Knot theory

Shintaro Fushida-Hardy
381B Sloan Hall, Stanford University, CA

This document contains knot theory notes, largely following Lickorish [Lic97]. Some solutions to exercises are also given. These notes were written as part of a reading course with Ciprian Manolescu in Spring 2020.

Contents

Chapter 1

Introduction to knots

The fundamental question in knot theory is the following: given two knots (1 dimensional loops embedded in 3 dimensional space), how can we determine whether or not the knots are the same? The tactic is to find knot invariants - quantities corresponding to each knot which are invariant under "untangling knots without allowing strands to pass through each other". (The basic legal moves are formalised as *Reidemeister moves.*) Thus whenever a given knot invariant exhibits different quantities for two knots, the two knots must be distinct. Knot invariants arise in many forms, including integers, polynomials, and homology theories. The game is to try to construct invariants which are useful (in the sense that they can actually be calculated), but complicated enough that they distinguish many inequivalent knots.

In this chapter we introduce some basic definitions and a few examples of knot invariants.

1.1 Basic definitions

Definition 1.1.1. A *link* L is a smoothly embedded closed 1-dimensional submanifold in $\mathbb{S}^3$ or $\mathbb{R}^3$. L is a *link of* m *components* if it is a collection of m simple closed curves. A *knot* is a link with one component. Equivalently, piecewise linear structures can be used to define knots and links.

Definition 1.1.2. Two links L_1 and L_2 are said to be *equivalent* if they satisfy either of the two equivalent properties:

- They are ambient isotopic. This means there is a homotopy $F_t : \mathbb{S}^3 \to \mathbb{S}^3$ such that each F_t is a diffeomorphism, F_0 is the identity, and F_1 maps L_1 onto L_2.

- There is an orientation preserving diffeomorphism $F : \mathbb{S}^3 \to \mathbb{S}^3$ mapping L_1 onto L_2.

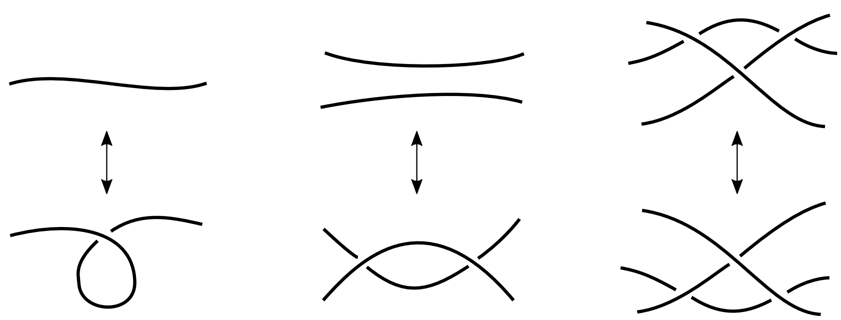
It's very difficult to draw knots in 3-dimensions. In particular, we have no way of drawing $\mathbb{S}^3$ in a meaningful way. However, by the following fundamental theorem it is always sufficient to draw figures in two dimensions.

Theorem 1.1.3. *Let L be a link in $\mathbb{R}^3$ or $\mathbb{S}^3$.*

1. *L admits a regular projection onto $\mathbb{R}^2$ or $\mathbb{S}^2$ respectively. (In fact, almost all projections are regular.) Such a projection is injective at all but finitely many points, which are the projections of only* two *points. By orienting the ambient space, the image of L under the projection is therefore a planar 4-regular graph whose vertices are decorated with crossing (over/under) information.*

2. *The ambient isotopy class of a link L is determined by a regular projection with crossing information.*

Combining points 1 and 2 above, the theory of knots and links is equivalently the theory of planar 4-regular graphs with crossing information at all of the vertices. These graphs are called *link diagrams*. Ambient isotopies of links descend to three basic transformations of link diagrams called *Reidemeister moves*, as shown in figure 1.1. Earlier we remarked



Type I   Type II   Type III

Figure 1.1: The three Reidemeister moves

that ambient isotopy is equivalent to orientation preserving diffeomorphism. Therefore considering orientation, we can produce new knots (which may or may not be equivalent to the original knot). We do this in two ways:

Definition 1.1.4. Let K be an oriented knot. The *reverse* of K, denoted rK, is the same knot (as smoothly embedded circle), but with the opposite orientation.

Definition 1.1.5. Let K be a knot. The *obverse* or *reflection* of K, denoted $\overline{K}$, is the image of K under an orientation reversing diffeomorphism of $\mathbb{S}^3$.

Example. $K, rK, \overline{K}, \overline{rK}$ are all distinct for the knot 9_{32}.

The basic operation for composing two knots is the connected sum, which we simply denote by $K_1 + K_2$. A knot is called the *unknot* if it bounds an embedded disk. Equivalently, the unknot is given by a knot diagram which is equivalent to the diagram with no crossings. By the definition of the connected sum, it is immediate that the unknot is a zero of addition of knots. We see later that no knot other than the unknot has an additive inverse.

Definition 1.1.6. A knot K is said to be *prime* if it is not the unknot, and $K = K_1 + K_2$ implies that K_1 or K_2 is the unknot.

To understand knots, we wish to understand all of the prime knots. In an attempt to tabulate prime knots, we first want some notion of complexity. This is given by our first knot invariant, the *crossing number*.

Definition 1.1.7. The *crossing number* of a knot K, denoted $c(K)$, is the minimum number of crossings required in a knot diagram representing K.

For example, the crossing number of the unknot is 0, and the crossing number of the trefoil knot is 3. It is easy to enumerate all possible knots with small crossing numbers, but the difficultly comes in determining which knots are prime and which knots are duplicates of other knots already in the list. (The most popular naming scheme for knots involves an enumeration using the crossing number.)

While the crossing number is very easy to define, and incredibly powerful for distinguishing knots (for example, it distinguishes the unknot from the trefoil knot, and these two knots from every other knot in existence), it is intractable. By its very definition (involving a “minimum”) it is difficult to compute given a general knot.

Another intractable but popular invariant is the *unknotting number*.

Definition 1.1.8. Let K be a knot. The *unknotting number*, denoted $u(K)$, is the minimum number of changes required in the crossing information of a knot diagram to make it equivalent to the unknot.

Example. The unknotting number of the unknot is 0. The unknotting numbers of 3_1 and 4_1 are 1. The unknotting number of 10_{11} is unknown as of the time of writing.

1.2 Some families of knots

We've now defined some invariants of knots and methods of manipulating existing knots, but we have yet to define any family of knots to study. One important family is the collection of *Pretzel links (and knots).*

Definition 1.2.1. The *pretzel link* $P(a_1, \dots, a_n)$ is the link defined by connecting n sets of a_i crossings in a spiral. When the sign of a_i is positive, the crossings are "anticlockwise upwards", and when a_i is negative, the crossing information is reversed. To clarify, figure 1.2 shows the Pretzel link $P(6, 6, -5, 5)$. By choosing different a_i, this proliferates an infinite number of links.

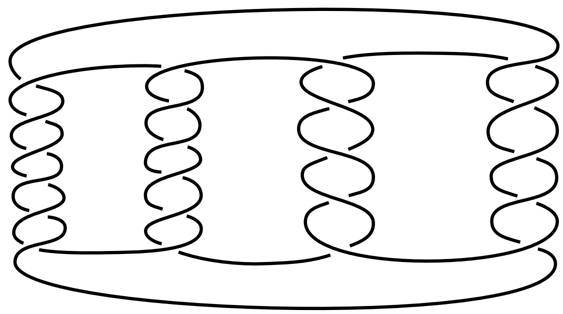


Figure 1.2: The pretzel link $P(6, 6, -5, 5)$

Definition 1.2.2. Another important family of links and knots are the so-called *rational links (and knots)*. The (p, q)-rational link is given by $C(a_1, \dots, a_n)$, where q/p has continued fraction

$$\frac{q}{p} = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots \cfrac{1}{a_{n-1} + \cfrac{1}{a_n}}}}.$$

The different ways of expressing q/p as a continued fraction all give rise to equivalent knots. Here $C(a_1, \dots, a_n)$ follows a similar construction to the pretzel links, with $C(2, 1, -2, 3)$ shown in figure 1.3.

Definition 1.2.3. A *braid of n strings* is a collection of n arcs traversing a box from left to right, keeping track of crossing information. Any braid gives rise to a link in a standard way by identifying the left and right edges of the box. All braids on n strings are elements of the *braid group* B_n, which has presentation

$$\langle \sigma_1, \dots, \sigma_{n-1} \mid [\sigma_i, \sigma_j] = 0 \text{ if } |i - j| \geq 2, \sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1} \rangle.$$

In the above, σ_i is realised as the braid where the ith arc crosses over the $i + 1$th arc, and all other arcs are unchanged.

Definition 1.2.4. The *Torus link* $T_{p,q}$ is formally the standard closed link obtained from the braid $(\sigma_1 \cdots \sigma_{p-1})^q$. Intuitively, it is the link obtained by wrapping p loops around a torus (in the toroidal direction), and applying q fractional twists (in the poloidal direction).

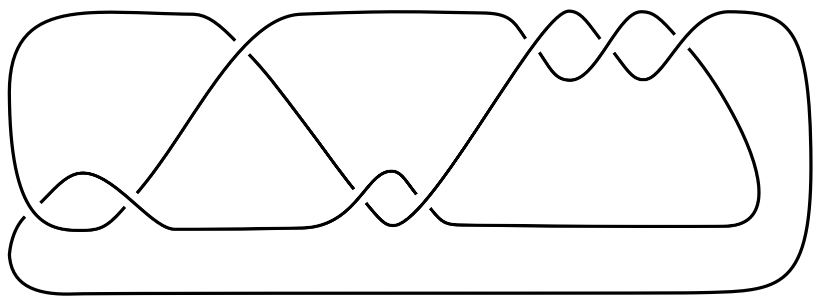


Figure 1.3: The (2/9)-rational knot, $C(2, 1, -2, 3)$.

Many methods of proliferating complicated knots in easy stages exist. Of course the connected sum gives a method of constructing more complicated knots from smaller pieces. Another important method is the *satellite knot*.

Definition 1.2.5. Let K, C be knots. Suppose K is embedded in a solid torus T, and $e : T \to \mathbb{S}^3$ is an embedding so that $e(T)$ is a regular neighbourhood of C. Then $e(K)$ is called a *satellite knot* with *pattern* K and *companion* C.

In fact, the connected sum $K = K_1 + K_2$ arises as a satellite knot with pattern K_1 and companion K_2.

1.3 Linking number

Suppose L is an oriented link. Then each crossing in its link diagram can be given a sign ± 1 depending on the crossing information, as in figure 1.4. The existence of type I

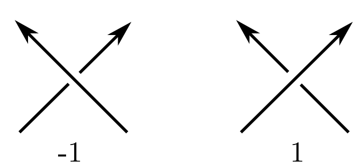


Figure 1.4: The standard sign convention for oriented crossings.

Reidemeister moves makes it clear that the signed count of crossings of an oriented knot is not a knot invariant. Moreover, we can't define something like the "minimum signed count" as applying enough type I moves will always ensure that the signed count can be zero. However, the signed count gives rise to an invariant for two-component links.

Definition 1.3.1. Let $L = L_1 \cup L_2$ be a two component oriented link. Then the *linking number* $\mathrm{lk}(L_1, L_2)$ is an invariant of L, defined to be half the signed count of crossings with one strand from L_1 and one from L_2.

The linking number embodies some homology theory as we soon see. The idea is to study *knot complements* as well as knots. Given a compact manifold X, with a knot K embedded in X, the *knot complement* of K is $X_K = X - \mathrm{int}(N)$, where N is a tubular neighbourhood of K in X.

In particular, if X is taken to be $\mathbb{S}^3$ (which is the most common choice), then X_K is a compact manifold has boundary a 2-torus, with homotopy type $\mathbb{S}^3 - K$. We can also consider simple closed curves on $\partial N \cong T^2$, and they turn out to be important in the understanding of the homology of X_K.

Definition 1.3.2. Let K be a knot in $\mathbb{S}^3$, with tubular neighbourhood N. A simple closed curve in ∂N is called a *meridian* of K if it is non-separating and bounds a disk in N (and hence bounds a disk that meets K at exactly one point). A simple closed curve in ∂N is called a *longitude* of K if it is homologous to K in N, but null-homologous in X_K.

The following theorem shows how we can understand the homology of X_K from K:

Theorem 1.3.3. *Let K be an oriented knot in $\mathbb{S}^3$, and X its knot complement. Then $H_1(X)$ is canonically isomorphic to $\mathbb{Z} = \langle [\mu] \rangle$, where μ is a meridian of K. Moreover, if C is any simple closed curve in X, then $[C] \in H_1(X)$ is $\mathrm{lk}(C, K)$. Finally, $H_2(X) = H_3(X) = 0$.*

This theorem generalises to complements of links:

Theorem 1.3.4. *Let L be an oriented link in $\mathbb{S}^3$ with m components, and X its link complement. Then $H_1(X)$ is canonically isomorphic to $\mathbb{Z}^m$, with generators given by a meridian for each component of L. Moreover,, $H_2(X) = \mathbb{Z}^{m-1}$.*

1.4 Exercises

Exercise 1.4.1. (Lickorish, 1.1) Show that 4_1 is equivalent to its reverse and reflection.

Solution: We first show that 4_1 is equivalent to its reflection, $\overline{4_1}$. This follows from a sequence of type II and type III Reidemeister moves, as shown in figure 1.5:

i. First, we drag the crossing labelled A to the top right of the diagram without changing any of the crossings (i.e. we apply an orientation preserving homeomorphism of the plane).

ii. Next we carry out a series of type II and type III Reidemeister moves on the segment labelled b, with the aim of pulling the strand to the top of the diagram.

iii. We continue applying type II and type III Reidemeister moves to b, along with two type I Reidemeister moves at the end.

iv. Finally we rotate the knot diagram by about 180° (in the plane), and wiggle the diagram to obtain what is clearly $\overline{4_1}$.

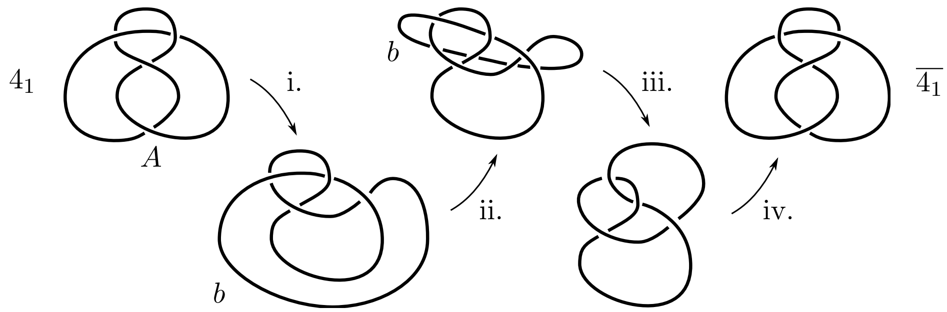


Figure 1.5: Proof that $4_1 = \overline{4_1}$.

To see that 4_1 is equivalent to its reverse, $r4_1$, assign an arbitrary orientation to the knot diagram for 4_1 shown in figure 1.5. Rotate the diagram about the vertical axis (in the plane of the figure) by 180°. This is an orientation preserving homeomorphism of the ambient space $\mathbb{R}^3$ which maps 4_1 onto its reverse. $\triangle$

Exercise 1.4.2. (Lickorish, 1.5) Show that every knot diagram can be changed into a diagram of the unknot by changing some crossings from "over" to "under". How many changes are necessary?

Solution: We first show that the unknotting number $u(K)$ is always finite. We then show that $u(K) \leq c(K)/2$, i.e. it is bounded above by half the crossing number.

The following proof is shown pictorially in figure 1.6. Consider a knot diagram K in the plane, where the plane is parametrised by (x, y)-coordinates. (Perturbing the diagram if necessary), the projection function $\pi_Y : K \to \mathbb{R}$ onto the y-coordinate has a unique global minimum, which we fix as our basepoint p. Moving right from p, every time a crossing is reached for the first time, change the crossing information so that we "pass over" rather than passing under.

Eventually we return to the basepoint. Given the new crossing information, we can parametrise an embedding of the knot in $\mathbb{R}^3$ as follows: first give the knot a parametrisation in time, i.e. K is the image of $\gamma : [0, \ell] \to \mathbb{R}^2$, parametrised by arc-length, with $\gamma(0) = \gamma(\ell) = p$. Next we use this to construct an embedding in $\mathbb{R}^3$, namely

$$e : [0, 2\ell] \to \mathbb{R}^3, \quad e(t) = \begin{cases} (\gamma(t), -t) & 0 \leq t \leq \ell \\ (p, t - 2\ell) & \ell \leq 2 \leq 2\ell. \end{cases}$$

(Note that this is not smooth at two points: $t = 0$ and $t = \ell$. However, e can be smoothed at these points by arbitrarily small perturbations.) Due to our assignment of crossing information, this is truly a knot represented by the modified knot diagram. Moreover, since p is the unique global minimum of π_Y, and $\partial_z e$ is strictly decreasing for (x, y, z) with $(x, y) \neq p$, and strictly increasing at (p, z), the projection of $e([0, 2\ell])$ onto the (y, z)-plane gives a knot diagram with no crossings. Therefore we have successfully unknotted K in finitely many moves (at most the number of crossings in K). Next we prove that $u(K)$ is

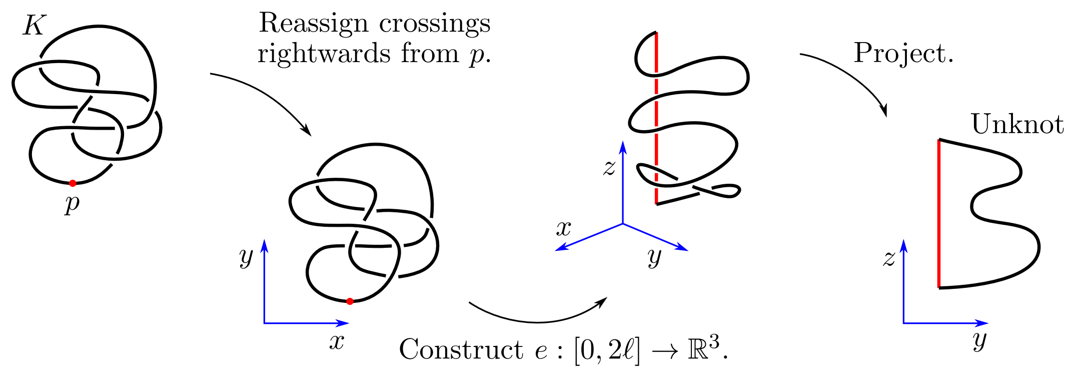


Figure 1.6: Proof that the unknotting number is finite.

bounded above by $c(K)/2$. Let K be a knot diagram exhibiting exactly $c(K)$ crossings. If the above procedure changes $N \leq c(K)/2$ crossings, we are done. Otherwise, follow the above procedure but "passing under" instead of passing over at each new new crossing. Then the end result is again the unknot (by constructing a similar embedding e' as above, but with $\partial_z e' = -\partial_z e$). Moreover, the crossing information is completely opposite that of the original procedure, so we must have changed $c(K) - N$ crossings. But this time $N > c(K)/2$, so $c(K) - N \leq c(K)/2$ as required. $\triangle$

Exercise 1.4.3. (Lickorish, 1.11) Let X_1 and X_2 be knot complements of non-trivial knots K_1 and K_2. Describe a homeomorphism $h : \partial X_1 \to \partial X_2$ so that $X_1 \sqcup_h X_2$ has the same homology as $\mathbb{S}^3$.

