bounds on volumes.

How powerful is volume as a knot invariant? It can be easy to calculate numerically, using the software SnapPy [**CDGW16**], for example. Such computations can be rigorously verified to lie in a fixed error range using interval arithmetic, as in [**HIK$^+$16**]. Thus computing volume is a useful tool for distinguishing knots with distinct volume. However, there are many distinct knots that cannot be distinguished by volume; they have the same volume. We give some methods of constructing such knots and links in Chapter 12.

Then, is there a better geometric knot invariant than volume to distinguish knots? In Chapter 14, we describe the *canonical decomposition* of a hyperbolic knot complement. This is a decomposition consisting of convex polyhedra. We will show that when two knots have the same canonical decomposition, they must necessarily have homeomorphic complements, and thus by the Gordon–Luecke theorem, they must be equivalent (up to reflection). Thus the canonical decomposition is a complete invariant for hyperbolic knots. Unfortunately, it is not easy to compute in general, and provable forms of canonical decompositions are only known for a few infinite families of knots, including 2-bridge knots [**Gue06a**]. Canonical decompositions of alternating knots are still unknown in general, for example.

Finally, we discuss very briefly one polynomial invariant. In most standard books on knot theory, there will be chapters on polynomial invariants, particularly the Alexander polynomial and the Jones polynomial. We will not treat such polynomials here; they arise from techniques that do not use hyperbolic geometry. There is one polynomial invariant of knots that depends heavily on hyperbolic geometry, however. This is the A-polynomial. We devote Chapter 15 to a discussion of the A-polynomial, its definition, and computation for a few examples. We will see that it relates to hyperbolic structures on a knot complement and the deformations of such structures.

0.2.7. The problem of relating geometric invariants to other invariants. What of the invariants that are being omitted from this book? We mentioned above Alexander and Jones polynomials. There are also more modern algebraic knot invariants, such as Khovanov homology and Floer homologies, and quantum invariants such as colored Jones polynomials.

Many open problems in knot theory, driving much of the ongoing research in the field, concern relating invariants of knots arising from other fields of mathematics to hyperbolic geometry and hyperbolic knot invariants. We will not discuss in detail these open problems, because defining nonhyperbolic invariants will take us too far afield. However, one motivating factor for writing this book was to help mathematicians, particularly students, get up to speed with their hyperbolic geometry, in order to investigate the relations of geometry to other invariants in knot theory.

0.3. Exercises

Exercise 0.15. Find a sequence of isotopy moves of the diagram of the Goeritz knot, Figure 0.5, that reduces it to the standard diagram of the unknot with no crossings.

Exercise 0.16. Download and install the software SnapPy [**CDGW16**]. Use it to sketch diagrams of a few knots, and determine whether the knot is hyperbolic. Do this for at least one hyperbolic knot and at least one nonhyperbolic knot.

Exercise 0.17. Convince yourself by drawing several examples that every 4-valent planar graph can be assigned over/under crossing information at each vertex to obtain an alternating knot. Now try to prove this fact. (This may require some graph theory.)

Exercise 0.18. Show that a connected sum of two knots is always a satellite knot.

Part 1

Foundations of Hyperbolic Structures

Chapter 1

Decomposition of the Figure-8 Knot

In this chapter, we begin developing tools to work with knots and links and the 3-manifolds they define. We give a geometric method, explained carefully by example, to decompose a knot or link complement into simple pieces. The methods here are an introduction to topological techniques in 3-manifold geometry and topology, and an introduction to some of the tools used in the field.

One goal of this chapter is to present a method that will allow us to pass from a knot or link *diagram* to a description of the knot or link *complement*. That is, we start with a 4-valent graph describing a knot or link K, and we obtain a mathematically rigorous decomposition of the 3-manifold $S^3 - K$ into simple 3-dimensional pieces, of which the pieces will be useful for applying tools from geometry and 3-manifold topology.

1.1. Polyhedra

Sometimes it is easier to study manifolds, including knot complements, if we split them into smaller, simpler pieces, for example 3-balls. We are going to decompose the figure-8 knot complement into two carefully marked 3-balls, namely ideal polyhedra. The diagram of the figure-8 knot that we use is shown in Figure 1.1. The decomposition we describe appears in Thurston's notes [**Thu79**], and with a little more explanation in [**Thu97**]. The procedure has been generalized to all link complements, for example in [**Men83**]. This work is essentially what we present below in the text and in exercises.