Solution: The boundary ∂X_1 is homeomorphic to $\mathbb{S}^1 \times \mathbb{S}^1$. Moreover, the two copies of $\mathbb{S}^1$ (in this product) can be chosen to be a meridian and longitude of K_1 respectively, which we denote by μ_1 and ν_1. Similarly, the boundary ∂X_2 is the product of a meridian and longitude μ_2 and ν_2 of K_2.

We define h to be a homeomorphism identifying μ_1 with ν_2, and ν_1 with μ_2. The idea is that $H_1(X_1)$ is generated by μ_1, but ν_2 is trivial in X_2. Similarly $H_1(X_2)$ is generated by μ_2, but ν_1 is trivial in X_1. Therefore both generators of $H_1(X_1)$ and $H_1(X_2)$ are trivial in X, so we expect $H_1(X)$ to be trivial.

It is immediate that $H_0(X)$ and $H_3(X)$ agree with $H_0(\mathbb{S}^3)$ and $H_3(\mathbb{S}^3)$. Finally, by the universal coefficient theorem and Poincaré duality, $H_2(X)$ is isomorphic to $H_1(X)$ modulo torsion. Thus if $H_1(X)$ vanishes, so does $H_2(X)$. In summary, to prove that $X = X_1 \sqcup_h X_2$ is a homology sphere, it suffices to prove that $H_1(X)$ is trivial.

We proceed by the long exact sequence of relative homology and the excision theorem. Specifically, we have the following diagram of exact sequences:

$$\begin{tikzcd}
H_1(X_1) \arrow[r] & H_1(X) \arrow[r] & H_1(X, X_1) \arrow[r] \arrow[d, "\cong"] & H_0(X_1) \\
H_1(\partial X_2) \arrow[r] & H_1(X_2) \arrow[r] & H_1(X_2, \partial X_2) \arrow[r] & H_0(\partial X_2).
\end{tikzcd}$$

Using H to denote reduced homology, $H_0(X_1) = H_0(\partial X_2) = 0$. We also know that $H_1(X_1)$ is generated by μ_1, and $H_1(X_2)$ by μ_2. Finally, $H_1(\partial X_2)$ is the abelian group generated by μ_2 and ν_2. Therefore the above diagram becomes

$$\begin{tikzcd}
\langle \mu_1 \rangle \arrow[r, "\varphi"] & H_1(X) \arrow[r] & H_1(X, X_1) \arrow[r] \arrow[d, "\cong"] & 0 \\
\langle \mu_2, \nu_2 \rangle \arrow[r] & \langle \mu_2 \rangle \arrow[r, "\psi"] & H_1(X_2, \partial X_2) \arrow[r] & 0.
\end{tikzcd}$$

Inspecting the bottom row, ψ is necessarily surjective. Moreover, exactness at $\langle \mu_2 \rangle$ forces ψ to be the zero map. Therefore $H_1(X_2, \partial X_2)$, and consequently $H_1(X, X_1)$, are trivial. But now this forces φ to be surjective. On the other hand, μ_1 is trivial in X (since it is identified with ν_2). Therefore φ is the zero map. It follows that $H_1(X)$ is trivial as required. △

Chapter 2

Prime factorisation

2.1 Seifert surfaces

Definition 2.1.1. A *Seifert surface* for an oriented link L in $\mathbb{S}^3$ is a connected compact oriented surface contained in $\mathbb{S}^3$ which has L as its boundary.

Theorem 2.1.2. *Every oriented link admits a Seifert surface.*

The preferred proof of this fact uses the so-called *Seifert algorithm.* The idea is to locally "uncross" all of the crossings so that a given knot diagram induces a collection of disjoint simple closed curves in the plane. Then each of these bounds an oriented disk (with the orientation induced from the boundary). Finally we "recross" the crossings to glue the disks together in such a way that the boundary of the resulting surface is exactly the given link. The resulting surface is oriented (and clearly compact and connected).

Remark. Any Seifert surface of a knot necessarily has one boundary component. In general, the number of boundary components is the number of link components.

Definition 2.1.3. The *genus* $g(K)$ of a knot K is defined by

$$g(K) = \min\{g(S) : S \text{ is a Seifert surface of } K\}.$$

Recall that the genus of a compact surface with boundary is the genus of the closed surface obtained by capping all boundary components with disks. Alternatively, we can define the genus of a compact surface via the Euler characteristic, using the formula

$$V - E + F = \chi(S) = 2 - 2g - b.$$

Corollary 2.1.4. *K is the unknot if and only if $g(K) = 0$.*

Clearly $g(0_1)$ is zero. On the other hand, any connected surface with one boundary component and genus 0 is necessarily a disk, so the boundary is 0_1.

Remark. In summary the *genus* is another knot invariant that distinguishes the unknot from all other knots. As remarked in chapter 1, invariants involving "min" are impractical for most computations. However, abstract properties of intractable invariants can have properties that still make them useful, as we see now.

Theorem 2.1.5. *Knot genus is additive:*

$$g(K_1 + K_2) = g(K_1) + g(K_2).$$

Proof. We give a proof outline. To see that $g(K_1 + K_2) \leq g(K_1) + g(K_2)$, consider Seifert surfaces F_1 and F_2 for K_1 and K_2 with minimal genus (embedded in $\mathbb{S}^3$). Observe that $F_1 \cup F_2$ is non-separating, as follows:

By exercise 2.3.2, given any knot C, $i(C, F_j) = \mathrm{lk}(C, K_j)$. Inspecting knot diagrams, we can always link K_j with copies of the unknot $0_1, 0_2$ to arrange $i(0_j, F_k) = \delta_{jk}$. Since the F_j are each connected embedded surfaces, this shows that $F_1 \cup F_2$ is not separating.

In particular, it follows that we can find an arc α from K_1 to K_2 in $\mathbb{S}^3 - (F_1 \cup F_2)$. Now choose a parallel arc α' from K_1 to K_2, so that α and α' bound a "strip" S (with twisting to match the orientations of the Seifert surfaces). The connected sum of K_1 and K_2 along α and α' is equipped with the Seifert surface $S \cup F_1 \cup F_2$, which has genus $g(K_1) + g(K_2)$. This proves that

$$g(K_1 + K_2) \leq g(S \cup F_1 \cup F_2) = g(F_1) + g(F_2) = g(K_1) + g(K_2).$$

Conversely, let F be a minimal genus Seifert surface for $K_1 + K_2$. Let $\mathbb{S}^2$ be a sphere embedded in $\mathbb{S}^3$, which intersects $K_1 + K_2$ transversely at two points a, b in such a way that K_1 and K_2 are contained in distinct components of $\mathbb{S}^3 - \mathbb{S}^2$. Then $\mathbb{S}^2$ intersects S along an arc α from a to b, and along finitely many disjoint circles.

Therefore F is cut into finitely many oriented components F_i. In particular, two of these satisfy $\partial F_1 = K_1 \cup n\mathbb{S}^1$ and $\partial F_2 = \partial K_2 \cup m\mathbb{S}^1$. Then $g(F_1) + g(F_2) \leq g(F)$. Although F_1 and F_2 are oriented, they are not Seifert surfaces since they have additional boundary components. Therefore we do surgery to create a new Seifert surface for $K_1 + K_2$ without increasing the genus which only intersects $\mathbb{S}^2$ along α.

Consider any component $\mathbb{S}^1 \subset S \cap \mathbb{S}^2$. Then $\mathbb{S}^1$ bounds two disks in $\mathbb{S}^2$, one of which contains α. The other disk may or may not intersect F. If it does, it does so along a circle. By induction, there exists $\mathbb{S}^1 \subset F \cap \mathbb{S}^2$ which bounds a disk which does not intersect F.

Consider a bicollar $\mathbb{S}^1 \times \mathbb{S}^0$ around this circle (in F). (One copy of $\mathbb{S}^1$ lies in each component of $\mathbb{S}^3 - \mathbb{S}^2$.) We now do surgery and replace the collar $\mathbb{S}^1 \times D^1$ with $D^2 \times \mathbb{S}^0$. The surgery is valid since F doesn't intersect the interior of either copy of D^2.

If this surgery is non-separating, then by the classification of surfaces, the genus of F decreases by 1. Therefore it must be separating. But $K_1 + K_2$ is connected and wasn't modified by the surgery, so it must belong to a single component of $F = \sqcup F_i$. Proceeding inductively, we produce a Seifert surface F' with boundary $K_1 + K_2$ which only intersects

$\mathbb{S}^2$ along α. Moreover, $g(F') \leq g(F)$. Therefore, cutting along α, we obtain Seifert surfaces F_1' and F_2' for K_1 and K_2 such that

$$g(K_1) + g(K_2) \leq g(F_1') + g(F_2') = g(F') \leq g(F) = g(K_1 + K_2).$$

This proves that $g(K_1) + g(K_2) = g(K_1 + K_2)$ as required. □

2.2 Prime decompositions

Using the additivity of genus, the following corollaries are almost immediate.

Corollary 2.2.1. *1. No non-trivial knots have additive inverses.*

2. There are infinitely many knots.

3. Any knot of genus 1 is prime.

4. Every knot is a finite sum of prime knots.

Proof. 1. If $K_1 + K_2 = 0_1$, then $g(K_1)$ and $g(K_2)$ are non-negative integers summing to 0, so $g(K_1) = g(K_2) = 0$. Therefore $K_1 = K_2 = 0_1$.

2. Let K be a non-trivial knot. Then for each $n \in \mathbb{N}$, $g(nK) = ng(K)$, and $g(K)$ is non-zero, so each of the nK are distinct knots.

3. Suppose K has genus 1, and $K = K_1 + K_2$. Then the genus of K_1 or K_2 is necessarily zero, so K_1 or K_2 is the unknot. Moreover, K is not the unknot since it has positive genus. Therefore K is prime.

4. Suppose K is not prime. Then $K = K_1 + K_2$, where K_1 and K_2 have strictly lower genus than K. By induction on the genus (and point 3 above), K is the connected sum of at most $g(K)$ prime knots. □

In fact, analogously to unique factorisation domains, we find that prime knots appear uniquely as factors of knots:

Theorem 2.2.2. *Suppose $K = P + Q$, where P is prime. Suppose moreover that $K = K_1 + K_2$. Then one of the following hold:*

- *There exists K_1' such that $K_1 = P + K_1'$, and $Q = K_1' + K_2$.*
- *There exists K_2' such that $K_2 = P + K_2'$, and $Q = K_2' + K_1$.*

Proof. We first establish notation. Let Σ_K be a 2-sphere embedded in $\mathbb{S}^3$ which realises the decomposition $K = K_1 + K_2$. Let Σ_P be a 2-sphere which realises the decomposition $K = P + Q$. Let B be the (closed) component of $\mathbb{S}^3$ carved out by Σ_P which "contains" P. All intersections are assumed to be transverse. In particular, $\Sigma_K \cap \Sigma_P$ is a disjoint union of (unknotted) circles. The proof idea is as follows:

1. The circles in $\Sigma_K \cap \Sigma_P$ each have linking number -1, 0, or 1 with K. First we do surgery to eliminate the circles with linking number 0.

2. The circles with linking number -1 or 1 bound components of $\Sigma_K \cap B$. We show that these surfaces are either disks or annuli.

3. We eliminate any components of $\Sigma_K \cap B$ which are disks.

4. The "trivial" annular components are eliminated in a similar manner to above. If all components are "trivial", we conclude that $\Sigma_K \cap \Sigma_P$ is empty, and deduce the conclusion of the theorem.

5. Given any "non-trivial" annular component, we use it to deduce the conclusion of the theorem.

1. Suppose $S \subset \Sigma_K \cap \Sigma_P$ is a component of $\Sigma_K \cap \Sigma_P$. This cuts Σ_K into two disks, which either intersect K once each, or one of which intersects K twice. In the former, S has linking number ± 1 with K, and in the latter, it has linking number 0 with K.

Let C be a component of $\Sigma_K \cap \Sigma_P$ in the latter case. C bounds a disk on Σ_K which doesn't intersect K. Without loss of generality, C is the *innermost* such circle, so that it bounds a disk whose interior doesn't intersect Σ_P. We do surgery: consider a bicollar $\mathbb{S}^1 \times \mathbb{S}^0$ of Σ_P around C. This bounds a copy of $\mathbb{S}^1 \times D^1 \subset \Sigma_P$ which we replace with $D^2 \times \mathbb{S}^0$. This is necessarily separating (since Σ_P has genus 0). Since K is connected and intersects Σ_P exactly twice, it intersects exactly one of the components obtained by surgery. Therefore we disregard the other component, and observe that the intersection $S \subset \Sigma_K \cap \Sigma_P$ has been eliminated. We inductively eliminate all components which have linking number 0 with K.

2. $\Sigma_K \cap B$ is a disjoint union of genus zero surfaces bound by circles each with linking number ± 1 with K. Suppose $S \subset \Sigma_K \cap B$ is a surface bound by n circles. Then $\Sigma_K \cap (\mathbb{S}^3 - B)$ is a copy of n disjoint disks, each intersecting K once each. This forces $n \leq 2$. By the classification of surfaces, either S is a disk or an annulus.

3. Suppose $S \subset \Sigma_K \cap B$ is a component which is a disk. It necessarily intersects K exactly once. On the other hand, ∂S cuts Σ_P into two disks $\Sigma_P' \sqcup_{\partial S} \Sigma_P''$ which also intersect K once each. This gives $B = B' \sqcup_S B''$, with $\partial B' = \Sigma_P' \sqcup_{\partial S} S$, and $\partial B'' = \Sigma_P'' \sqcup_{\partial S} S$. This gives an expression $P = P' + P''$, but P is prime, so either P' or P'' is trivial. Assuming P' is trivial, we replace B and Σ_P with B'' and $\Sigma_P'' \sqcup_{\partial S} S$. Then $\Sigma_K \cap B$ no longer has the component S (by a small perturbation using the orientation). Inductively, all "disk" components are eliminated.

4. Now any component $S \subset \Sigma_K \cap B$ is an annulus, A'. This has a boundary $C_1 \sqcup C_2 \subset \Sigma_K \cap \Sigma_P$. This also bounds an annulus A on Σ_P. Observe that A' cuts B into two pieces, one of which has boundary $A \cup A'$. We denote this component of B by M.

On the other hand, $\Sigma_P - A$ consists of two disjoint disks, each intersecting K exactly once each. Choose one of the disks, say D_1. Then thickening D (into the ball) we obtain a 3-ball $\widetilde{D}$ whose boundary intersects K twice, with $\widetilde{D} \cap K$ unknotted. But now $\widetilde{D} \cup M$ is topologically a ball, and its boundary intersects K exactly twice. Since P is a prime knot contained in B, either P is contained in $\widetilde{D} \cup M$ or its intersection with $\widetilde{D} \cup M$ is trivial. In the latter case, we update B by replacing it with $B - (\widetilde{D} \cup M)$. This decreases the number of components of $\Sigma_K \cap \Sigma_P$.

If all components bounded by two annuli are trivial as above, by induction we eliminate all components of $\Sigma_K \cap \Sigma_P$. Then (without loss of generality) Σ_P lies inside the component of $\mathbb{S}^3 - \Sigma_K$ containing K_1. This expresses K_1 as $P + K_1'$. On the other hand, it expresses Q has $K_1' + K_2$.

5. Finally suppose we find a non-trivial component $\widetilde{D} \cup M$. Without loss of generality M is contained in the component of $\mathbb{S}^3$ (carved out by Σ_K) containing K_1. Surgically replace the other component by a trivial ball-arc pair, by noting that its intersection with B was already trivial. Then K is replaced with K_1, while $\widetilde{D} \cup M$ is still a solid ball realising the sum $K = P + K'$. This proves that $K_1 = P + K_1'$.

Finally to show that $Q = K_1' + K_2$, we follow a similar procedure: surgically replace $\widetilde{D} \cup M$ with a solid torus T, in such a way that $\widetilde{D} \cup T$ is still a ball. (Specifically glue ∂D to a longitude of T). This replaces K with Q. On the other hand, the component of Σ_K containing K_2 has been unchanged, and we have realised $Q = K_1' + K_2$. This completes the proof. □

Corollary 2.2.3. *Suppose P is prime, and $P + Q = P + K$. Then $Q = K$.*

Proof. By the previous theorem, there are two cases to consider. First suppose there exists P' such that $P = P + P'$, and $Q = P' + K$. Since P is prime, P' must be the unknot, so $P' + K = K$. Therefore $Q = K$. Next suppose there exists K' such that $K = P + K'$ and $Q = K' + P$. Then we immediately have $Q = K$. □

Theorem 2.2.4. *Every knot has a unique (up to order) factorisation into primes.*

Proof. Suppose $K = P_1 + \cdots + P_n = Q_1 + \cdots Q_m$ are factorisations of K into prime knots. By the previous theorem, P_1 is a summand of Q_1 or a summand of $Q_2 + \cdots + Q_m$. But now by inductively applying the theorem, P_1 is a summand of Q_i for some i. Since each Q_i is prime, it follows that $P_1 = Q_i$. Without loss of generality, choose $i = 1$. By the previous corollary, we can now "cancel" the P_1, so that $P_2 + \cdots + P_n = Q_2 + \cdots + Q_m$. Inductively we find that $P_j = Q_j$ for each $1 \leq j \leq n$. But now $Q_{n+1} + \cdots + Q_m = 0$, so each Q_j for $j > n$ is the unknot. This proves that $n = m$, and moreover that the P_j are equal to Q_j (up to order). □

2.3 Exercises

Exercise 2.3.1. (Lickorish, 2.1) Prove that a non-trivial torus knot is prime by considering the way in which a 2-sphere meeting the knot at two points intersects the torus that contains the knot.

Solution: Suppose K is a torus knot, and write $K = K_1 + K_2$. Then there is a 2-sphere S which meets K (transversely) at exactly two points, with K_1 and K_2 contained in distinct components of $\mathbb{S}^3 \setminus S$ (according to the Schönflies theorem). Since K is a torus knot, it is itself embedded on the surface of a torus T. By the transversality theorem, we assume $T \cap S$ is transverse, so that it is a 1-manifold. $T \cap S$ is necessarily compact and without boundary, so it is a disjoint union of simple closed curves.

Since $T \cap S$ is a disjoint union of embedded circles, either K meets two components of $T \cap S$ once each, or K meets one component of $T \cap S$ twice. We first show that the former is impossible. Note that the simple closed curves in the first case cannot be null-homotopic: in that instance the Jordan curve theorem ensures that each curve intersects K an even number of times. It follows that the curve is necessarily a meridian, so if it intersects the torus knot exactly once, this contradicts the non-triviality of K.

Next suppose K intersects a single component γ of $T \cap S$, twice. By the Schönflies theorem, the signed count of intersections of K and S must vanish. Therefore $\gamma \cap K$ contains two points, with signs 1 and -1. Suppose first that γ is not null-homotopic. Then γ is non-separating, and cutting along γ gives a cylinder. K induces two arcs on the cylinder, each with end points on the same boundary component. Any such arc forms a bigon with the boundary, so gluing the cylinder back to an arc, K is guaranteed to bound a disk. This contradicts the non-triviality of K. It follows that γ is null-homotopic, so it necessarily bounds a disk D. Recall our expression $K = K_1 + K_2$. We have shown that one of tne of the components K_i is induced from an arc in D, so K_i is the unknot. This proves that K is prime. $\triangle$

Exercise 2.3.2. (Lickorish, 2.4) Suppose F is a Seifert surface for an oriented knot K, and let C be an oriented simple closed curve in $F - K$. Prove that $\operatorname{lk}(C, K) = 0$.

Solution: We will prove a more general result. Namely, if K_1, K_2 are any two oriented knots, then $\operatorname{lk}(K_1, K_2)$ is the signed intersection number $i(K_1, F)$ of K_1 with any Seifert surface F of K_2.

To prove this, we first show that $lk(K_1, K_2) = i(K_1, F)$, where F is the Seifert surface constructed by the Seifert algorithm. Next we show that $i(K_1, F)$ is independent of the choice of Seifert surface.

Consider the link diagram of K_1 and K_2. Let F be the Seifert surface of K_2 obtained by the Seifert algorithm. Let K_2' be the "uncrossed" link corresponding to K_2, and F' the "uncrossed surface" bound by K_2' which appears in the construction of F. Since K_1 and K_2 are in general position, and all orientations are preserved, the signed count of crossings of K_1 and K_2 is the same as that of K_1 and K_2'. Similarly the signed count of intersections of F with K_1 is equal to that of F'. It suffices to show that $i(K_1, F') = \operatorname{lk}(K_1, K_2')$.

Since K_2' is a collection of disjoint unknots, and F' is a collection of disks bounded by the components of K_2', it suffices to show equality of linking and intersection numbers for a single component. Assume without loss of generality that K_2' is the unknot, and F' is a disk. (Working in the projection), the intersection of K_1 with F' is a union of arcs with their endpoints on K_2'. Fix an arc γ. There are three cases: either the signed count of intersections is -2, 0, or 2. The signed count are -2 and 2 exactly when the crossings for the arc alternate, which corresponds to the arc passing through the surface F' at one point. On the other hand, the signed count is 0 exactly when the crossings of the arc and K_2' don't alternate, and hence the arc doesn't pass through F'. Therefore the signed counts of intersections $i(K_1, F')$ is exactly $\mathrm{lk}(K_1, K_2')$ as required, from which it follows that $i(K_1, F) = \mathrm{lk}(K_1, K_2)$.

It remains to prove that $i(K_1, F)$ is independent of the choice of Seifert surface. This is a homological result: suppose F, F' are both Seifert surfaces of K_2. Then F and F' represent 2-chains in homology. Reversing the orientation of F' and gluing it to F along their common boundary (K_2) gives an oriented surface without boundary, and hence a 2-cycle. Therefore its image under ∂ must vanish. But in particular this requires that the algebraic intersection number of $F \sqcup_{K_2} F'$ with any knot vanishes. Therefore $i(K_1, F) = i(K_1, F')$ as required. $\triangle$

Exercise 2.3.3. (Lickorish, 2.9) Show that connected sums of links are not well defined, and that $L_1 + L_2 = L_1 + L_3$ does not imply that $L_2 = L_3$.

Solution: First let L_1 be a trivially linked unknot and trefoil knot (i.e. a trefoil and unknot with linking number 0). Let L_2 be a trefoil knot. Then $L_1 + L_2$ is not well defined, since it either gives two knots with crossing numbers of 3, or two knots where one of them is the unknot. These cannot be equivalent links.

We now write $L = L_1 + L_2$ to mean an *internal sum*; in other words L is already understood, and L_1 and L_2 are links such that L is their sum given a choice of components to carry out the operation (even though the connected sum of links isn't well defined as an external operation.) We show that $L_1 + L_2 = L_1 + L_3$ doesn't imply that $L_2 = L_3$. Concretely, let L be a link with two components; a connected sum of two trefoils with the same chirality, and an unlinked trefoil again with the same chirality. (Hereafter all trefoil knots are chosen to have the same chirality.) Now if L_1 is a trefoil, and L_2 consists of an unlinked unknot and the sum of two trefoils, we see that $L = L_1 + L_2$. On the other hand, if L_3 consists of two unlinked trefoils, again $L = L_1 + L_3$. But for the same reason as earlier, L_2 is not equivalent to L_3. $\triangle$

Chapter 3

Jones and HOMFLY polynomials

3.1 Jones polynomial

Definition 3.1.1. The *Kauffman bracket* is a map

$$\langle - \rangle : \text{ link diagrams } \to \mathbb{Z}[A, A^{-1}]$$

determined uniquely by

- $\langle 0_1 \rangle = 1$,
- $\langle L \sqcup 0_1 \rangle = (-A^2 - A^{-2})\langle L \rangle$,
- $\langle \times \rangle = A\langle)(\rangle + A^{-1}\langle \asymp \rangle$

Lemma 3.1.2. It is straight forward to verify that the Kauffman bracket is invariant under type II and type III Reidemeister moves. On the other hand, type I Reidemeister moves have the following effect:

$$\langle \propto \rangle = -A^{3}\langle \smile \rangle, \quad \langle \propto \rangle = -A^{-3}\langle \smile \rangle.$$

Example. Let H denote the usual diagram of a Hopf link. Then by two applications of the above lemma, we find that

$$\langle H \rangle = -A^4 - A^{-4}.$$

Using this calculation, it straight forward to show that

$$\langle 3_1 \rangle = A^{-7} - A^{-3} - A^5,$$

where 3_1 is the usual (rotationally symmetric) diagram of the trefoil knot. Finally if 4_1 is the usual diagram of the figure-eight knot (e.g. as used in exercise 1.4.1 (Lickorish 1.1)), then

$$\langle 4_1 \rangle = A^{-1}\langle 3_1 \rangle - A^4\langle H \rangle = A^{-8} - A^{-4} + 1 - A^4 + A^8.$$

Definition 3.1.3. The *writhe* of an oriented diagram D, denoted $w(D)$, is the signed sum of the crossings of D (with signs assigned as in figure 1.4). This is invariant number type II and III Reidemeister moves, but changes by -1 under a type I Reidemeister move (and $+1$ under the mirrored Reidemeister move).

Theorem 3.1.4. *Let D be an oriented link diagram. Then*

$$(-A)^{-3w(D)}\langle D\rangle$$

is a link invariant of D.

Proof. If D is modified by type II or III Reidemeister moves, neither the writhe nor Kauffman bracket change, so the above epression is unchanged. If $D \rightsquigarrow D'$ is a type I Reidemeister move, by the earlier lemma,

$$(-A)^{-3w(D)}\langle D\rangle = (-A)^{-3(w(D')+1)}(-A^3\langle D'\rangle) = (-A)^{-3w(D')}\langle D'\rangle.$$

□

Definition 3.1.5. The *Jones Polynomial* is the link invariant $V(L)$ defined by

$$V(L) = \Big((-A)^{-3w(D)}\langle D\rangle\Big)_{t^{1/2}=A^{-2}} \in \mathbb{Z}[t^{1/2}, t^{-1/2}]$$

for any link diagram D of L.

For *knots* (and more generally links with an odd number of components), the Jones polynomial is truly a polynomial in the sense that it lies in $\mathbb{Z}[t, t^{-1}]$. The Jones polynomial does not detect the orientation of knots, since if the orientation of every component of a link is reversed, the signed counts are unchanged. It is currently unknown whether or not there is a non-trivial knot K such that $V(K) = 1$.

Example. The writhe of the trefoil is 3. Therefore its Jones polynomial is

$$V(3_1) = -t^{-4} + t^{-3} + t^{-1}.$$

The writhe of the figure-8 knot is 0. Therefore its Jones polynomial is

$$V(4_1) = t^{-2} - t^{-1} + 1 - t + t^2.$$

Theorem 3.1.6. *The Jones polynomial is characterised by normalisation and a skein relation. More precisely, the Jones polynomial is the unique function*

$$V : \{\ \textit{oriented links in } \mathbb{S}^3\} \to \mathbb{Z}[t^{1/2}, t^{-1/2}]$$

satisfying

- $V(0_1) = 1$,
- $t^{-1}V(L_+) - tV(L_-) + (t^{-1/2} - t^{1/2})V(L_0) = 0$, *for any* $L_0, L_\pm$ *which have identical link diagrams at all but one crossing, at which point* L_+ *has a positive crossing,* L_- *a negative crossing, and* L_0 *an un-crossing (in the unique orientation-compatible way).*

Proof. A straight forward computation shows that the above properties are satisfied by the Jones polynomial. For the converse, it suffices to show that the above properties uniquely define $V(L)$ for all link diagrams L. We proceed by induction on the number of crossings of a link diagram. The base case is known; $V(0_1) = 1$. For the inductive step, suppose $V(L)$ is known for all links L with at most n crossings. Suppose L' is a link with $n+1$ crossings. By changing k crossings for some k, we obtain a diagram of an unlink U. By the skein relations, $V(L')$ is expressed in terms of $V(U)$ and $V(L_1), \ldots, V(L_k)$ where each L_i is a link diagram with n crossings. Therefore the problem reduces to knowing $V(U)$ for all unlinks U.

Using the fact that V is a link invariant, we can use the skein relations to inductively conclude that the unlink with c components has invariant $(-t^{-1/2}-t^{1/2})^{c-1}$. This completes the proof. □

Proposition 3.1.7. Two immediate properties of the Jones polynomial are

- $V(K_1 + K_2) = V(K_1)V(K_2)$,
- $V(K)(t) = V(\overline{K})(t^{-1})$.

Proof. For the first fact, simply compute the Jones polynomial of K_1 first. For the second fact, observe that reflections do not change the Kauffman bracket, but they change the sign of the writhe. □

Corollary 3.1.8. *The trefoil is distinct from its reflection.*

Using the multiplicative property of the Jones polynomial, one might like to attempt to prove that knots admit prime decompositions. The proof would look something like this:

Let K *be a knot. If* K *isn't prime, it has a decomposition* $K = K_1 + K_2$ *into non-trivial knots. Then* $V(K) = V(K_1)V(K_2)$*. Since* $\mathbb{Z}[t^{1/2}, t^{-1/2}]$ *is a unique factorisation domain, by induction* $V(K)$ *can only be factored into finitely many terms, at which point each factor is the Jones polynomial of a prime knot.*

The reason this proof fails is because there is no guarantee that $V(K_i)$ isn't a unit if K_i is non-trivial. In fact, we noted earlier that it is open whether or not there exist non-trivial knots whose Jones polynomial is 1. Conversely, it is clear that the "uniqueness" aspect of a unique factorisation domain will not guarantee uniqueness of prime decompositions, since different knots can have the same Jones polynomial. In summary, we obtain the following question:

Question. Does there exist a non-trivial knot K such that $V(K)$ is a *unit* in $\mathbb{Z}[t, t^{-1}]$?

3.2 HOMFLY polynomial

Theorem 3.2.1. *There is a unique link invariant*

$$P : \{ \text{ oriented links in } \mathbb{S}^3 \} \to \mathbb{Z}[\ell^{\pm 1}, m^{\pm 1}]$$

such that

- $P(0_1) = 1$,
- $\ell P(L_+) + \ell^{-1} P(L_-) + m P(L_0) = 0$.

P is called the HOMFLY polynomial.

Proof. We proceed with a proof by induction on the number of crossings. At each step we must ensure the skein relation and invariance under Reidemeister moves. For each n, let $\mathcal{D}_n$ denote the set of link diagrams in the plane with at most n crossings. We will prove that for all $n \in \mathbb{N}$,

1. The skein relation holds for any three diagrams in $\mathcal{D}_n$ related in the assigned way.
2. $P(D)$ is unchanged by Reidemeister moves on D involving at most n crossings.
3. Any *ascending diagram* D of a link in $\mathcal{D}_n$ with k components has

$$P(D) = \Big(\frac{-\ell - \ell^{-1}}{m}\Big)^{k-1} = \mu^{k-1}.$$

By an ascending diagram, we mean a link diagram where each component has been ordered and assigned a base point, so that tracing out the components in order, each crossing is first encountered as an underpass.

Base case: $\mathcal{D}_0$ consists only of unlinked diagrams of the unknot. We verify the above three properties for $n = 0$:

The first result holds vacuously; no diagrams in $\mathcal{D}_0$ contain crossings. The second result also holds vacuously; all Reidemeister moves involve at least 1 crossing. Finally the the third property can be forced, since members of $\mathcal{D}_0$ are precisely diagrams consisting of unlinks of k components.

For the inductive step, suppose that $P : \mathcal{D}_{n-1} \to \mathbb{Z}[\ell^{\pm 1}, m^{\pm 1}]$ has been defined, so that the above properties hold. We define $P : \mathcal{D}_n \to \mathbb{Z}[\ell^{\pm 1}, m^{\pm 1}]$ as follows:

(i) Let D be a diagram with n crossings and k components in some order.

(ii) Choose a base point in each component, and let αD denote the associated ascending diagram. Define $P(\alpha D) = \mu^{k-1}$.

(iii) Changing crossings one at a time, αD can be modified to achieve D. With each crossing change, there is a bi-product with $n-1$ crossings (according to the Skein relations). By the inductive hypothesis, the value of P on such a bi-product diagram is known. $P(D)$ is then defined using the Skein relations (applied to $P(\alpha D)$ and each of the bi-products).

Each step required a choice: we must show that $P(D)$ is well defined by proving that $P(D)$ is unaffected by the component order, base point position, and crossing-change order. We start with the choice in (iii).

Let $D_{\alpha\beta}$ be a diagram with $\alpha, \beta \in \{+, 0, -\}$. The indices α and β represent the signs of fixed crossings. To prove that the order in which crossings are changed doesn't affect P, we transpose the operations. We obtain the following four skein relations:

$$\begin{aligned}
\ell P(D_{++}) + \ell^{-1} P(D_{-+}) + mP(D_{0+}) &= 0 \\
\ell P(D_{-+}) + \ell^{-1} P(D_{--}) + mP(D_{-0}) &= 0 \\
\ell P(D_{++}) + \ell^{-1} P(D_{+-}) + mP(D_{+0}) &= 0 \\
\ell P(D_{+-}) + \ell^{-1} P(D_{--}) + mP(D_{0-}) &= 0.
\end{aligned}$$

For example, to get from D_{--} to D_{++}, we can go via D_{-+} or D_{+-}. To prove that the order doesn't matter, we must show that

$$mP(D_{0+}) - \ell^{-2} mP(D_{-0}) = mP(D_{+0}) - \ell^{-2} mP(D_{0-}).$$

By the inductive hypothesis, the skein relation holds, so we have

$$\begin{aligned}
\ell P(D_{0+}) + \ell^{-1} P(D_{0-}) + mP(D_{00}) &= 0 \\
\ell P(D_{+0}) + \ell^{-1} P(D_{-0}) + mP(D_{00}) &= 0.
\end{aligned}$$

The result follows.

Next we prove that there is no dependence on the choice in (ii), i.e. the choice of base point for each component is irrelevant. Suppose b is the original base point of a component, and b' is a new base point, which is placed one crossing x after b. Let βD denote the ascending diagram of D using this new base point. If the crossing x involved a distinct component, then $\alpha D = \beta D$, so nothing changes. If the crossing consists of segments from our chosen component, then βD differs from αD be reversing this crossing. But annulling this crossing as in the skein relation gives another ascending diagram, with exactly one more component. (This is easily seen by considering the "heights" of the four segments meeting at a crossing.) Therefore

$$\ell P(\beta D) + \ell^{-1} \mu^{k-1} + m\mu^k = 0.$$

This gives $P(\beta D) = \mu^{k-1}$. This shows that $P(\alpha D) = P(\beta D)$. But above we showed that the order in which crossings are changed doesn't affect the value of $P(D)$, so one such process is to change from βD to αD, and then obtain D. Since $P(\alpha D) = P(\beta D)$, the final value for $P(D)$ is again unchanged.

So far we have shown that the value of $P(D)$ does not depend on the base points of components, or the order in which crossings are changed. The only possible dependence is on the order of the components. It is also clear that P satisfies the skein relation for any three related diagrams with at most n crossings. Before proving that P is independent of the order of components, it remains to prove that P is invariant under Reidemeister moves.

The idea for each of the types of Reidemeister moves is to push everything to the start of the algorithm. We prove the invariance under type I Reidemeister moves, as type II and III are similar. Suppose D is modified by a type I Reidemeister move which involves at most n crossings. Then the only possibility is that a crossing was removed. Let D' denote the diagram from D after removing a crossing with the type I move. Since the choice of base point in each component has been shown to be immaterial, we can choose the base point to be immediately before the crossing. Then αD still contains the crossing, and the algorithm to calculate $P(D)$ from αD doesn't affect the crossing. Moreover, it follows that calculating $P(D')$ from $\alpha D'$ follows precisely the same crossing changes. Therefore it suffices to show that $P(\alpha D) = P(\alpha D')$. But this is immediate, since both αD and $\alpha D'$ are ascending diagrams with the same number of components.

Proving invariance under type II and type III moves in a similar fashion, it remains to prove that $P(D)$ is invariant under the choice of component ordering. We do not describe the details here, but the idea is as follows: suppose we choose a new ordering, and obtain an ascending diagram βD. But now we can re-label the ordering, consider βD is a diagram obtained by changing crossings in αD. Then $P(\beta D)$ is obtained from $P(\alpha D) = \mu^{k-1}$. Since β was arbitrary, it suffices to prove that $P(\beta D) = \mu^{k-1}$ to prove that $P(D)$ doesn't depend on the choice of component order. □

3.3 Exercises

Exercise 3.3.1. (Exercise to myself) Can I modify the definition of the Kauffman bracket to find a link invariant?

Solution: For any (possibly multivariable) polynomials $P, Q, R \in \mathbb{Z}[A_i]$ let I be a map from link diagrams to $\mathbb{Z}[A_i]$ defined by

(a) $I(0_1) = 1$,

(b) $I(L \sqcup 0_1) = PI(L)$,

(c) $I(\slashoverback) = QI(\smoothing) + RI(\hsmoothing)$.

Then by relations (b) and (c),

$$I(\rightkink) = (R + PQ)I(\frown), \quad I(\leftkink) = (PR + Q)I(\frown).$$

To achieve link invariance, we must have $R + PQ = PR + Q = 1$. This forces $RP^2 = P + R - 1$, so in particular

$$2 \deg P + \deg R \leq \max\{\deg P, \deg R\}.$$

This forces P to be an integer (so that it has degree 0 or $-\infty$). Thus write $P = n \in \mathbb{Z}$, so that

$$R + nQ = nR + Q = 1.$$

Next we consider type II moves. By using invariance of I under type I moves (and (b) and (c)), it is straight forward to show that we must have

$$Q + R = 0.$$

Substitution into the earlier relations give $(1-n)R = (n-1)R = 1$. This requires R non-zero, which forces $1-n = n-1$. The only solution is $n = 1$, which contradicts $(1-n)R = 1$. Therefore there are no link invariants satisfying the generalised Kauffman properties. $\triangle$

Exercise 3.3.2. (Exercise to myself) Is there a map from link diagrams to $\mathbb{Z}[t^{1/2}, t^{-1/2}]$ which satisfies the skein relation and normalisation (characterising the Jones polynomial) which *isn't* a link invariant?

Solution: For any Laurent polynomial $P(A) \in \mathbb{Z}[A, A^{-1}]$ let B_P be a map from link diagrams to $\mathbb{Z}[A, A^{-1}]$ defined by

(a) $B_P(0_1) = 1$,

(b) $B_P(D \sqcup 0_1) = P(A) B_P(D)$,

(c) $B_P(\times) = A B_P()() + A^{-1} B_P(\asymp)$.

This uniquely determines a function on link diagrams by induction: suppose the function is defined for links with at most n crossings. Then given a link with $n+1$ crossings, (c) reduces the link to two copies with n crossings each.