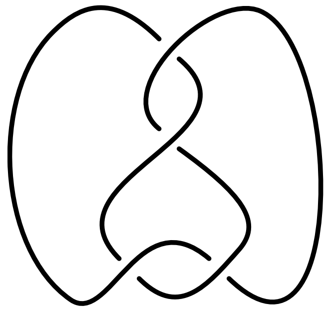


Figure 1.1. A diagram of the figure-8 knot.

Definition 1.1. A *polyhedron* is a closed 3-ball whose boundary is labeled with a finite graph, containing a finite number vertices and edges, so that complementary regions, which are called *faces*, are simply connected.

An *ideal polyhedron* is a polyhedron with all vertices removed. That is, to form an ideal polyhedron, start with a regular polyhedron and remove the points corresponding to vertices.

We will cut $S^3 - K$ into two ideal polyhedra. We will then have a description of $S^3 - K$ as a gluing of two ideal polyhedra. That is, given a description of the polyhedra, and gluing information on the faces of the polyhedra, we may reconstruct the knot complement $S^3 - K$. Although we use the example of the figure-8 knot, in the exercises, you will walk through the techniques below to determine decompositions of other knot complements into ideal polyhedra, and to generalize to all knots.

1.1.1. Overview. Start with a diagram of the knot. There will be two polyhedra in our decomposition. These can be visualized as two balloons: One balloon expands above the diagram, and one balloon expands below the diagram. As the balloons continue expanding, they will bump into each other in the regions cut out by the graph of the diagram. Label these regions. In Figure 1.2, the regions are labeled A, B, C, D, E, and F. These will correspond to faces of the polyhedra.

The faces meet up in edges. There is one edge for each crossing. It runs vertically from the knot at the top of the crossing to the knot at the bottom (or the other way around). The balloon expands until faces meet at edges. Figure 1.3 shows how the top balloon would expand at a crossing. The edge is drawn as an arrow from the top of the crossing to the bottom. Faces labeled T and U meet across the edge. Rotating the picture 180° about the edge, we would see an identical picture with S meeting V.

It may be helpful to examine the meeting of faces at an edge by a 3-dimensional model. Henry Segerman has come up with a paper model to

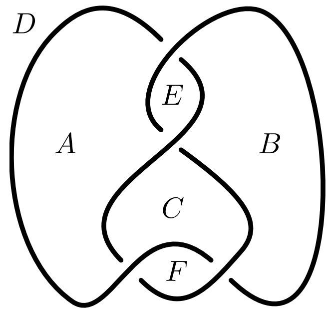


Figure 1.2. Faces for the figure-8 knot complement.

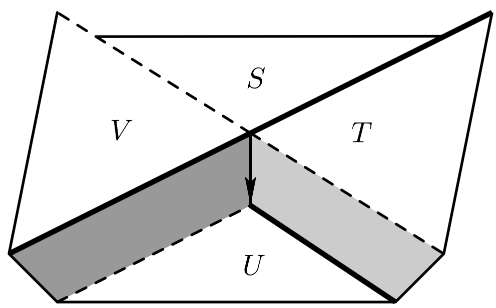


Figure 1.3. The knot runs along diagonals. Faces labeled U and T meet at the edge shown, marked by an arrow.

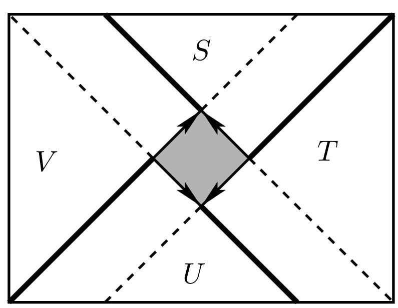


Figure 1.4. Cut out the shaded square. Start with a pair of parallel lines. Fold the thick part of the line in a direction opposite that of the dashed part of the line. Fold parallel thick and dashed lines in opposite directions. Correct folding results in a model that looks like Figure 1.3.

illustrate the phenomenon of Figure 1.3. Start with a sheet of paper labeled as in Figure 1.4. Cut out the shaded square in the middle. Now fold the paper until it looks like that in Figure 1.3. By rotating the paper model, we can see how the faces meet up.

Stringing crossings such as this one together, we obtain the complete polyhedral decomposition of the knot. This is the geometric intuition behind the polyhedral expansion. We now explain a combinatorial method to describe the polyhedra.

1.1.2. Step 1. Sketch faces and edges into the diagram.

Recall that a diagram is a 4-valent graph lying on a plane, the plane of projection. The regions on the plane of projection that are cut out by the graph will be the faces, including the outermost unbounded region of the plane of projection. We start by labeling these, as in Figure 1.2.

Edges come from arcs that connect the two strands of the diagram at a crossing. These are called *crossing arcs.* For ease of explanation, we are going to draw each edge four times, as follows. Shown on the left of Figure 1.5 is a single edge corresponding to a crossing arc. Note that the edge is ambient isotopic in S^3 to the three additional edges shown on the right in Figure 1.5.

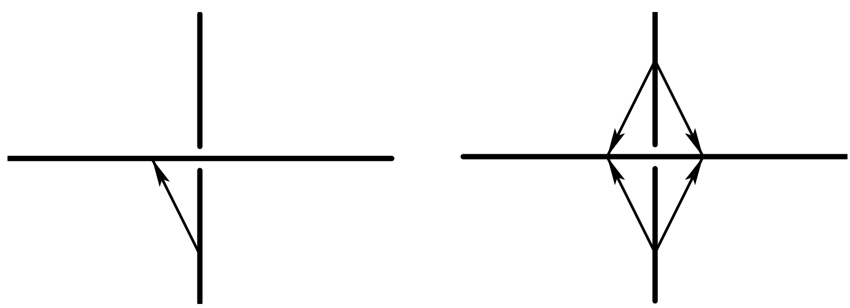


Figure 1.5. A single edge.

The reason for sketching each edge four times is that it allows us to easily visualize which edges bound the faces that we have already labeled. In Figure 1.6, we have drawn four copies of each of the four edges we get from crossing arcs of the diagram of the figure-8 knot. Note that the face labeled A, for example, will be bordered by three edges, one with two tick marks, one with a single tick mark, and one with no tick marks.

Remark 1.2. Orientations on the edges can be chosen to run in either direction; that is, arrows on the edges can run from overcrossing to undercrossing or vice versa, as long as we are consistent with orientations corresponding to the same edge. We have chosen the orientations in Figure 1.6 to simplify a later step, and to match a figure in Chapter 4. The opposite choice for any edge is also fine.

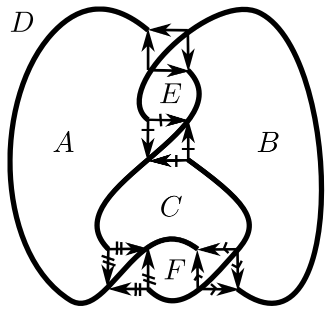


Figure 1.6. Edges of the figure-8 knot complement.

1.1.3. Step 2. Shrink the knot to ideal vertices on the top polyhedron.

Now we come to the reason for using *ideal* polyhedra, rather than regular polyhedra. Notice that the edges stretch from a part of the knot to a part of the knot. However, the manifold we are trying to model is the knot complement, $S^3 - K$. Therefore, the knot K does not exist in the manifold. An edge with its two vertices on K must necessarily be an ideal edge; that is, its vertices are not contained in the manifold $S^3 - K$.

Since the knot is *not* part of the manifold, we will shrink strands of the knot to ideal vertices. That is, we retract each knot strand to a single point. This may cause some confusion at first, because the strand of the knot is not homeomorphic to a single point. However, we are considering the *complement* of the strand. The complement of the strand on the boundary of the ball is homeomorphic to the complement of a single point on the boundary of the ball, so we replace strands by ideal vertices (single removed points).

Focus first on the polyhedron on top. Each component of the knot we "see" from inside the top polyhedron will be shrunk to a single ideal vertex. These visible knot components correspond to sequences of overcrossings of the diagram. Compare to Figure 1.3 — note that at an undercrossing, the component of the knot ends in an edge, but at an overcrossing the knot continues on. Moreover, note that at an overcrossing, the knot passes the same edge twice, once on each side.

In terms of the four copies of the edge in Figure 1.5, when we consider the polyhedron on top, we may identify the two edges which are isotopic along an overstrand, but not those isotopic along understrands. See Figure 1.7.

Shrink each overstrand to a single ideal vertex. The result is a pattern of faces, edges, and ideal vertices for the top polyhedron, shown in Figure 1.8. Notice that the face D is a disk containing the point at infinity.

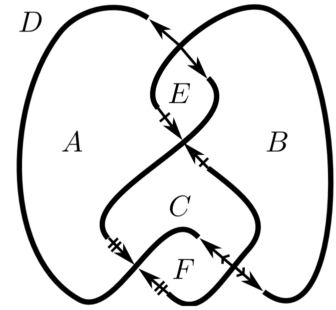


Figure 1.7. Isotopic edges in top polyhedron identified.

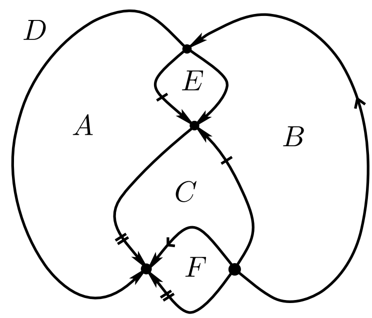


Figure 1.8. Top polyhedron, viewed from the inside.