Note that B_P is never invariant under both type I Reidemeister moves and its "dual move", by applications of (b) and (c). (Precisely, this follows from the argument in the previous exercise.) We can now define a modified Jones polynomial:

$$V_P(D) = \Big((-A)^{-3w(D)} B_P(D) \Big)_{t^{1/2} = A^{-2}} \in \mathbb{Z}[t^{1/2}, t^{-1/2}].$$

In general this is *not* a link invariant, although it is a well defined map on link diagrams. It remains to verify that V_P is normalised and satisfies the skein relation. Normalisation is immediate, and the skein relation follows from the fact that

$$w(L_+) - 1 = w(L_0) = w(L_-) + 1$$

and property (c). $\triangle$

Chapter 4

Alternating links

4.1 General properties

In this section we prove two important properties of *alternating links*, which essentially say that split-ness and primality can be read off alternating link diagrams. These motivate the idea that alternating diagrams are in some sense "minimal", which is pursued in the next section. Formally, the relevant definitions and theorems are as follows:

Definition 4.1.1. A link L is *alternating* if it admits an alternating link diagram D. That is, tracing around any component, the crossings alternate between underpasses and overpasses.

Definition 4.1.2. A link $L \subset \mathbb{S}^3$ is *split* if there exists $\mathbb{S}^2 \subset \mathbb{S}^3$ such that non-empty components of L lie in both components of $\mathbb{S}^3 - \mathbb{S}^2$, with $L \cap \mathbb{S}^2 = \varnothing$. A link diagram D is *split* if there is a simple closed curve γ in $\mathbb{R}^2$ such that non-empty components of D lie in both components of $\mathbb{R}^2 - \gamma$, with $D \cap \gamma = \varnothing$.

Definition 4.1.3. A link $L \subset \mathbb{S}^3$ (which is not the unknot) is *prime* if for any $\mathbb{S}^2 \subset \mathbb{S}^3$ intersecting L transversely at exactly two points, L intersects a component of $\mathbb{S}^3 - \mathbb{S}^2$ along an unknotted arc. A link diagram D is *prime* if for any simple closed curve γ in $\mathbb{R}^2$ intersecting D transversely at exactly two points, D intersects a component of $\mathbb{R}^2 - \gamma$ along a link diagram of a trivial ball-arc pair. D is said to be *strongly prime* if the above intersection is guaranteed to be an arc with no crossings.

The two main theorems are the following:

Theorem 4.1.4. *Let L be a link with an alternating diagram D. Then L is split if and only if D is split.*

Theorem 4.1.5. *Let L be a link with an alternating diagram D. Then L is prime if and only if D is prime.*

To prove these theorems we fix some notation. A link diagram D will be taken to be a subset of $\mathbb{S}^2$ (rather than $\mathbb{R}^2$). Then $D \subset \mathbb{S}^2 \subset \mathbb{S}^3$, and the corresponding link L can be isotoped so that it agrees with D everywhere except on small balls at the crossings of D. We place a ball centered at each crossing, and declare that L lies on the spheres bounding these balls (so as to "blow up" the crossings). These balls are called *bubbles*. $\mathbb{S}^2$ separates every bubble into a *Northern hemisphere* and *Southern hemisphere* by fixing an orientation.

We next define S_+ and S_- to be the spheres obtained from $\mathbb{S}^2$ by perturbing by the bubbles. Formally, define S_+ to be the union of "$\mathbb{S}^2$ − bubbles" with all of the Northern hemispheres, and S_- to be the union of "$\mathbb{S}^2$ − bubbles" with Southern hemispheres. We further define B_+ to be the unique component of $\mathbb{S}^3 - S_+$ which is disjoint from all the bubbles, and B_- to be the component of $\mathbb{S}^3 - S_+$ disjoint from the bubbles.

Finally we fix the notion of a surface in $\mathbb{S}^3$ in *general position* with respect to our data. Let F be a surface in $\mathbb{S}^3$ transverse to L. By perturbing F, it can be assumed to intersect L only on $\mathbb{S}^2$ (so that F never meets L on the surface of bubbles). Moreover, F can be taken to be transverse to S_+ and S_-. Then F meets S_+ and S_- in the union of disjoint simple closed curves.

Next we consider the intersection of F with a bubble. By the assumption that F meets L only along $\mathbb{S}^2$, it must be the case that F meets each bubble in a union of "saddles". The boundary of each saddle is comprised of four arcs (forming a circle), with two arcs in each hemisphere. We say that F as described is a surface in *general position*.

Definition 4.1.6. Suppose $F \subset \mathbb{S}^3$ is a surface in general position with respect to the data $(D, L, S_+, S_-, \mathbb{S}^2)$. We say that F is in *standard position* if:

(a) Each of $F \cap B_+$ and $F \cap B_-$ is a disjoint union of disks.

(b) No component of $F \cap S_+$ or $F \cap S_-$ meets a bubble in more than one arc.

(c) Each component of $F \cap S_+$ and $F \cap S_-$ meets some saddle or meets L.

Lemma 4.1.7. Let D be a non-split diagram for L. Suppose F is a 2-sphere separating some components of L. Then F can be replaced with another 2-sphere with the same property, but in standard position.

Proof. We give a proof outline. Suppose F is in general position. We first explain (a). The same argument applies for both B_+ and B_-. Consider the components of $F \cap S_+$: these are each simple closed curves. Of these, consider the ones that do not bound a disk in B_+. Choose the inner-most of these curves. This curve bounds a disk on S_+ which can be pushed to a disk in B_+. Since our chosen curve was inner-most, the interior of this disk is disjoint from F. Cutting along the disk, we decompose F into two disjoint spheres. At least one of these components separates L, and we have removed the curve. By induction, all curves in $F \cap S_+$ and $F \cap S_-$ that do not bound disks in $F \cap B_+$ and $F \cap B_-$ can be eliminated.

For (c), suppose a component C of $F \cap S_+$ or $F \cap S_-$ doesn't meet a saddle. Then C necessarily lies in $S' = \mathbb{S}^2 -$ bubbles. But F doesn't intersect D, so this forces C to not intersect D. In particular C cannot separate D by assumption.

On the other hand, C bounds two disks whose union is F, and these disks necessarily lie in $F \cap B_+$ and $F \cap B_-$. Since C doesn't separate D, F doesn't separate L. This is a contradiction, so C must meet a saddle. In particular, it meets a saddle *or* L.

Finally, we leave (b) for when I have more time! □

Lemma 4.1.8. Suppose L, with diagram D, is not a split link. Suppose F is a 2-sphere meeting L transversely at two points, separating $\mathbb{S}^3$ into two balls, neither of which intersects L in trivial ball-arc pairs. Then F can be replaced with a 2-sphere in standard position.

Proof. The proof of this lemma is analogous to the previous lemma. □

The first of these two lemmas is the key ingredient in the proof of alternating links being split if and only if alternating diagrams are split, and the second lemma is key for alternating links being prime if and only if alternating diagrams are prime. An important property of alternating links is that we can deduce the following: suppose F is in standard position, and D is alternating. Consider a component C of $F \cap S_+$, and give it an orientation. Suppose C enters the surface of a bubble, with the bubble on its left. Then the next time C enters the surface of a bubble, it must appear on the right, and so on, alternating. The one exception is that $F \cap S_+$ can intersect D. (For example, the pattern can be left, right, left, intersection with D, left, right,...)

Theorem 4.1.9. *Let L be a link with an alternating diagram D. Then L is split if and only if D is split.*

Proof. If D is split, it is immediate that L is split. Conversely, suppose L is split but D is not split. By an earlier lemma, there exists a 2-sphere F separating some components of L in standard position with (D, L, S_+, S_-). Choose an inner-most component C of $F \cap S_+$, i.e. a curve C which bounds a disk on $\mathbb{S}_+$ whose interior doesn't meet F. By property (c) (of standard position surfaces), C meets a saddle. To return back to the original position, C must meet at least two saddles, so in particular it meets at least one saddle on the left and one on the right.

Now consider the saddle that meets C on the left. There is another arc in F bounding the saddle, and by property (b), this arc belongs to a distinct component C' of $F \cap S_+$. On the other hand, C also meets a saddle on the right, which is also bounded by an arc from a component C'' of $F \cap S_+$. This shows that both sides of C intersect F, so C cannot be inner-most. This proves that in fact F cannot intersect S_+ at all. Similarly $F \cap S_-$ is empty. It follows that F is contained in B_+, B_-, or a bubble. In any case F cannot separate L.

This is a contradiction, so it must be the case that whenever L is split, so is D. □

Theorem 4.1.10. *Let L be a link with an alternating diagram D. Then L is prime if and only if D is prime.*

Proof. This result follows an analogous proof by making use of the second standard-position lemma. $\square$

4.2 Applications of the Jones polynomial

The main goal of this section is to prove that any (appropriately reduced) alternating diagram necessarily exhibits the crossing number of the corresponding link. This was a long standing conjecture - one of the *Tait conjectures* - and first resolved in 1987. (The resolution used the Jones polynomial, as we do now.) Another of the Tait conjectures we resolve here that the *writhe* of a reduced alternating diagram is an invariant of the knot. (More precisely, any two reduced alternating diagrams of a given knot have the same writhe.)

Definition 4.2.1. Suppose D is an alternating diagram. D is *reduced* if it has no *removable crossings*. That is, the diagram cannot be expressed as $A \times B$ (or its mirror image).

Next we wish to introduce a measure of redundancy for link diagrams. To this end we develop some notation to compare link diagrams.

Definition 4.2.2. Let D be a link diagram. Order the n crossings of D. A function $s : \{1, \ldots, n\} \to \{-1, 1\}$ is called a *state* of D. Given a state s, let sD denote the diagram obtained by replacing crossings as follows:

- If $s(i) = 1$, and $\times$ is the ith crossing, replace it with $)($.
- If $s(i) = -1$, and $\times$ is the ith crossing, replace it with $\asymp$.

For any s, $|sD|$ denotes the number of components of sD. (Note that sD is an unlink.)

Proposition 4.2.3. Let D be a link diagram. Then the Kauffman bracket is given by

$$\langle D \rangle = \sum_{s \text{ state}} A^{\sum_i s(i)} (-A^{-2} - A^2)^{|sD|-1}.$$

Proof. It suffices to show that the above formula satisfies the definition of the Kauffman bracket.

1. Let D denote 0_1. Then there is a unique state (the empty state), so $\sum_i s(i) = 0$. On the other hand, $sD = D$ has one component, so $|sD| - 1 = 0$. It follows that the expression on the right gives 1 as required.

2. Suppose $D' = D \sqcup 0_1$. Then all of the states are unchanged: the only difference on the right side is that $|sD'| = |sD| + 1$. Therefore the right side is multiplied by a factor of $-A^{-2} - A^2$, as required.

3. Finally observe that $\times$ has two states, giving the diagrams $)($ and $\asymp$. Suppose D' and D'' are obtained my modifying a crossing $\times$ in D to each of the un-crossings. Every state s' on D' extends to a state s on D, with $|s'D'| = |sD|$, and $\sum_i s'(i) = \sum_i s(i) - 1$. Analogous results hold for D''. Combining these, we have

$$\begin{aligned}\langle D \rangle &= \sum_{s \text{ state of } D} A^{\sum_i s(i)} (-A^{-2} - A^2)^{|sD|-1} \\ &= \sum_{s' \text{ state of } D'} A^{\sum_i s'(i)+1} (-A^{-2} - A^2)^{|s'D'|-1} \\ &\quad + \sum_{s'' \text{ state of } D''} A^{\sum_i s''(i)-1} (-A^{-2} - A^2)^{|s''D''|-1} \\ &= A\langle D' \rangle + A^{-1} \langle D'' \rangle.\end{aligned}$$

This completes the proof that the expression on the right agrees with the definition of the Kauffman bracket. □

Definition 4.2.4. We write s_+ and s_- to denote the constant states sending all crossings to 1 and -1 respectively. D is said to be *plus-adequate* if $|s_+D| > |sD|$ for all s with $\sum_i s(i) = n - 2$, and *minus-adequate* if $|s_-D| > |sD|$ for all s with $\sum_i s(i) = 2 - n$. D is called *adequate* if it is both plus and minus adequate.

The idea is that whenever two arcs replace a crossing (to from s_+D from D) we require that all of these pairs belong to distinct components. (Similarly for s_-D.)

Proposition 4.2.5. A reduced alternating link diagram is adequate.

Proof. Suppose D is a reduced alternating link diagram. Since D is alternating, giving it a chess-board colouring, s_+D consists of boundaries of black regions (wlog) with corners rounded off, and s_-D consists of boundaries of white regions.

Suppose for a contradiction that D is not plus-adequate, so changing some crossing of s_+D gives a diagram with at least as many components. Then this crossing is removable: consider a push-off S from the component formed by changing the given crossing of s_+D. This realises the removability of the crossing, since the alternating condition requires that S does not interact with any other crossings. □

Definition 4.2.6. Given a Laurent polynomial P, we write MP and mP to denote the maximum and minimum degrees of terms of P. The *breadth* of a polynomial, denoted BP, is defined by $MP - mP$.

Lemma 4.2.7. Let D be a link diagram with n crossings. Then

- $M\langle D \rangle \leq n + 2|s_+D| - 2$, with equality if D is plus-adequate.
- $m\langle D \rangle \geq -n - 2|s_-D| + 2$, with equality if D is minus-adequate.

Proof. For each state s, we write

$$\langle D|s\rangle = A^{\sum_i s(i)}(-A^{-2} - A^2)^{|sD|-1}.$$

By an earlier proposition, $\langle D\rangle = \sum_s \langle D|s\rangle$. Since $\sum_i s_+(i) = n$, $M\langle D|s_+\rangle = n+2|s_+D|-2$. Now consider any other state s. This is achieved from s_+ by changing one point at a time, i.e. we have a sequence $s_0, \ldots, s_k$ such that

$$s_0 = s_+, s_k = s, s_{r-1}(i) = s_r(i) \text{ for all but exactly one } i \in \{1, \ldots, n\}.$$

For each r, we have

$$\sum_i s_r(i) = n - 2r, \quad |s_r D| = |s_{r-1}D| \pm 1.$$

It follows that $M\langle D|s_{r-1}\rangle - M\langle D|s_r\rangle$ is either $2-2$ or $2+2$, i.e. 0 or 4. In particular, $M\langle D|s_{r-i}\rangle \geq M\langle D|s_r\rangle$, so by induction

$$M\langle D|s\rangle \leq n + 2|s_+D| - 2.$$

Since s was arbitrary, summing all of the above gives

$$M\langle D\rangle \leq n + 2|s_+D| - 2.$$

Next assume that D is plus-adequate. Then in the previous sequence, $|s_+D| = |s_1D| + 1$ is guaranteed, $M\langle D|s_1\rangle < M\langle D|s_+\rangle$. It follows that s_+ is the *only* state achieving degree $n+2|s_+D|-2$, so summing over all states cannot cancel the largest degree term. Therefore

$$M\langle D\rangle = n + 2|s_+D| - 2.$$

Next we inspect the second condition. Recall that the mirror image of a diagram corresponds to the substitution of A^{-1} for A in the Kauffman polynomial. Moreover, it corresponds to replacing each state s by reversing its image. It follows that the second statement is simply the first applied to $\overline{D}$. $\square$

Corollary 4.2.8. *If D is adequate, then*

$$B\langle D\rangle = M\langle D\rangle - m\langle D\rangle = 2n + 2|s_+D| + 2|s_-D| - 4.$$

To use this result, we need some further estimates on $|s_+D| + |s_-D|$.

Lemma 4.2.9. Let D be a non-split link diagram with n crossings. Then

$$|s_+D| + |s_-D| \leq n + 2.$$

Proof. This follows from induction on n. If $n = 0$, the inequality reads $2 \leq 2$, which is clearly true. Next suppose the result holds for diagrams with $n - 1$ crossings. Choose a crossing of D. Replacing the crossing with)(or $\asymp$, one of the outcomes D' must be connected (essentially by the Jordan curve theorem). Then without loss of generality, $s_+D' = s_+D$, and $|s_-D'| = |s_-D| \pm 1$. By the inductive hypothesis, it follows that

$$|s_+D| + |s_D| = |s_+D'| + |s_-D'| \mp 1 \leq (n-1) + 2 \mp 1 \leq n + 2.$$

$\square$

This result can tightened:

Lemma 4.2.10. Let D be a connected n-crossing diagram. If D is alternating, then $|s_+D| + |s_-D| = n + 2$.

Proof. Consider a chess-board colouring of the regions bound by D. Then $|s_+D|$ is the number of components of s_+D, but each component bounds a black region (wlog). Similarly $|s_-D|$ counts the number of white regions. Therefore $|s_+D| + |s_-D|$ is the number of faces of D.

On the other hand, recall that $V - E + F = 2$ by Euler's formula. Since D is 4-valent, $E = 2V$. Therefore

$$|s_+D| + |s_-D| = F = V + 2 = n + 2$$

as required. $\square$

Theorem 4.2.11. *Let D be a connected n-crossing diagram of an oriented link L. Then $B(V(L)) \leq n$. If D is alternating and reduced, then $B(V(L)) = n$.*

Proof. Under the substitution $t = A^{-4}$, we have

$$V(L) = (-A)^{-3w(D)} \langle D \rangle.$$

This gives

$$4B(V(L)) = B((-A)^{-3w(D)} \langle D \rangle) = B(\langle D \rangle) \leq 2n + 2|s_+D| + 2|s_-D| - 4$$

by lemma 4.2.7. Moreover, by lemma 4.2.9, this gives

$$4B(V(L)) \leq 2n + 2|s_+D| + 2|s_-D| - 4 \leq 2n + 2(n+2) - 4 = 4n.$$

It follows that $B(V(L)) \leq n$. Now suppose that D is alternating and reduced. Then D is adequate, so by corollary 4.2.8, the first inequality above is strict. Moreover, by lemma 4.2.10, the second inequality is also strict. This gives the second claim. $\square$

Corollary 4.2.12 (Tait's conjecture)**.** *Suppose L is a link that admits a reduced alternating diagram. Then this diagram exhibits the crossing number of L.*

Proof. Suppose L is split. Then the alternating diagram D is necessarily split by one of the main theorems from the previous section. Consider the connected components D_i of D. Again by the theorem from the previous section, the link L_i corresponding to D_i is not split, and hence admits only connected link diagrams.

By the previous theorem, $B(V(L_i)) = n_i$, where n_i is the number of crossings in the reduced alternating diagram D_i. But by the first part of the previous theorem, any diagram of L_i has at least $B(V(L_i))$ crossings. It follows that any diagram of L has at least $\sum_i n_i$ crossings. But D has exactly $\sum_i n_i$ crossings as required. $\square$

4.3 Exercises

Exercise 4.3.1. (Lickorish 4.2) Prove that the Whitehead link is not a split link.

Solution: Each component of the Whitehead link W is an unknot. Since any two link diagrams are related by Reidemeister moves, the two componenets of any diagram of the Whitehead link will always be unknots. In particular, if the Whitehead link was a split link, it would have a diagram $0_1 \sqcup 0_1$. Therefore

$$V(W) = -t^{1/2} - t^{-1/2}.$$

On the other hand, a straight forward calculation shows that the Jones polynomial of the Whitehead link is

$$V(W) = t^{-3/2}(-1 + t - 2t^2 + t^3 - 2t^4 + t^5).$$

$\triangle$

Exercise 4.3.2. (Lickorish 4.5) Show that $K_1 + K_2$ is an alternating knot if and only if K_1 and K_2 are alternating.

Solution: We first note that the connected sum of any two alternating knots is an alternating knot. Suppose K_1 and K_2 are alternating, and let D_1 and D_2 to be alternating diagrams for K_1 and K_2 respectively.

The "boundary" of the diagrams D_1 and D_2 are comprised of segments whose endpoints are under-crossings and over-crossings. Each segment inherits an orientation from the embedding in the plane, as well as an orientation from the crossings at the end points. We now sign each segment: $+$ if the two induced orientations agree, and $-$ if they disagree.