1.1.4. Step 3. Shrink the knot to ideal vertices for the bottom polyhedron.

Notice that underneath the knot, the picture of faces, edges, and vertices will be slightly different. In particular, when finding the top polyhedron, we collapsed overstrands to a single ideal vertex. When you put your head underneath the knot, what appear as overstrands from below will appear as understrands on the usual knot diagram.

One way to see this difference is to take the 3-dimensional model constructed in Figure 1.4. Figure 1.3 shows the view of the faces meeting at an edge from the top. If you turn the model over to the opposite side, you will see how the faces meet underneath. Figure 1.9 illustrates this. Note that U now meets V, and S meets T.

In terms of the combinatorics, edges of Figure 1.5 that are isotopic by sliding an endpoint along an understrand are identified to each other on the bottom polyhedron, but edges only isotopic by sliding an endpoint along an overstrand are not identified.

As above, collapse each knot strand corresponding to an understrand to a single ideal vertex. The result is Figure 1.10.

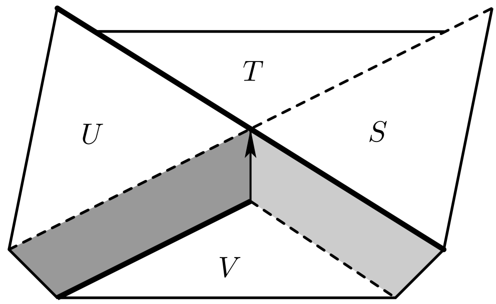


Figure 1.9. 3-dimensional model, opposite side as in Figure 1.3. Now faces V and U meet along an edge.

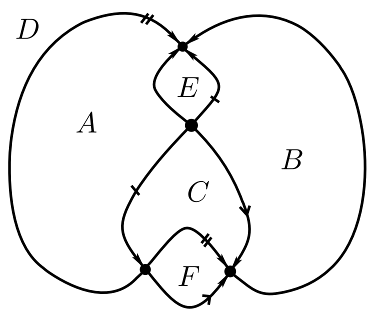


Figure 1.10. Bottom polyhedron, from the outside.

One thing to notice: we sketched the top polyhedron with our heads inside the ball on top, looking out. If we move the face D away from the point at infinity, then it wraps *above* the other faces shown in Figure 1.8.

On the other hand, we sketched the bottom polyhedron with our heads outside the ball on the bottom. If we move the face D away from the point at infinity, it wraps *below* the other faces shown in Figure 1.10.

1.1.5. Rebuilding the knot complement from the polyhedra. Figures 1.8 and 1.10 show two ideal polyhedra that we obtained by studying the figure-8 knot complement. We claim that they glue to give the figure-8 knot complement. That is, attach face A on the bottom polyhedron to the face labeled A on the top polyhedron, ensuring that the edges bordering face A match up. Similarly for the other faces.

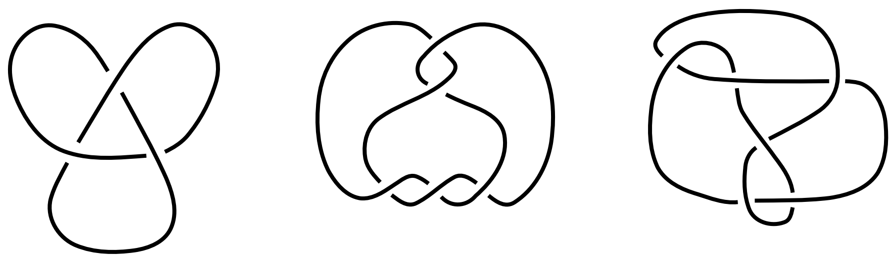
This process of gluing faces and edges gives exactly the complement of the knot. By construction, faces glue to give the faces illustrated in Figure 1.6, and edges glue to give the edges there, except that when we have finished, all four edges in an isotopy shown in that figure have been glued together.

1.2. Generalizing: Exercises

This polyhedral decomposition works for any knot or link diagram, to give a polyhedral decomposition of its complement.

Exercise 1.3. As a warm-up exercise, determine the polyhedral decomposition for one (or more) of the knots shown in Figure 1.11. Sketch both top and bottom polyhedra.

Your solution should consist of two ideal polyhedra, i.e. marked graphs on the surface of a ball, with faces and edges marked according to the gluing pattern. For example, the complete diagrams in Figures 1.8 and 1.10 form the solution for the figure-8 knot.