If at least one boundary segment from D_1 and D_2 have the same sign, then the naive connected sum between these segments gives an alternating diagram of $K_1 + K_2$. If every boundary segment of D_1 has positive sign and every boundary segment of D_2 has negative sign, we create a new alternating diagram D_2' as follows:

D_2 can be viewed as a graph in the plane. Then D_2 defines an "exterior" face F_1, and choosing any boundary segment of D_2, another face F_2. Choose push-offs C_1, C_2 of

∂F_1 and ∂F_2 disjoint from D_2. The curves C_i bound an annulus, in which D_2 is not embedded. An isotopy of the annulus in the 3-sphere is given by swapping the inner and outer boundaries. (Visually, this corresponds to rotating a solid torus by 180° along a meridian.) Projecting this back onto the plane, we obtain a new alternating diagram D_2', in which our chosen boundary segment is still a boundary segment, but has opposite sign. Therefore the connected sum of D_1 and D_2' via this segment gives an alternating diagram of $K_1 + K_2$.

This completes the proof of one direction. For the other direction, recall that a knot K admitting an alternating diagram is prime if and only if the alternating diagram is prime. With this in mind, suppose $K = K_1 + K_2$ is alternating, and let D be an alternating diagram of K. Since K isn't prime, D is not a prime diagram, so there exists a simple closed curve γ realising a non-trivial connected sum of D. Moreover, (considering chess board colourings for example), one can show that the connected summands are themselves alternating. Inductively, we conclude that D can be expressed as a connected sum of alternating prime knots. Since prime decompositions are unique, this gives an expression for both K_1 and K_2 as connected sums of alternating prime knots. By the first direction, K_1 and K_2 are alternating as required. △

Exercise 4.3.3. (Lickorish 5.3) Show that

$$c(K_1 + K_2) = c(K_1) + c(K_2)$$

for alternating knots K_1 and K_2.

Solution: Since K_1 and K_2 are alternating, there exist minimal crossing alternating diagrams D_1 and D_2 for K_1 and K_2 (by the main theorem of this chapter). In particular, D_1 and D_2 are reduced. By the proof from above, by isotoping the diagram D_2 if necessary, we can construct an alternating diagram D by taking a connected sum $D_1 \# D_2$ along a path. Moreover, since D_1 and D_2 are reduced, so is D. It follows that D exhibits the crossing number of $K_1 + K_2$ (by the main theorem of this chapter). That is,

$$c(K_1) + c(K_2) = c(D_1) + c(D_2) = c(D) = c(K_1 + K_2).$$

△

Chapter 5

The Alexander polynomial

5.1 Homological definition of the Alexander polynomial

The Alexander polynomial is the first known knot polynomial, and is well understood via homology theory. We begin by introducing relevant terminology, and then define the polynomial.

Definition 5.1.1. Let F, E, M be R-modules, and suppose F and E are free with finite bases $\{f_i\}$ and $\{e_j\}$. A short exact sequence

$$F \xrightarrow{\alpha} E \xrightarrow{\varphi} M \to 0$$

is called a *finite presentation* of M. Given our choice of bases for E and F, let A be a matrix representing α, i.e.

$$\alpha f_i = \sum_j A_{ij} e_j.$$

Then A is called a *presentation matrix* for M. We think of $\{e_j\}$ as a basis for M, and $\{f_i\}$ as relations.

Theorem 5.1.2. *Any two presentation matrices A_1 and A_2 differ by a sequence of the following matrix moves and their inverses:*

- *Standard row/column operations:*
  - *Adding a scalar multiple of one row/column to another row/column.*
  - *Permuting two rows or two columns.*
- *Addition of an extra column of zeros.*
- *Replacement of A with $\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}$.*

Definition 5.1.3. Let R be a commutative ring, and M a module over R with $m \times n$ presentation matrix A. The *rth elementary ideal* $\mathcal{E}_r$ is the ideal of R generated by all of the $(m-r+1) \times (m-r+1)$ minors of A. By convention, for $r \leq 0$, we declare $\mathcal{E}_r = 0$. For $r > m$, we declare $\mathcal{E}_r = R$.

Recall that an $a \times a$ *minor* of A is the determinant of an $a \times a$ submatrix of A. Therefore the rth elementary ideal is generated by determinants of all sub-matrices after removing $r-1$ rows and $n-m+r-1$ columns.

Proposition 5.1.4. Elementary ideals satisfy the following properties:

- For each r, $\mathcal{E}_r$ is independent of the choice of presentation matrix. (This follows from elementary properties of determinants together with the previous theorem classifying presentation matrices for a given R-module M.)
- For each r, $\mathcal{E}_{r-1} \subset \mathcal{E}_r$.
- If A is an $m \times m$ (square) matrix, there is a unique $m \times m$ minor, namely $\det A$. Therefore $\mathcal{E}_1 = \langle \det A \rangle$.

We now develop the relevant homology theory in order to define the Alexander polynomial. This first proposition is an instance of Alexander duality:

Proposition 5.1.5. Let $\Sigma \subset \mathbb{S}^3$ be a connected, compact, orientable surface with non-empty boundary. Then $H_1(\mathbb{S}^3 - \Sigma; \mathbb{Z})$ and $H_1(\Sigma; \mathbb{Z})$ are isomorphic. Moreover, there is a unique non-degenerate bilinear form

$$\beta : H_1(\mathbb{S}^3 - \Sigma; \mathbb{Z}) \times H_1(\Sigma; \mathbb{Z}) \to \mathbb{Z}$$

satisfying $\beta([c], [d]) = \mathrm{lk}(c, d)$ for any oriented simple closed curves c, d in $\mathbb{S}^3 - F$ and F respectively.

Note that any Seifert surface satisfies the premises of F in the above proposition. With this in mind, we next aim to define the *Seifert form*.

Suppose F is a Seifert surface for a link L. Let N be a tubular neighbourhood of L. Let X denote the closure of $\mathbb{S}^3 - N$. Then $X \cap F$ is just a copy of F with a collar around the boundary of F removed, so we can identify $X \cap F$ with F. F then has a tubular neighbourhood $F \times [-1, 1]$ in X, so that every meridian of a component of L enters at $F \times \{-1\}$ and leaves from $F \times \{1\}$. Define embeddings $i^{\pm} : F \to \mathbb{S}^3 - F$ by $i^{\pm}(x) = x \times \{\pm 1\}$.

Definition 5.1.6. Let F be a Seifert surface of an (oriented) link L. The *Seifert form* is defined to be

$$\alpha : H_1(F; \mathbb{Z}) \times H_1(F; \mathbb{Z}) \to \mathbb{Z}, \quad \alpha : (x, y) \mapsto ({i^-}_*(x), y).$$

By the previous proposition, the Seifert form is well defined, and $\alpha([a,b]) = \mathrm{lk}(a^-, b)$ where $a^- = i^- a$.

Since F is an oriented surface, $H_1(F;\mathbb{Z})$ is freely generated by $2g+n-1$ simple closed curves, which we denote by $\{f_i\}$. That is,

$$H_1(F;\mathbb{Z}) = \langle [f_1], \ldots, [f_{2g+n-1}] \rangle = \bigoplus_{i=1}^{2g+n-1} \mathbb{Z}.$$

Since $\beta : H_1(\mathbb{S}^3 - F;\mathbb{Z}) \times H_1(F;\mathbb{Z}) \to \mathbb{Z}$ is non-degenerate, it defines a dual basis for $H_1(\mathbb{S}^3 - F;\mathbb{Z})$. That is, we can choose homology classes $\{[e_i]\}$ in $H_1(\mathbb{S}^3 - F;\mathbb{Z})$ so that $\beta([e_i],[f_j]) = \mathrm{lk}(e_i, f_j) = \delta_{ij}$.

Definition 5.1.7. Let $\{[f_i]\}$ be a basis for $H_1(F;\mathbb{Z})$. The matrix A representing α in this basis is called a *Seifert matrix* of F.

Observe that A satisfies

$$A_{ij} = \alpha([f_i],[f_j]) = \mathrm{lk}(f_i^-, f_j) = \mathrm{lk}(f_i, f_j^+).$$

Now let $\{[e_i]\}$ be the β-dual basis of $\{[f_i]\}$ for $H_1(\mathbb{S}^3 - F;\mathbb{Z})$. Then $\mathrm{lk}(e_i, f_j) = \delta_{ij}$, so in $\mathbb{S}^3 - F$ the above immediately gives

$$[f_i^-] = \sum_j A_{ij}[e_j], \quad [f_j^+] = \sum_i A_{ij}[e_i].$$

(These can be verified by substitution into the above expression, noting that β is non-degenerate.)

Definition 5.1.8. Let X be the link complement of L in $\mathbb{S}^3$ as above. The *infinite cyclic cover* X_∞ of X is constructed as follows:

1. Let Y be the cut manifold (with boundary) of X obtained by cutting along F (where F is viewed as a submanifold of X).
2. Note that Y is a cobordism, with boundary F^- and F^+ (where the signs are determined by the orientation of F).
3. Let X_∞ be the manifold obtained by gluing infinitely many copies of (Y_i, F_i^+, F_i^-) end to end, so that Y_i and Y_{i+1} are glued together along F_i^- and F_{i+1}^+.

Note that the orientations are chosen so that if μ was a meridian of a component of L, then the orientation of μ agrees with the order of the indices Y_i.

$X_\infty \to X$ has a canonical covering map (which mirrors that of $\mathbb{R} \to \mathbb{S}^1$.)

Proposition 5.1.9. $H_1(X_\infty, \mathbb{Z})$ has a canonical module structure over $\mathbb{Z}[t, t^{-1}]$.

Proof. We first note that X_∞ is obtained by $\cdots \sqcup_F Y_i \sqcup_F Y_{i+1} \sqcup_F \cdots$, so there is an automorphism $t \in \mathrm{Aut}(X_\infty)$ defined by translation by 1 to the right. The infinite cyclic group $\langle t \rangle$ is now a group of automorphisms, in fact, Deck transformations, of X_∞. These descend to automorphisms of $H_1(X_\infty;\mathbb{Z})$. On the other hand, $H_1(X_\infty;\mathbb{Z})$ is an abelian group, so it is naturally a $\mathbb{Z}$-module. Therefore $H_1(X_\infty;\mathbb{Z})$ is a $\mathbb{Z}\langle t \rangle = \mathbb{Z}[t,t^{-1}]$ module. □

Theorem 5.1.10. *Let F be a Seifert surface for an oriented link L, and let A be a Seifert matrix of the corresponding Seifert form (in any basis for $H_1(F;\mathbb{Z})$). Then $tA - A^T$ is a matrix presenting $H_1(X_\infty;\mathbb{Z})$ as a $\mathbb{Z}[t,t^{-1}]$ module. (In particular, $H_1(X_\infty;\mathbb{Z})$ is a finitely presented $\mathbb{Z}[t,t^{-1}]$-module.)*

Proof. The proof will be omitted, but the idea is that if F has free basis $\{1 \otimes [f_i]\}$ and E has free basis $\{[e_i] \otimes 1\}$ as $\mathbb{Z}[t,t^{-1}]$-modules, then $tA - A^T$ presents α_* the short exact sequence

$$F \xrightarrow{\alpha_*} E \to H_1(X_\infty;\mathbb{Z}).$$

□

We are finally ready to define the Alexander polynomial and related invariants.

Theorem 5.1.11. *The $\mathbb{Z}[t,t^{-1}]$-module $H_1(X_\infty;\mathbb{Z})$ is an invariant of oriented links. It is called the* Alexander invariant *or* Alexander module.

Recall from an earlier proposition that elementary ideals of a finitely presented R-module are invariants of the module. Therefore the Alexander module gives further invariants:

Definition 5.1.12. The *rth Alexander ideal* is the rth elementary ideal of $H_1(X_\infty;\mathbb{Z})$ as a $\mathbb{Z}[t,t^{-1}]$-module. The *rth Alexander polynomial* is a generator of the smallest principal ideal of $\mathbb{Z}[t,t^{-1}]$ that contains the rth Alexander ideal. The first Alexander polynomial, denoted $\Delta_L(t)$, is called *the* Alexander polynomial.

Note that we have only defined Alexander polynomials up to multiplication by a unit; $\pm t^k$. In the next section we see that the Alexander polynomial can be calculated via Skein relations, if we fix a normalisation.

5.2 Alternative definitions of the Alexander polynomial

In this section we explore three more methods of computing the Alexander polynomial. The first two methods are useful given a link diagram. In particular, the second method relates the Alexander polynomial to the HOMFLY polynomial, and so we also have a connection with the Jones polynomial. The third method allows us to compute the Alexander polynomial of a knot given a presentation of its knot group.

First alternative approach: diagram incidence matrix.

Let D be an oriented diagram of a link L. Suppose D has n crossings, $v_0, \ldots, v_{n-1}$. By Euler's formula $V - E + F = 2$, noting that a link diagram is 4-regular, we have

$$F = 2 - V + E = 2 - V + 2V = V + 2 = n + 2.$$

That is, the diagram partitions the plane into $n+2$ regions, which we label $r_0, \ldots, r_{n+1}$. We now construct an $n \times (n+2)$ incidence matrix for the vertices and faces:

$$M_{ij} = \begin{cases} 0 & r_j \text{ doesn't meet } v_i \\ -t & r_j \text{ meets } v_i \text{ on the left, before undercrossing} \\ 1 & r_j \text{ meets } v_i \text{ on the right, before undercrossing} \\ t & r_j \text{ meets } v_i \text{ on the left, after undercrossing} \\ -1 & r_j \text{ meets } v_i \text{ on the right, after undercrossing.} \end{cases}$$

M is not square, so we cannot compute its determinant. However, once can remove two columns from the matrix corresponding to adjacent regions to obtain a square matrix $\widetilde{A}$. Then $\det \widetilde{A}$ is the *Alexander polynomial*.

Remark. The Alexander polynomial constructed above is only determined up to a unit. Removing different choices of columns from M will scale $\det \widetilde{A}$ by $\pm t^k$.

The matrix $\widetilde{A}$ is called the *Alexander matrix*, and is equivalent to $tA - A^T$ from the previous section, where A is the Seifert matrix of L. Therefore $\widetilde{A}$ captures the data of the Alexander module.

Second alternative approach: skein relations.

In the chapter on the Jones polynomial, we observed that a skein relation together with normalisation determines the Jones polynomial uniquely. Similarly, a certain normalisation of the Alexander polynomial (called the Conway polynomial) is determined uniquely by a skein relation:

Definition 5.2.1. The *Conway polynomial* is the unique link invariant

$$\nabla : \{ \text{ oriented links in } \mathbb{S}^3\} \to \mathbb{Z}[z, z^{-1}]$$

such that

- $\nabla_{0_1}(z) = 1$,
- $\nabla_{L_+}(z) - \nabla_{L_-}(z) = z\nabla_{L_0}(z)$.

Theorem 5.2.2. *Let A be a Seifert matrix for L. Then $\det(t^{1/2}A - t^{-1/2}A^T)$ is a well defined invariant of L, in that there is no ambiguity of units. Moreover,*

1. $\nabla_L(t^{-1/2} - t^{1/2}) = \det(t^{1/2}A - t^{-1/2}A^T)$, *so* $\det(t^{1/2}A - t^{-1/2}A^T)$ *is obtained by reparametrising the Conway polynomial.*

2. If A is an $r \times r$ matrix, then

$$t^{-r/2}\Delta_L(t) = t^{-r/2}\det(tA - A^T) = \det(t^{1/2}A - t^{-1/2}A^T),$$

so this invariant is obtained by specifying a normalisation of the Alexander polynomial.

In particular, the Conway polynomial is a normalisation of the Alexander polynomial.

Remark. The HOMLFY polynomial $P(\ell, m)$ was defined to be the unique oriented link invariant satisfying

- $P(0_1) = 1$,
- $\ell P(L_+) + \ell^{-1}P(L_-) + mP(L_0) = 0$.

Therefore the Conway polynomial is obtained by a substitution into the HOMFLY polynomial:

$$\nabla(z) = -iP(i, -iz).$$

Similarly, the Jones polynomial is obtained by substitution into the HOMFLY polynomial:

$$V(L)(t) = -iP(it^{-1}, i(t^{-1/2} - t^{1/2})).$$

Third alternative approach: group presentations.

Definition 5.2.3. Let K be a knot in $\mathbb{S}^3$. Then its *knot group* is $\pi_1(\mathbb{S}^3 - K)$.

We find that knot groups admit balanced presentations (called *Wirtinger presentations*) that can be readily computed from knot diagrams. We now describe this procedure.

1. Let D be an oriented diagram of a knot K. If D has n crossings, D is composed of n arcs. (At each crossing, we consider an incoming arc to be the same as the outgoing arc if the arc follows an overpass. We consider the incoming and outgoing arcs to be distinct if the arc follows an underpass.) Choose a basepoint of D, and label each arc: $x_1, \ldots, x_n$.

2. The symbols $x_1, \ldots, x_n$ are *generators* of the knot group. The relations are given by each crossing. Suppose x_i, x_j, x_k meet at a crossing, with x_i passing over, x_j incoming, and x_k outgoing, so that the crossing is positively oriented. Then we add a relation

$$x_i x_j = x_k x_i.$$

If the crossing is negatively oriented, we add the relation

$$x_j x_i = x_i x_k.$$

This gives a presentation

$$\pi_1(\mathbb{S}^3 - K) = \langle x_1, \ldots, x_n | r_1, \ldots, r_n \rangle.$$

Observe that any one relation is determined by the other $n-1$ relations. Similarly, any one generator can be removed by Tietze transformations.

Definition 5.2.4. The *Fox derivative* is a notion of differentiation for words in a free group. Explicitly, for each x_i in $F(x_j)$, the fox derivative is a map $\partial_{x_i} : F(x_j) \to \mathbb{Z}[F(x_j)]$ defined by

- $\partial_{x_i} 1 = 0$
- $\partial_{x_i} x_j = \delta_{ij}$
- $\partial_{x_i} x_j^{-1} = -\delta_{ij} x_j^{-1}$
- $\partial_{x_i} w x_j = \partial_{x_i} w + w\delta_{ij}$.

Using this notion of a derivative, we can readily compute the *Jacobian* of a presentation P,

$$J(P)_{ij} = \partial_{x_j} r_i = \begin{pmatrix} \partial_{x_1} r_1 & \cdots & \partial_{x_n} r_1 \\ \vdots & \ddots & \vdots \\ \partial_{x_1} r_n & \cdots & \partial_{x_n} r_n \end{pmatrix}.$$

There is a unique map $T : \mathbb{Z}[F(x_j)] \to \mathbb{Z}[t, t^{-1}]$ defined by sending each x_i to t. This map naturally extends to matrices, and in particular we denote by J_T the Jacobian with each x_i replaced with t.

Theorem 5.2.5. *Let P, Q be finite presentations of a group G. Then $J_T(P)$ and $J_T(Q)$ are* equivalent *in that they differ by a sequence of the following moves:*

- *Standard row/column operations (see the start of the chapter).*
- *Addition of an extra row of zeros.*
- *Replacement of A with $\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}$.*
- *Multiplication of any row or column by $\pm t^k$.*

It follows that the determinant of $J_T(P)$ is a well defined invariant of G, up to multiplication by a unit.

Theorem 5.2.6. *Let P be a presentation of the knot group $\pi_1(\mathbb{S}^3 - K)$. Then the rth elementary ideals of $J_T(P)$ are the rth Alexander ideals. In particular, the Alexander polynomial is the generator of the smallest principal ideal containing the minors of the largest square blocks of $J_T(P)$.*

Earlier we remarked that the Wirtinger presentation of a knot group is redundant in a symmetric way (in that we can remove any relation and any generator). This is realised in the calculation of the Alexander polynomial using a Wirtinger presentation P: one can remove any row and column from $J_T(P)$, and the determinant of the remaining square matrix is the Alexander polynomial.

5.3 Properties of the Alexander polynomial

One of the main uses of the Alexander polynomial is in the study of the genus of links. Some properties relating the Alexander polynomial to genus will be explored in this section, along with other general properties.

Theorem 5.3.1. *Let L be an oriented link.*

- *$\Delta_L(t)$ is associate to $\Delta_L(t^{-1})$.*
- *If L is a knot, $\Delta_L(1) = \pm 1$. If L has more than one component, $\Delta_L(1) = 0$.*
- *$\Delta_L(t)$ is associate to $\Delta_{\overline{L}}(t)$ and $\Delta_{rL}(t)$.*
- *For knots K_1 and K_2, $\Delta_{K_1+K_2}(t)$ is associate to $\Delta_{K_1}(t)\Delta_{K_2}(t)$.*
- *If L is split, then $\Delta_L(t) = 0$.*

Proof. The first claim follows from the formula $\Delta_L(t) = \det(tA - A^T)$, and the properties of the determinant and how it interacts with transposition.

The second claim comes from realising that $\Delta_L(t) = \det(A - A^T)$, and the ijth entry of $A - A^T$ counts the algebraic intersections of f_i and f_j on F.

The third claim comes from the fact that if A is a Seifert matrix for L, then $-A$ is a Seifert matrix for $\overline{L}$, and A^T is a Seifert matrix for rL.

The fourth claim comes from the fact that $\mathrm{diag}(A_1, A_2)$ is a Seifert matrix for $K_1 + K_2$, if A_1 and A_2 are Seifert matrices for K_1 and K_2 respectively.

FInally, the fifth claim is most easily seen by using the characterisation of the Alexander polynomial via skein relations. Taking the split link to be $L_1 \sqcup L_2 = L_0$, let L_+ and L_- be obtained by "cross-gluing" components in each of L_1 and L_2 with the correct orientation. That L_+ and L_- are diagrams of the same link, so $\nabla_{L_0}(t) = \nabla_{L_+}(t) - \nabla_{L_-}(t) = 0$. □

As a corollary, a certain normal form of the Alexander polynomial of knots is

$$\Delta_L(t) = a_0 + a_1(t^{-1} + t) + a_2(t^{-2} + t^2) + \cdots,$$

where each a_i is an integer, and a_0 is odd. This was used by Alexander (before Conway introduced the Conway normalisation).

Corollary 5.3.2. *The Alexander polynomial doesn't distinguish the granny knot from the square knot.*

Theorem 5.3.3. *Let L be a link with c components. Let g be the genus of L. Then*

$$2g + c - 1 \geq \text{ breadth } \Delta_L(t).$$

In particular, for a knot K, the genus is bounded below by half the breadth of Δ_K.

Proof. Let F be a Seifert surface of L with genus g. Then $H_1(F;\mathbb{Z})$ has $2g+c-1$ generators, and hence its Seifert matrix A is a $(2g+c-1)\times(2g+c-1)$ matrix. It follows that $\det(tA - A^T)$ is a *polynomial* with degree at most $2g+c-1$. □

Theorem 5.3.4. *Let K, C be knots. Suppose K is embedded in a solid torus T, and $e : T \to \mathbb{S}^3$ is an embedding so that $e(T)$ is a regular neighbourhood of C. Then $e(K)$ is the* satellite knot *with pattern K and companion C. Choose $n > 0$ so that K represents n times a generator of $H_1(T;\mathbb{Z})$. Then*

$$\Delta_{e(K)}(t) = \Delta_K(t)\Delta_C(t^n).$$

Equality is up to multiplication by a unit, but genuine equality holds when Δ is given the Conway normalisation.

Corollary 5.3.5. *A Whitehead double of any knot has Alexander polynomial equal to 1 (up to multiplication by units). (Thus the Alexander polynomial cannot distinguish satellite knots from the unknot.)*

Proof. A *Whitehead double* is a satellite knot with pattern K as in figure 5.1. Let C be

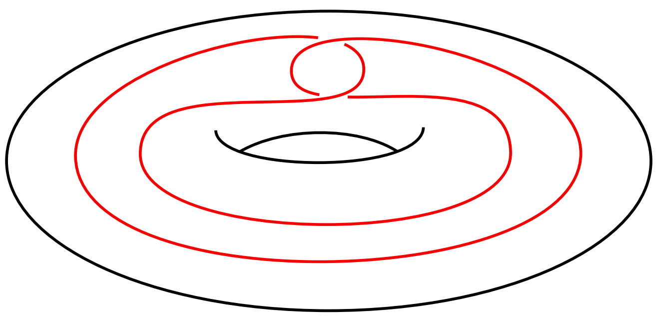


Figure 5.1: Pattern for the Whitehead double.

the *companion knot*, so that the Whitehead double of C is $e(K)$. By the previous theorem, write

$$\Delta_{e(K)}(t) = \Delta_K(t)\Delta_C(t^n).$$

Since K represents the trivial first homology class in $H_1(T)$, $n = 0$, so $\Delta_C(t^n) = \Delta_C(1) = \pm 1$. On the other hand, K is the unknot, so $\Delta_K(t)$ is a unit. It follows that $\Delta_{e(K)}$ is itself a unit. □

Theorem 5.3.6. *The Deck transformation $t : X_\infty \to X_\infty$ induces a map $t_* : H_1(X_\infty;\mathbb{Q}) \to H_1(X_\infty;\mathbb{Q})$. ($H_1(X_\infty;\mathbb{Q})$ is an infinite dimensional vector space.) Then the Alexander polynomial of the link (whose link complement has infinite cyclic cover X_∞) is the characteristic polynomial of t_*.*

One final important property is that the Alexander module also bounds unknotting numbers.

Theorem 5.3.7. *If the rth elementary ideal of the Alexander module of a knot K is properly contained in $\mathbb{Z}[t, t^{-1}]$, then the unknotting number of K is bounded below by r:*

$$u(K) \geq r.$$

5.4 Exercises

Exercise 5.4.1. (Exercise to myself) Compute the Alexander polynomial of the trefoil knot using the four different approaches outlined in this chapter.

Solution: 1. First we use the homological approach. The Seifert surface F obtained from the usual diagram of the trefoil knot has minimal genus 1. Therefore generators for $H_1(F;\mathbb{Z})$ are given by two simple closed curves f_1, f_2 on F. Choosing orientations, for one example pair we have

$$A_{ij} = \mathrm{lk}(f_i, f_j^+) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Therefore $\det(tA - A^T) = t^2 - t + 1$.

2. Next we use the incidence matrix approach. Consider a standard diagram of the trefoil knot, so that it has five regions. We order the regions by middle, leaf, leaf, leaf, exterior. Then the incidence matrix is a 3×5 matrix. To compute the Alexander polynomial, we can drop two adjacent regions, so our incidence matrix will be a 3×3 matrix with columns: *middle, leaf, leaf.* This gives

$$\widetilde{A} = \begin{pmatrix} 1 & -t & 0 \\ 1 & 0 & -1 \\ 1 & -1 & -t \end{pmatrix}.$$

The determinant of this polynomial is $-t^2 + t - 1$.

3. Now we use the skein relation to inductively determine the Conway polynomial of the trefoil. First we conclude from the skein relations that the unlink U of two components has vanishing Conway polynomial. This is proven more generally in the previous section for any split link.

But now if 3_1 denotes the trefoil knot, H a Hopf link, U the unlink with two components, and 0_1 the unknot, the skein relations give

$$\nabla_{3_1}(z) - \nabla_{0_1}(t) = z\nabla_H(z), \quad \nabla_H(z) - \nabla_U(z) = z\nabla_{0_1}(z).$$

It follows that $\nabla_{3_1}(z) = z^2 + 1$. By substitution of $z = t^{-1/2} - t^{1/2}$, this gives

$$\Delta_{3_1}(t) = \nabla_{3_1}(t^{-1/2} - t^{1/2}) = t - 1 - t^{-1}.$$

4. Finally we use the approach with Fox derivatives. A Wirtinger presentation of the knot group of the trefoil is given by

$$\langle x_0, x_1, x_2 \mid x_0x_1 = x_2x_0, x_1x_2 = x_0x_1, x_2x_0 = x_1x_2\rangle.$$

This can be vastly simplified using Tietze transformations, through which we obtain

$$\pi_1(\mathbb{S}^3 - 3_1) = \langle x, y \mid x^{-1}y^{-1}x^{-1}yxy\rangle.$$

The Jacobian of this presentation is a 1×2 matrix, specifically

$$J = (-x^{-1} - x^{-1}y^{-1}x^{-1} + x^{-1}y^{-1}x^{-1}y, x^{-1}y^{-1}x^{-1}yx + x^{-1}y^{-1}x^{-1} - x^{-1}y^{-1}).$$

Hence $J_T = (-t^{-1} - t^{-3} + t^{-2}, t^{-1} + t^{-3} - t^2)$. The first Alexander ideal is therefore $\langle t^{-1} + t^{-3} - t^2\rangle$, so the Alexander polynomial is $t^{-1} - t^2 + t^{-3}$.

In summary, our four methods have obtained:

$$t^2 - t + 1, \quad -t^2 + t - 1, \quad t - 1 - t^{-1}, \quad t^{-1} - t^2 + t^{-3}.$$

These all differ by units, as required. △

Exercise 5.4.2. (Lickorish 6.4) Show that for a knot K, $\Delta_K(t) = 1$ if and only if $H_1(X_\infty; \mathbb{Z}) = 0$.

Solution: If $H_1(X_\infty; \mathbb{Z}) = 0$, it is presented by the 1×1 unit matrix. The determinant is 1.

Conversely, suppose $H_1(X_\infty; \mathbb{Z}) \neq 0$. In particular, K cannot be the unknot, so K has genus at least 1. Therefore K has a Seifert matrix A of size $2g \times 2g$, with $g \geq 1$. A presentation matrix of $H_1(X_\infty; \mathbb{Z})$ is now given by the square matrix $tA - A^T$, so $\det(tA - A^T) = \Delta_K(t)$. Suppose for a contradiction that $\det(tA - A^T) = 1$. Consider the presentation below:

$$F \xrightarrow{\alpha} E \xrightarrow{\beta} H_1(X_\infty; \mathbb{Z}) \to 0.$$

$tA - A^T$ is a matrix representing α. Since $tA - A^T$ is invertible, α is surjective. Therefore by exactness $\ker \beta = E$. But β cannot be the zero map, since $H_1(X_\infty; \mathbb{Z}) \neq 0$. △

Exercise 5.4.3. (Lickorish 6.5) Which polynomials can arise as the Alexander polynomial of a genus 1 knot?

Solution: Let A be a Seifert matrix for a genus 1 Seifert surface. Then A is a 2×2 matrix, and

$$(A - A^T)_{ij} = \mathrm{lk}(f_i^-, f_j) - \mathrm{lk}(f_j^-, f_i) = \mathrm{lk}(f_i^-, f_j) - \mathrm{lk}(f_i^+, f_j).$$

The expression on the right gives the algebraic intersection number of f_i and f_j, so

$$A - A^T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Therefore A is of the form

$$A = \begin{pmatrix} a & b \\ b-1 & c \end{pmatrix}, \quad a, b, c \in \mathbb{Z}.$$

We now verify that any such matrix arises as the Seifert matrix for some knot. To this end, consider the Pretzel knot $P(q, r, s)$ where q, r, s are odd integers. Then $P(q, r, s)$ has a canonical Seifert surface F, obtained by attaching three strips (with q, r, and s half-twists) between two disks. Let f_1 be the anticlockwise loop in F passing through the strips with q and r twists, and let f_2 be the clockwise loop in F passing through the strips with r and s twists. The corresponding Seifert matrix is

$$A' = \frac{1}{2}\begin{pmatrix} q+r & r+1 \\ r-1 & r+s \end{pmatrix}.$$

Therefore given a matrix A with entries $a, b, b-1$, and c as above, this is realised as the Seifert matrix of the pretzel knot

$$P(2a-2b+1, 2b-1, 2c-2b+1).$$

We now have

$$\begin{aligned} \det(tA - A^T) &= \det \begin{pmatrix} a(t-1) & tb-b+1 \\ tb-b+t & c(t-1) \end{pmatrix} \\ &= ac(t^2-2t+1) - (b^2t^2 + bt^2 + t - 2b^2t + 2bt + b^2 - b) \\ &= (ac - b^2 + b)t^2 + (-2ac - 1 + 2b^2 - 2b)t + (ac - b^2 + b). \end{aligned}$$

Considering Alexander polynomials in the normal form

$$\Delta_K(t) = a_0 + a_1(t + t^{-1}) + a_2(t^2 + t^{-2}) + \cdots,$$

the above gives $a_0 = 2ac - 2b^2 + 2b + 1$, $a_1 = b^2 - ac - b$, and $a_i = 0$ for $i > 1$.

We now observe that $ac = b^2 - b - a_1$, so by substitution, $a_0 = -2a_1 + 1$. Therefore all possible Alexander polynomials of genus 1 knots (up to multiplication by a unit) are given by

$$\Delta_K(t) = a_1(t + t^{-1}) - 2a_1 + 1$$

where a_1 is an unconstrained integer. △

Chapter 6

Further applications of Seifert surfaces

6.1 Conway polynomial

In the previous section, we took for granted the fact that the *Conway normalisation* of the Alexander polynomial is well defined. We can understand why this is the case (without reference to skein relations) by studying how Seifert surfaces of a given link relate to each other.

Theorem 6.1.1. *Suppose F_1 and F_2 are Seifert surfaces of an oriented link L. Then there is a sequence*

$$F_1 = \Sigma_1, \Sigma_2, \ldots, \Sigma_N = F_2$$

of Seifert surfaces of L, such that for each i, Σ_i is obtained from Σ_{i-1} by surgery along an arc (or vice versa), or they are related by an isotopy of $\mathbb{S}^3$.

We don't prove this, but we give a description of what it means for an oriented surface Σ to be obtained from Σ' by surgery along an arc.

1. Let $\Sigma \subset \mathbb{S}^3$ be an oriented surface. Let $\gamma : [0,1] \to \mathbb{S}^3$ be an arc meeting Σ transversely, exactly at $\gamma(0)$ and $\gamma(1)$, so that $\gamma(\varepsilon)$ and $\gamma(1-\varepsilon)$ are on the same side of Σ.

2. Let $D^1 \times 0$ denote the image of γ (so that $\gamma(t) = t \times 0$, $D^1 = [0,1]$). Consider a tubular neighbourhood $D^1 \times D^2$ of $D^1 \times 0$, which meets Σ exactly at $\mathbb{S}^0 \times D^2$.

3. Define $\Sigma' = (\Sigma - \mathbb{S}^0 \times D^2) \sqcup_{\mathbb{S}^0 \times \mathbb{S}^1} (D^1 \times \mathbb{S}^1)$. Then Σ' is obtained from Σ by surgery along the arc γ.

The condition that $\gamma(\varepsilon)$ and $\gamma(1-\varepsilon)$ are on the same side of Σ is required to ensure that the surface obtained after the surgery is itself oriented.

Next we investigate how the corresponding Seifert matrices are related.

Definition 6.1.2. Let A be a square matrix over $\mathbb{Z}$. An *elementary enlargement* of A is a matrix B of either of the following two forms:

$$\begin{pmatrix} A & 0 & 0 \\ \eta^T & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} A & \xi & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

In the above, η, ξ are arbitrary vectors. Similarly obtaining A from either of the two matrices above is called an *elementary reduction.*

Matrices A, A' over $\mathbb{Z}$ are called *S-equivalent* if they are related by a sequence of elementary enlargements, elementary reductions, and conjugation by unimodular matrices.

Theorem 6.1.3. *Any two Seifert matrices of an oriented link L are S-equivalent.*

Corollary 6.1.4. *The Conway normalisation of the Alexander polynomial is well defined.*

Recall that the *Conway normalisation* refers to the definition

$$\Delta_L(t) = \det(t^{1/2}A - t^{-1/2}A^T)$$

for a Seifert matrix A of L. This is related to the original definition of the Alexander polynomial by multiplication by units in $\mathbb{Z}[t^{1/2}, t^{-1/2}]$.

Proof. Given that the Alexander polynomial is well defined up to multiplication by units, it remains to show that modifying the underlying Seifert surface doesn't affect the Conway normalisation. In other words, if $\Delta(t) = \det(t^{1/2}A - t^{-1/2}A^T)$, then $\Delta(t)$ is invariant under changing A by an S-equivalence. It's immediate that $\Delta(t)$ is invariant under conjugation of A by unimodular matrices. Similarly one can show that $\Delta(t)$ is invariant under elementary enlargements and elementary reductions of A. $\square$

Theorem 6.1.5. *The Conway polynomial $\nabla_L(z)$ defined by $\Delta_L(t) = \nabla_L(t^{-1/2} - t^{1/2})$ enjoys the following properties, for an oriented link L with c components.*

1. *∇_L is characterised by the following normalisation and skein relation:*
   - $\nabla(0_1) = 1$,
   - $\nabla_{L_+} - \nabla_{L_-} = z\nabla_{L_0}$.
2. *If L is split, $\nabla_L(z) = 0$.*
3. *Write $\nabla_L(z) = \sum_{i=0}^{\infty} a_i(L) z^i$. Then*

(a) $a_i(L) = 0$ for $i \equiv c \mod 2$,

(b) $a_i(L) = 0$ for $i < c - 1$.

4. *If $c = 1$, $a_0(L) = 1$.*

5. *If $c = 2$, $a_1(L)$ is the linking number of the two components of L.*

6. *If L_+, L_-, L_0 are as in skein relations, and L_+ and L_- have one component each, then L_0 has two components, and $a_2(L_+) - a_2(L_-)$ is the linking number of these two components.*

Proof. Points 1 and 2 have already been established.

For point 3, we use the skein relations. Suppose $i \equiv c \mod 2$, and suppose both (a) and (b) holds for all diagrams with at most n crossings. (The base case is immediate, by inspecting the Conway polynomials of the unknot and unlinks.) Let D be a diagram of a link L with c components and $n+1$ crossings. By the skein relations, we can write

$$\nabla_L(z) = \nabla_{L'}(z) + z\nabla_{L_1}(z) + \cdots + z\nabla_{L_m}(z)$$

where the L_i are a sequence of links with $c-1$ components and n crossings. The link L' is an unlink obtained after uncrossing m crossings. Since the Conway polynomial of the unlink vanishes, we've written $\nabla_L(z) = z(\sum \nabla_{L_i}(z))$. Both (a) and (b) follow.

For point 4, observe that $a_0(L) = \nabla_L(0)$. But evaluating the skein relation at 0 shows that changing any crossing of L doesn't modify $a_0(L)$. By unknotting, we obtain $a_0(L) = 1$.

For point 5, suppose L has components L_1 and L_2, with $\mathrm{lk}(L_1, L_2) = n$. Assume without loss of generality that $n \geq 0$. Consider the skein relation with $L_+ = L$, and change one of the crossings of the two components. Then L_- has two components L_1' and L_2', and L_0 has one component. Then

$$\mathrm{lk}(L_1, L_2) - \mathrm{lk}(L_1', L_2') = 1 = a_0(L_0) = a_1(L) - a_1(L_-).$$

By induction, the linking number $\mathrm{lk}(L_1, L_2)$ is equal to $a_1(L) - a_1(\widetilde{L})$. where L is a link with two components with crossing number 0. In particular, we could have chosen L to be a split link, so $a_1(L) = \mathrm{lk}(L_1, L_2)$.

Point 6 follows immediately from point 5. □

6.2 Signatures

Another knot invariant obtained from Seifert matrices are *signatures.* Unlike the Alexander polynomial, the signatures detect reflections.