(a) Trefoil. (b) The 5_2 knot. (c) The 6_3 knot.

Figure 1.11. Three examples of knots.

Exercise 1.4. The examples of knots we have encountered so far are all alternating, as in Definition 0.14. The diagram of the knot 8_{19} in Figure 1.12 is not alternating. In fact, the knot 8_{19} has no alternating diagram.

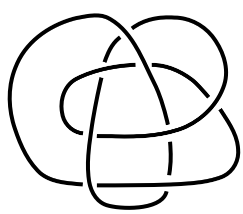


Figure 1.12. The knot 8_{19}, which has no alternating diagram.

Determine the polyhedral decomposition for the given diagram of the knot 8_{19}. Note: as above, many ideal vertices are obtained by shrinking overstrands to a point. However, you will have to use, for example, Figure 1.3 to determine what happens between two understrands.

Exercise 1.5. Recall that the *valence* of a vertex in a graph is the number of edges that meet that vertex. The valence of an ideal vertex is defined similarly.

(a) If a knot diagram is alternating, we obtain a very special ideal polyhedron. In particular, all ideal vertices will have the same valence. What is it? Show that the ideal vertices for an alternating knot all have this valence.
(b) What are the possible valences of ideal vertices in general, i.e. for nonalternating knots? For which $n \geq 0 \in \mathbb{Z}$ is there a knot diagram whose polyhedral decomposition yields an ideal vertex of valence n? Explain your answer, with (portions of) knot diagrams.

Exercise 1.6. In the polyhedral decomposition for alternating knots, the polyhedra are given by simply labeling each ball with the projection graph of the knot and declaring each vertex to be ideal.
(a) Prove this statement for any alternating knot. That is, prove that the decomposition gives polyhedra whose edges match the projection graph of the diagram.
(b) Show that for nonalternating knots, this is false. That is, the decomposition does not give polyhedra whose edges match the projection graph of the diagram.

Exercise 1.7. A graph admits a *checkerboard coloring* if all the complementary regions can be colored either white or shaded, with white faces meeting shaded faces across the edges. Any 4-valent graph can be checkerboard colored, particularly projection graphs of knot diagrams.

In the case of an alternating knot, faces are identified from the top polyhedron to the identical face on the bottom polyhedron, and the identification is by a *gear rotation*: white faces on the top are rotated once counter-clockwise and then glued to the corresponding face on the bottom; shaded faces on the top are rotated once clockwise and then glued. This is shown for the figure-8 knot in Figure 1.13. Prove that for the decomposition of any alternating knot, faces are identified by a gear rotation.

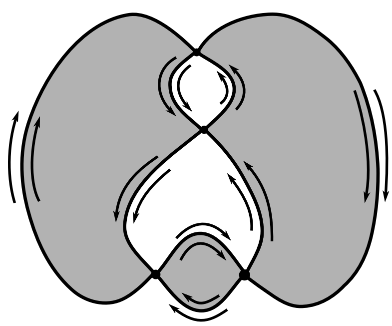


Figure 1.13. Checkerboard coloring and "gear rotation" for the figure-8 knot.

Exercise 1.8. The diagrams we have encountered so far are all reduced, as in Definition 0.6, but we can follow the above procedure for nonreduced diagrams. For example, we can obtain a polyhedral decomposition for diagrams which contain a nugatory crossing.

Show that the polyhedral decomposition of a knot diagram will contain a monogon, i.e. a face whose boundary is a single edge and a single vertex, if and only if the diagram has a simple nugatory crossing.

Exercise 1.9. Recall that a bigon is a region of a graph bounded by exactly two edges and exactly two vertices. Note that when a bigon appears in our polyhedral decomposition, the two edges of the bigon must be isotopic to each other. Hence, we sometimes will remove bigon faces from the polyhedral decomposition, identifying their two edges.

Let bigons be bygone. — William Menasco

For the figure-8 knot, sketch the two polyhedra that we get when bigon faces are removed. How many edges are there in this new, bigon-free decomposition? The resulting polyhedra are well-known solids in this case. What are they?

For each of the polyhedra obtained in Exercise 1.3, sketch the resulting polyhedra with bigons removed.

Exercise 1.10. Suppose we start with an alternating knot diagram with at least two crossings, and do the polyhedral decomposition above, collapsing bigons at the last step. What are possible valences of vertices? Sketch the diagram of a single alternating knot that has all possible valences of ideal vertices in its polyhedral decomposition.

What valences of vertices can you get if you don't require the diagram to be alternating but collapse bigons? Can you find 1-valent vertices? For any $n > 4 \in \mathbb{Z}$, can you find n-valent vertices?