Definition 6.2.1. Let L be an oriented link. The *signature* of L is

$$\sigma(L) = \mathrm{signature}(A + A^T)$$

where A is a Seifert matrix of L.

To prove that this is a well defined invariant, it suffices to show that the signature of $A + A^T$ is invariant under S-equivalence. By Sylvester's law of inertia, the signature is invariant under conjugation by unimodular matrices. One can explicitly show that the signature is also invariant under elementary enlargement and elementary reductions.

Proposition 6.2.2. The signature satisfies the following properties:

- $\sigma(L) = \sigma(rL)$,
- $\sigma(L) = -\sigma(\overline{L})$,
- $\sigma(K_1 + K_2) = \sigma(K_1) + \sigma(K_2)$.

Proof. These are immediate from the fact that if A is a Seifert matrix for L, then A^T is a Seifert matrix for rL, and $-A$ is a Seifert matrix for $\overline{L}$. Finally $\operatorname{diag}(A_1, A_2)$ is a Seifert matrix for $K_1 + K_2$, if A_1 and A_2 are Seifert matrices for K_1 and K_2 respectively. $\square$

This can be generalised in the following way:

Definition 6.2.3. The ω-signature $\sigma^*_\omega(L)$ is defined for an oriented link L and unit-length $\omega \in \mathbb{C}$ by

$$\sigma^*_\omega(L) = \text{signature}((1-\omega)A + (1-\overline{\omega})A^T).$$

where A is a Seifert matrix of L. This is defined for $\omega \neq 1$ and ω which are not zeros of the Alexander polynomial of L. This is extended further as follows. The *Tristram-Levine* signature is the function

$$\sigma : \mathbb{S}^1 \to \mathbb{Z}; \quad \omega \mapsto \sigma_\omega(L) := \lim \frac{\sigma^*_{\omega^+}(L) + \sigma^*_{\omega^-}(L)}{2}$$

where ω^+ and ω^- are unit length complex numbers tending to ω with arguments above and below that of ω.

The Tristram-Levine signature a well defined invariant of oriented links, and also satisfies the three properties mentioned above concerning the (classical) signature.

6.3 Slice knots

Definition 6.3.1. A knot $K \subset \mathbb{S}^3$ is a *topologically slice knot* if it is the boundary of a disk D^2 locally flatly embedded in B^4. $K \subset \mathbb{S}^3$ is *smoothly slice* if it is the boundary of a smoothly embedded disk in B^4.

By *locally flat*, we mean that D^2 has a tubular neighbourhood $D^2 \times D^2$, which meets $\mathbb{S}^3$ exactly on $K \times D^2$.

Definition 6.3.2. A knot $K \subset \mathbb{S}^3$ is a *ribbon knot* if it bounds a disk D^2 immersed in $\mathbb{S}^3$, which only has *ribbon singularities*. That is, D^2 intersects itself transversely along arcs, and the preimage of each such arc in the disk consists of two arcs: one is completely contained in the interior of the disk, and the other has its endpoints on K.

A famous conjecture asks whether all slice knots are ribbon knots. This is currently unsolved, but as of 2011 it is true for three-strand pretzel knots and knots with bridge number 2. A possible family of counterexamples has been suggested. The converse is known to be true.

Proposition 6.3.3. Let F be a genus g Seifert surface for a slice knot K in $\mathbb{S}^3$. There is a basis for $H_1(F;\mathbb{Z})$ such that the Seifert matrix in this basis has the form

$$\begin{pmatrix} 0 & P \\ Q & R \end{pmatrix}$$

consisting of $g \times g$ blocks of integers.

A proof is given in Lickorish. An immediate corollary is a useful necessary condition for sliceness:

Theorem 6.3.4. *If K is a slice knot, then the Conway-normalised Alexander polynomial is of the form $f(t)f(t^{-1})$, where f is a polynomial with integer coefficients.*

Proof. Consider a Seifert matrix $A = \begin{pmatrix} 0 & P \\ Q & R \end{pmatrix}$. Then

$$\begin{aligned}
\det(t^{1/2}A - t^{-1/2}A^T) &= \det \begin{pmatrix} 0 & t^{1/2}P - t^{-1/2}Q^T \\ t^{1/2}Q - t^{-1/2}P^T & t^{1/2}R - t^{-1/2}R^T \end{pmatrix} \\
&= -\det(t^{1/2}P - t^{-1/2}Q^T)\det(t^{1/2}Q - t^{-1/2}P^T) \\
&= \det(tP - Q^T)\det(t^{-1}P - Q^T).
\end{aligned}$$

□

Theorem 6.3.5. *If K is a slice knot, then its signature (and Tristram-Levine signature) vanish.*

Proof. This follows from inspecting the Seifert matrix A with the designated form. □

Which knots are slice? One way this can be determined is by considering the knot to be a level set of a Morse function on the disk. Without loss of generality, the disk has only critical points of index 0 and 1. A critical point of index 0 is bounded by an unknot. A critical point of index 1 is represented in a link diagram by moving between the following two local configurations in figure 6.1: Using this process, we can show that the square knot $3_1 + \overline{3_1}$ is slice. More generally, all knots of the form $K + \overline{K}$ are slice.

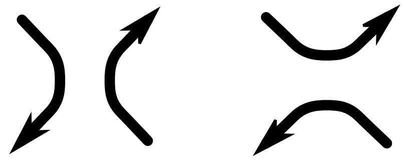


Figure 6.1: Change in level sets for index 1 critical points

Definition 6.3.6. Oriented knots K_0 and K_1 are *concordant* if $K_0 + r\overline{K_1}$ is slice.

To see that concordance is an equivalence relation, we can phrase it in terms of cobordisms.

Theorem 6.3.7. *Oriented knots K_0 and K_1 are concordant if and only if they bound $\mathbb{S}^1 \times [0,1]$ embedded in $\mathbb{S}^3 \times [0,1]$, such that $\mathbb{S}^1 \times 0 = K_0$ and $\mathbb{S}^1 \times 1 = K_1$.*

Theorem 6.3.8. *The* knot concordance group*, denoted $\mathcal{C}$, is the group*

$$\mathcal{C} = \{\textit{concordance classes of oriented knots}\}.$$

This is a group under the operation of taking connected sums. The unknot is the additive inverse.

Remark. Isotopy is stronger than concordance, which is stronger than homotopy. To see that these are strict, note that slice knots are exactly the members of the concordance class of the unknot. Since non-trivial slice knots exist, isotopy is strictly stronger than concordance. On the other hand, not all knots are slice, so concordance is strictly stronger than homotopy.

Theorem 6.3.9. *Given any slice knot K, there exists a ribbon knot R such that $K + R$ is ribbon.*

Recall that every ribbon knot is slice, but it is not currently known if every slice knot is ribbon.

Definition 6.3.10. The *slice genus* $g^*(K)$ of a knot K is the minimum genus of an oriented surface F locally-flatly embedded in B^4 with boundary K.

It is immediate that the 4-genus of a knot is a lower bound of the genus, and any slice knot has vanishing 4-genus. The previous theorem generalises as follows:

Theorem 6.3.11. *Let K be a knot. Then*

$$|\sigma_\omega(K)| \leq 2g^*(K).$$

In fact, the 4-ball genus is a lower bound of the unknotting number!

Theorem 6.3.12. *Let K be a knot. Then*

$$g^*(K) \leq u(K).$$

Proof. The idea is that every crossing change corresponds to the addition or removal of the genus of an oriented surface bound by the knot. Earlier we noted that passing an index 1 critical point of a Morse function on a slice surface corresponds to changing un-crossings as in figure 6.1. Changing a *crossing* can be achieved by two such moves, so each crossing change adds or removes genus. It follows that at least $g^*(K)$ crossing changes are required to obtain a genus-0 surface, so $g^*(K) \leq u(K)$. □

Remark. We have now obtained two lower bounds for the unknotting number of a knot:

- If the rth Alexander polynomial of K is not a unit, $r \leq u(K)$.
- $|\sigma_\omega(K)| \leq 2g^*(K) \leq 2u(K)$.

6.4 Exercises

Exercise 6.4.1. (Lickorish 8.8) Prove that the unknotting number of 8_2 is 2.

Solution: It is easy to see that the unknotting number is at most 2, by reversing two of the crossings in the usual diagram of 8_2.

A Seifert matrix of the Seifert surface of 8_2 obtained from the Seifert algorithm is

$$A = \begin{pmatrix} 1 & & & & & \\ -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & -1 & 1 & & \\ & & & -1 & 1 & \\ & & & & -1 & -1 \end{pmatrix}.$$

Then the 5×5 minor of $tA - A^T$ (in the top right) is 1, so the second Alexander ideal of 8_2 is trivial. This shows that we cannot use the "rth Alexander polynomial" bound. Instead, using A, we compute the signature of 8_2. With A as above, $A + A^T$ has five positive eigenvalues and one negative eigenvalue, so $\sigma(8_2) = 4$. But then $g^*(8_2) \geq 2$. Since the 4-genus is a lower bound of the uncrossing number, this gives $u(8_2) \geq 2$. △

Exercise 6.4.2. (Lickorish 8.9) Prove that the unknotting number of the sum of n copies of the trefoil is n.

Solution: The trefoil knot has Seifert matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Therefore

$$\sigma_i(3_1) = \sigma((1-i)A + (1+i)A^T) = \sigma \begin{pmatrix} 2 & 1-i \\ 1+i & 2 \end{pmatrix} = 2.$$

More generally, a Seifert matrix for the connected sum of n copies of the trefoil is given by $A_n = \bigoplus_n A$. The signature of $(1-i)A_n + (1+i)A_n^T$ is then $\sum_n \sigma((1-i)A + (1+i)A^T) = \sum_n \sigma(3_1) = \sum_n 2 = 2n$. It follows that $u(\#^n 3_1) \geq \sigma_i(\#^n 3_1)/2 = 2n/2 = n$.

To make this inequality an equality, simply observe that the connected sum of n trefoils can be made into an unknot by uncrossing any crossing in each copy of 3_1 in the diagram. $\triangle$

Chapter 7

Branched covers and the Goeritz matrix

7.1 Covering spaces and the determinant of a link

Recall that in the chapter concerning the Alexander polynomial, we introduced the *infinite cyclic cover* $X_\infty \to X$ of the link complement X. This is *infinite cyclic* in the sense that the Deck transformations of X_∞ are $\mathbb{Z} = \langle t \rangle$.

Any loop in X with linking number 0 with L corresponds to a loop which has algebraic intersection number 0 with a Seifert surface F of L. By the construction of X_∞, we see that a loop in X lifts to a loop in X_∞ if and only if the loop has linking number 0 with L.

If L is a knot, by the second to last theorem in the first chapter, we have that $H_1(X)$ is cyclic, and generated by a meridian of L. Moreover, given any curve C in X, $[C] \in H_1(X)$ is $\mathrm{lk}(C, L)$. Therefore $p_*\pi_1(X_\infty)$ is the kernel of the natural map $\pi_1(X) \to H_1(X)$. By Hurewicz, the kernel is also the commutator subgroup of $\pi_1(X)$.

Definition 7.1.1. The *cyclic double cover* of X is $\widehat{X}_2 \to X$ induced from X_∞ by modding out by $\langle t^2 \rangle$.

Of course, $\widehat{X}_2$ can be obtained by gluing two copies of Y, where Y is X cut along a Seifert surface F. Unlike the infinite cyclic cover, a loop in X lifts to a loop in $\widehat{X}_2$ if and only if it has linking number zero mod 2 with L.

Definition 7.1.2. The *cyclic double cover of* $\mathbb{S}^3$ *branched over* L is constructed as follows:

1. Let $\widehat{X}_2 \to X$ be the cyclic double cover of $X = \mathbb{S}^3 - L$.

2. Any component L_i of L has a tubular neighbourhood N in $\mathbb{S}^3$, homeomorphic to a solid torus. A longitude lifts to a loop in $\widehat{X}_2$, and the square of a meridian lifts to a loop in $\widehat{X}_2$ (since the square has linking number 0 mod 2 with L_i).

3. ∂N can be identified with $\mathbb{S}^1 \times \mathbb{S}^1$, parametrised by longitude and meridian. The restriction of $\widehat{X}_2 \to X$ to ∂N is given by $(z_1, z_2) \mapsto (z_1, z_2^2)$.

4. This map extends to a map on $\widehat{X}_2 \sqcup_{p^{-1}\partial N} (\mathbb{S}^1 \times D^2) \to X \sqcup_{\partial N} (\mathbb{S}^1 \times D^2)$ by the formula $(z_1, z_2) \mapsto (z_1, z_2^2)$. This is a *branched cover*, as it is a 2:1 covering map on X, but degenerates to a 1:1 map over L_i.

5. For each of the n components of L, glue a solid torus to $\widehat{X}_2$. Similarly glue n solid tori into X. We obtain a double cover $X_2 \to \mathbb{S}^3$, but branched over L.

Note that the construction is independent of the orientation of L, since $\mathbb{Z}/2\mathbb{Z}$ cannot detect orientation. The above construction automatically generalises to $m : 1$ branched covers for any m.

Definition 7.1.3. The *group* of a cover $p : E \to X$ is $p_*\pi_1(E)$.

Theorem 7.1.4. *Let X_2 be the cyclic double cover of $\mathbb{S}^3$ branched over L. Suppose A is a Seifert matrix of L (given some orientation). Then $A + A^T$ presents $H_1(X_2; \mathbb{Z})$ as a $\mathbb{Z}$-module.*

Observe that this result parallels the earlier result that $H_1(X_\infty; \mathbb{Z})$ is presented by $tA - A^T$ as a $\mathbb{Z}[t, t^{-1}]$-module.

Corollary 7.1.5. *Let X_2 be the cyclic double cover of $\mathbb{S}^3$ branched over L. The order of $H_1(X_2)$ is*

$$|H_1(X_2)| = |\det(A + A^T)| = |\Delta_L(-1)|.$$

Proof. Since $H_1(X_2)$ is a finitely presented abelian group, it's a direct sum of cyclic groups. The corresponding presentation matrix M is diagonal. Each diagonal entry is the order of the corresponding summand, so the determinant of M is the order of $H_1(X_2)$. (We use the convention that an infinite group has order 0 to ensure that everything is well defined). The determinant of any two square presentation matrices differ by ± 1, so we obtain the first equality. The second equality is immediate from the homological definition of the Alexander polynomial. □

Remark. Recall that for any K we can write

$$\Delta_K(t) = a_0 + a_1(t^{-1} + t) + \cdots .$$

For a knot K, $\Delta_K(1) = \pm 1$, so $\Delta_K(-1)$ is necessarily an odd integer. It follows from the above formula that $H_1(X_2)$ is always finite whenever X^2 is a cyclic double cover branched over a knot.

Definition 7.1.6. The *determinant* of a link is the invariant $|\Delta_L(-1)|$.

By the previous corollary, the determinant is the size of the first homology group of the double cover of $\mathbb{S}^3$ branched over L. Currently it isn't clear that this is a useful invariant, but we soon see that the determinant is very easy to compute via Goeritz matrices.

7.2 Gordon-Litherland form and Goeritz matrix

In an earlier chapter we introduced the *Seifert form*

$$\alpha : H_1(F;\mathbb{Z}) \times H_1(F;\mathbb{Z}) \to \mathbb{Z},$$

defined to count the linking number of loops with push-offs of other loops, with orientations induced from F. We now introduce the *Gordon-Litherland form* $\mathcal{G}_F : H_1(F;\mathbb{Z}) \times H_1(F;\mathbb{Z}) \to \mathbb{Z}$ which is well defined for all surfaces with boundary a link (rather than just oriented surfaces). The idea is to push-off in "both directions" and count linking numbers.

Let $L \subset \mathbb{S}^3$ be a link, and F any surface embedded in $\mathbb{S}^3$ with boundary L. (F need not be orientable.) F admits a fibre bundle $p : E \to F$ with fibre I, where E is a tubular neighbourhood of F in $\mathbb{S}^3$, and locally a product of F with I. This induces a "boundary bundle" with fibre $\partial I = \{-1, 1\}$. The total space of this induced bundle is the orientable double over of F, which we denote by $\widetilde{F}$.

Definition 7.2.1. Let $L \subset \mathbb{S}^3$ be a link, and F any surface embedded in $\mathbb{S}^3$ with boundary L. Let $p : \widetilde{F} \to F$ be the orientable double cover of F, embedded in $\mathbb{S}^3$. The *Gordon-Litherland form* is the map

$$\mathcal{G}_F : H_1(F;\mathbb{Z}) \times H_1(F;\mathbb{Z}) \to \mathbb{Z}$$

defined by $\mathcal{G}_F([f],[g]) = \mathrm{lk}(p^{-1}f, g)$.

In the case where F is orientable, $p^{-1}f$ consists of two loops, i^+f and i^-f (in the notation used when defining the Seifert form). If F is not orientable, $p^{-1}f$ consists of a single loop.

Remark. The cyclic double cover $\widehat{X}_2$ of $X = \mathbb{S}^3 - L$ can be constructed by gluing two copies of $\mathbb{S}^3 - F$ together, where F is any (not necessarily orientable) surface with boundary L. Similarly we can then define the cyclic double cover X_2 of $\mathbb{S}^3$ branched over L.

Theorem 7.2.2. *Let $L \subset \mathbb{S}^3$ be a link, and F a surface embedded in $\mathbb{S}^3$ with boundary L. Any matrix representing the Gordon-Litherland form is a presentation matrix for $H_1(X_2)$.*

Conversely, we noted earlier that $A + A^T$ presents $H_1(X_2)$, where A is a Seifert matrix. It is clear that $A + A^T$ presents $\mathcal{G}_F$ when F is orientable, from our earlier remark that $p^{-1}f$ then consists of two loops, i^+f and i^-f.

Definition 7.2.3. The *Goeritz matrix* for a link is constructed as follows:

1. Let D be a connected diagram of a link L. Give D a chessboard colouring. Let $R_0, \ldots, R_n$ be the white regions of D.

2. At any crossing c, there is an associated sign $\zeta(c)$. If a white region is on the left before an underpass, $\zeta(c) = 1$. Otherwise $\zeta(c) = -1$.

3. A *pre-Goeritz matrix* is an $(n+1) \times (n+1)$ matrix defined by

$$g_{ij} = \sum \zeta(c)$$

   where the sum is taken over crossings where regions R_i and R_j meet, for $i \neq j$.

   The diagonal terms are defined by

$$g_{ii} = -\sum_{j \neq i} g_{ij}.$$

4. The *Goeritz matrix* is an $n \times n$ matrix obtained by deleting a column and corresponding row. By convention we delete the column and row indexed by 0, to obtain an $n \times n$ matrix $g_{ij} : 1 \leq i, j \leq n$.

Theorem 7.2.4. *Any Goeritz matrix for a link diagram D (associated to the white regions) represents, with respect to some basis, the Gordon-Litherland form $\mathcal{G}_F$, where F is the surface obtained from the black regions of the colouring of D.*

Corollary 7.2.5. *The determinant of L is equal to $|\det G|$, where G is any Goeritz matrix for L.*

Since G is easy to write down, it can be a useful invariant.

Example. By comparing various knot polynomials, it has been established that the trefoil and figure 8 knots are distinct. However, one "technically easy" way to see this is by comparing determinants.

A pre-Goeritz matrix for the trefoil knot is

$$P = \begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix}.$$

Therefore a Goeritz matrix is $G = (-3)$, so that $|\det G| = 3 = \det 3_1$.

On the other hand, a pre-Goeritz matrix for the figure-8 knot is

$$P = \begin{pmatrix} -3 & 2 & 1 \\ 2 & -3 & 1 \\ 1 & 1 & -2 \end{pmatrix}.$$

Therefore $|\det G| = 5 = \det 4_1$. It follows that the trefoil and figure 8 knots are distinct.

Chapter 8

Arf invariant

8.1 Quadratic forms and the classical Arf invariant

Let V be a finite dimensional vector space over a field k, not of characteristic 2. Let $B : V \times V \to k$ be a symmetric bilinear form. Fixing a basis $\{e_1, \ldots, e_n\}$ of V, B can be expressed as a symmetric matrix with entries b_{ij}.

Now given any vector $x = (x_1, \ldots, x_n)$, we have

$$B(x,x) = \sum_{i=1}^{n} b_{ii}x_i^2 + 2\sum_{i<j} b_{ij}x_ix_j.$$

This is a polynomial in n entries, homogeneous of degree 2. Therefore we call this a *quadratic form*. More formally, the above is map satisfying the following homogeneity property:

$$\widetilde{B} : V \to k, \quad \widetilde{B}(ax) = a^2\widetilde{B}(x).$$

Moreover, it is easy to verify that

$$B(x,y) = \frac{1}{2}(\widetilde{B}(x+y) - \widetilde{B}(x) - \widetilde{B}(y)).$$

Definition 8.1.1. Let k be a field, and V a finite dimensional vector space over k. A *quadratic form* is a map $\varphi : V \to k$ such that $\varphi(ax) = a^2\varphi(x)$ for all $x \in V, a \in k$, and such that

$$(x,y) \mapsto \varphi(x+y) - \varphi(x) - \varphi(y)$$

is a symmetric bilinear form. φ is said to be *non-degenerate* if the associated bilinear form is non-degenerate.

Remark. When k is not of characteristic 2, the data of a symmetric bilinear form is equivalent to that of a quadratic form: a bilinear form B determines a quadratic form ψ

by $B(x,x) = \psi(x)$. Conversely, B is recovered from ψ by

$$B : (x,y) \mapsto \frac{1}{2}(\psi(x+y) - \psi(x) - \psi(y)).$$

Remark. When k is of characteristic 2, the property $\varphi(ax) = a^2\varphi(x)$ for all $x \in V, a \in k$ is implied by the requirement that $\varphi(x+y)-\varphi(x)-\varphi(y)$ be bilinear. Explicitly, by bilinearity,

$$0\varphi(x) = 0 = \varphi(x+0y) - \varphi(x) - \varphi(0y) = \varphi(0y).$$

On the other hand, it is immediate that $1\varphi(x) = \varphi(1x)$.

Definition 8.1.2. Let $\varphi : V \to \mathbb{Z}/2\mathbb{Z}$ be a non-degenerate bilinear form. Then φ is either of *type I* or *type II*, which we now describe.

1. Since φ is non-degenerate, the associated symmetric bilinear form B is non-degenerate. Over $\mathbb{Z}/2\mathbb{Z}$, this is equivalently a skew-symmetric bilinear form, and hence B can be expressed in a symplectic basis $\{e_1, f_1, \ldots, e_n, f_n\}$. That is,

   $$B = \bigoplus_{i=1}^{n} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

   where V is $2n$-dimensional. In this basis, we have

   $$\varphi(x_1e_1 + y_1f_1 + \cdots) = \sum_{i=1}^{n} x_i^2\varphi(e_i) + \sum_{i=1}^{n} y_i^2\varphi(f_i) + \sum_{i=1}^{n} x_iy_i.$$

2. We now attempt to remove the square terms. Suppose $\varphi(e_i)$ is non-zero, and $\varphi(f_i) = 0$. Then replace $\{e_i, f_i\}$ with

   $$\{g_i, h_i\} = \{e_i + f_i, f_i\}$$

   in a new basis. This is also symplectic, and we have

   $$x_i^2 + x_iy_i = x_i'y_i'$$

   where $\sum_i x_i'g_i + y_i'h_i = \sum_i x_ie_i + y_if_i$. Therefore we can eliminate any square terms that do not appear in pairs.

3. Suppose $\varphi(e_i), \varphi(f_i), \varphi(e_j), \varphi(f_j)$ are all non-zero. Define a new basis for V in which we replace $\{e_i, f_i, e_j, f_j\}$ with

   $$\{g_i, h_i, g_j, h_j\} = \{(e_i + e_j + f_i), (e_i + e_j + f_j), (e_i + f_i + f_j), (e_j + f_i + f_j)\}.$$

   Then we obtain another symplectic basis, and square terms have been removed:

   $$x_i^2\varphi(e_i) + x_j^2\varphi(e_j) + y_i^2\varphi(f_i) + y_j^2\varphi(f_j) + x_iy_i + x_jy_j = x_i'y_i' + x_j'y_j',$$

   where $\sum_i x_i'g_i + y_i'h_i = \sum_i x_ie_i + y_if_i$. This establishes that pairs of non-vanishing squares can be cancelled against other non-vanishing pairs of squares.

4. By eliminating pairs in the above fashion, and reordering the symplectic basis if required, any φ can be expressed in either of the following ways:

$$\varphi(x_1e_1 + \cdots + y_nf_n) = x_1y_1 + \cdots + x_ny_n$$
$$\varphi(x_1e_1 + \cdots + y_nf_n) = x_1y_1 + \cdots + x_ny_n + x_n^2 + y_n^2.$$

If φ can be written in the former way, φ is of *type I.* If φ can be written in the latter way, φ is of *type II.*

Definition 8.1.3. The *Arf invariant* of a non-degenerate quadratic form $\varphi : V \to \mathbb{Z}/2\mathbb{Z}$, denoted $c(\varphi)$, is defined by

$$c(\varphi) = \begin{cases} 0 & \text{if } \varphi \text{ is of type I} \\ 1 & \text{if } \varphi \text{ is of type II.} \end{cases}$$

Proposition 8.1.4. Let $\varphi : V \to \mathbb{Z}/2\mathbb{Z}$ be a non-degenerate quadratic form. The following values are equal:

1. The Arf invariant $c(\varphi)$ of φ.

2. The value 0 or 1 attained more often by φ as it ranges over the 2^{2n} elements of V.

3. The value $\sum_{i=1}^{n} \varphi(e_i)\varphi(f_i)$ where $\{e_1, f_1, \ldots, e_n, f_n\}$ is any symplectic basis.

Proof. We first show that 1 and 2 are equal.

Let V be of dimension $2n$. We proceed by induction on n to show that if φ is of type I, then the value 1 is attained by φ $2^{2n-1} - 2^{n-1}$ times, and if φ is of type II, 1 is attained $2^{2n-1} + 2^{n-1}$ times.

Suppose $n = 1$, and φ is of type I. Then φ attains 1 once. On the other hand, if φ is of type II, it attains 1 three times. This verifies the base case.

Next suppose that type I quadratic forms on V^{2k} attain 1 $2^{2k-1} - 2^{k-1}$ times. Let φ be a type I quadratic form on $V^{2(k+1)}$. Then φ induces a type I quadratic form φ' on V^{2k} by forgetting the first two coordinates. If the first two coordinates of x are both 1, then $\varphi(x) = 1$ provided that $\varphi'(x') = 0$. This happens $2^{2k-1} + 2^{k-1}$ times. If either of the first two coordinates of x are 0, then $\varphi(x) = 1$ provided that $\varphi'(x') = 1$. This happens $3(2^{2k-1} - 2^{k-1})$ times. In total, we have

$$2^{2k-1} + 2^{k-1} + 3(2^{2k-1} - 2^{k-1}) = 2^{2(k+1)-1} - 2^{(k+1)-1},$$

as required. Similarly a straight forward calculation verifies the result for φ of type II.

A similar inductive proof shows that 2 is equal to 3. $\square$

8.2 The Arf invariant for links

Definition 8.2.1. We say that a link L is *proper* if $\mathrm{lk}(L_i, L - L_i)$ is even for any component L_i of L.

Suppose L is an oriented link, with Seifert surface F. Then a quadratic form

$$q : H_1(F; \mathbb{Z}/2\mathbb{Z}) \to \mathbb{Z}/2\mathbb{Z}$$

can be defined by $q(x) = \alpha(x, x) \mod 2$, where α is the Seifert form. The symmetric bilinear form

$$(x, y) \mapsto q(x + y) - q(x) = q(y)$$

is represented by $A - A^T$, and counts the intersection number of transverse curves. In general this bilinear form is degenerate, but indices a non-degenerate form on the quotient $H_1(F; \mathbb{Z}/2\mathbb{Z})/H_1(\partial F; \mathbb{Z}/2\mathbb{Z})$. Suppose L is a proper link. Then

$$q([L_i]) = \mathrm{lk}(L_i^-, L) = \mathrm{lk}(L_i, L - L_i) = 0 \mod 2.$$

It follows that q induces a well defined quadratic form on $H_1(F; \mathbb{Z}/2\mathbb{Z})/H_1(\partial F; \mathbb{Z}/2\mathbb{Z})$.

Definition 8.2.2. Let L be an oriented proper link. The *Arf invariant* $\mathcal{A}(L)$ is the Arf invariant of the quadratic form

$$q : H_1(F; \mathbb{Z}/2\mathbb{Z})/i_* H_1(\partial F; \mathbb{Z}/2\mathbb{Z}) \to \mathbb{Z}/2\mathbb{Z}$$

described above. Here i denotes the inclusion $\partial F \to F$, the Arf invariant is independent of the choice of F.

Proposition 8.2.3. The Arf invariant satisfies the following properties:

1. $\mathcal{A}(0_1) = 0$.
2. $\mathcal{A}(L_1 + L_2) = \mathcal{A}(L_1) + \mathcal{A}(L_2)$.
3. If L and L' are proper links differing at a single point as shown in figure 8.1, they have the same Arf invariant.

Property 2 above follows from algebra: given quadratic forms φ and ψ, $c(\varphi \oplus \psi) = c(\varphi) + c(\psi)$.

Recall that a method of proving that a knot is slice is by taking a knot diagram and applying the move shown in figure 6.1 (which is just figure 8.1 with orientations) to reduce it to a collection of unlinked unknots. This suggests that the Arf invariant may not only be an invariant of links up to isotopy, but maybe even concordance. We see later in an exercise that this is indeed the case.

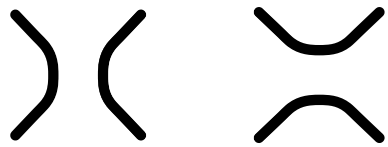


Figure 8.1: Arf invariant invariance

Theorem 8.2.4. *Let L be any oriented link. Then*

$$V(L)(i) = \begin{cases} (-\sqrt{2})^{\#L-1}(-1)^{\mathcal{A}(L)} & \text{if } L \text{ is proper} \\ 0 & \text{otherwise.} \end{cases}$$

Remark. This is an interesting result because it sheds light on why knot theorists struggled to extend the definition of the Arf invariant to links which are not proper.

Proof. We give a proof strategy. Recall that the Jones polynomial is characterised by

- $V(0_1) = 1$, and
- $t^{-1}V(L_+) - tV(L_-) + (t^{-1/2} - t^{1/2})V(L_0) = 0$.

Therefore to prove that the Arf invariant satisfies the claimed identity, it suffices to show that for $\widehat{A}(L) = (-\sqrt{2})^{\#L-1}(-1)^{\mathcal{A}(L)}$,

- $\widehat{A}(0_1) = 1$, and
- $\widehat{A}(L_+) + \widehat{A}(L_-) + \sqrt{2}\widehat{A}(L_0) = 0$.

The proof proceeds by case work, by verifying this formula for L_+, L_- proper and L_0 not proper, and so on. This can all be achieved by using the third property of the Arf invariant in the previous proposition. $\square$

The Arf invariant also relates to the Alexander polynomial in a clean way:

Theorem 8.2.5. *Let K be a knot. Then $\mathcal{A}(K) \equiv a_2(K)$ modulo 2, where $a_2(K)$ is the coefficient of z^2 in the Conway polynomial $\nabla_K(z)$. Moreover,*

$$\mathcal{A}(K) = \begin{cases} 0 & \text{if } \Delta_K(-1) \equiv \pm 1 \bmod 8 \\ 1 & \text{if } \Delta_K(-1) \equiv \pm 3 \bmod 8. \end{cases}$$

Proof. This follows from the observation that $\mathcal{A}(L_+) - \mathcal{A}(L_-) \equiv \mathrm{lk}(L_0)$ modulo 2, when L_+ is a knot. Note that $\mathrm{lk}(L_0)$ denotes the linking number of the two components of L_0.

Since L_+ and L_- have one component, they are automatically proper. On the other hand, L_0 is proper if and only if $\mathrm{lk}(L_0)$ is even. If L_0 is proper, then by the third property in the previous proposition, $\mathcal{A}(L_0) = \mathcal{A}(L_+) = \mathcal{A}(L_-)$. Therefore

$$\mathcal{A}(L_+) - \mathcal{A}(L_-) \equiv 0 \equiv \mathrm{lk}(L_0) \mod 2.$$

Similarly in the case where L_0 isn't proper, then one can show that exactly one of $\mathcal{A}(L_+)$ and $\mathcal{A}(L_-)$ vanish. Therefore

$$\mathcal{A}(L_+) - \mathcal{A}(L_-) \equiv 1 \equiv \mathrm{lk}(L_0) \mod 2.$$

But, recalling the properties of the Conway polynomial, we also have $a_2(L_+) - a_2(L_-) = \mathrm{lk}(L_0)$. This proves the first part of the proposition.

For the second part, write $\Delta_K(-1) = \nabla_K(-2i)$. (This holds in the Conway normalisation.) Then $\Delta_K(-1) \equiv 1 - 4a_2(K)$ modulo 8. This completes the proof. □

Remark. Recall that $|\Delta_K(-1)|$ is the *determinant* of K. This shows that the Arf invariant of a knot can be determined from the determinant.

Corollary 8.2.6. *Let K be a slice knot. Then $\mathcal{A}(K) = 0$.*

Proof. By the Fox-Milnor condition, $\Delta_K(t) = f(t)f(t^{-1})$ for some polynomial f. In particular, $\Delta_K(-1) = f(-1)^2$. Moreover, $\Delta_K(-1)$ is odd. Therefore $|\Delta_K(-1)|$ is the square of an odd integer, which is necessarily ± 1 modulo 8. The result now follows from the previous theorem. □

Theorem 8.2.7. *Two knots are said to be* pass equivalent *if their diagrams are related by a sequence of* pass moves*. Every knot is pass equivalent to either the trefoil or the unknot. Moreover, the trefoil and unknot are not pass equivalent. If K is pass equivalent to the unknot, it has Arf invariant zero. If K is pass equivalent to the trefoil, it has Arf invariant one.*

8.3 Exercises

Exercise 8.3.1. (Lickorish 10.2) Determine, directly from a Seifert matrix, the Arf invariant of a pretzel knot $P(p, q, r)$, where p, q, r are odd integers.

Solution: Recall that a Seifert matrix for $P(p, q, r)$ is given by

$$A = \frac{1}{2}\begin{pmatrix} p+q & q+1 \\ q-1 & q+r \end{pmatrix},$$

by choosing an appropriate basis $\{f_1, f_2\}$ of simple closed curves in the "canonical" Seifert surface. Using the Seifert matrix we can compute the determinant of the knot:

$$\Delta_{P(p,q,r)}(-1) = \det(A + A^T) = \det \begin{pmatrix} p+q & q \\ q & q+r \end{pmatrix} = pq + pr + qr.$$

The Arf invariant is now determined by the value of $pq + pr + qr$ mod 8.

For example, the trefoil knot is the $P(1,1,1)$ pretzel knot, so $pq + pr + qr = 3$ modulo 8. It follows that $\mathcal{A}(3_1) = 1$. On the other hand, the pretzel knot $P(-2,3,7)$, also called the Fintushel–Stern knot, has $pq + pr + qr = -6 - 14 + 21 = -1$ modulo 8, so its Arf invariant is zero. $\triangle$

Exercise 8.3.2. (Lickorish 10.3) Prove that cobordant knots have the same Arf invariant.

Solution: Here *cobordant* means *concordant*. Recall that K_1 and K_2 are concordant if and only if $K_1 + r\overline{K_2}$ is slice. Then $\mathcal{A}(K_1 + r\overline{K_2}) = 0$. Since the Arf invariant is additive, it follows that $\mathcal{A}(K_1) = \mathcal{A}(r\overline{K_2})$. Finally we note that the Alexander polynomial detects neither reflections nor reversals, so $\Delta_{r\overline{K_2}} = \Delta_{K_2}$. By the previous theorem, it follows that $\mathcal{A}(r\overline{K_2}) = \mathcal{A}(K_2)$. Therefore the Arf invariant is a concordance invariant. $\triangle$

Chapter 9

Knot groups

In the chapter concerning the Alexander polynomial, we introduced the knot group of a knot K;

$$\pi_1(\mathbb{S}^3 - K).$$

We then described the *Wirtinger presentation* and how to determine the Alexander polynomial from it via Fox's free differential calculus. In this chapter we further explore the knot group.

9.1 A knot group is a $K(G, 1)$

In this section we show that knot groups are Eilenberg Mac Lane spaces. We begin by stating two important theorems from the study of 3-manifold topology.

Theorem 9.1.1 (The loop theorem)**.** *Let M be a 3-manifold with boundary. Suppose $i_* : \pi_1(\partial M) \to \pi_1(M)$ is* not *injective. Then there is an embedding $e : D^2 \to M$ with $e^{-1}(\partial M) = \partial D^2$ such that $e : \partial D^2 \to \partial M$ is* not *homotopic to a constant map.*

Theorem 9.1.2 (The sphere theorem)**.** *Let M be an orientable 3-manifold. Suppose $\pi_2(M) \neq 0$. Then there is a (smooth) embedding $\mathbb{S}^2 \to M$ that represents a non-trivial element of $\pi_2(M)$.*

We now apply these theorems to the study of knots.

Lemma 9.1.3. Let K be a non-trivial knot, and $X = \mathbb{S}^3 - K$. Then $i_* : \pi_1(\partial X) \to \pi_1(X)$ is an inclusion.

Proof. Suppose conversely that $\pi_1(\partial X) \to \pi_1(X)$ is not injective. Then by the loop theorem, there is an embedding $e : D^2 \to X$ so that ∂D^2 maps into the torus ∂X, so that ∂D^2 represents a non-trivial element in $\pi_1(\partial X)$.

We show that $e(D^2)$ can be perturbed to give a Seifert surface of K, which shows that K is unknotted. On one hand, $e(\partial D^2)$ represents a non-trivial element of $H_1(\partial X)$. On the other hand, ∂D^2 bounds D^2 in X, so it must be trivial in $H_1(X)$. Since the only simple closed curve representing a non-trivial element of the kernel of $H_1(\partial X) \to H_1(X)$ is is a longitude of K, ∂D^2 is a longitude of K. The result follows. □

Corollary 9.1.4. *A knot K is the unknot if and only if $\pi_1(\mathbb{S}^3 - K)$ is infinite cyclic.*

Proof. If K is the unknot, the Wirtinger presentation of $\pi_1(\mathbb{S}^3 - K)$ is $\langle x \rangle \cong \mathbb{Z}$.

Conversely, suppose K is not the unknot. Then by the previous lemma, there is an injection $\mathbb{Z}^2 \cong \pi_1(\partial X) \to \pi_1(X)$, where X is the complement of K. Therefore $\pi_1(X)$ cannot be cyclic. □

Theorem 9.1.5. *Fix a knot K, and $X = \mathbb{S}^3 - K$. Let G be the group $\pi_1(X)$. Then X is a $K(G, 1)$.*

Recall that a $K(G, n)$ is a path connected space X such that $\pi_n(X) = G$, and all other homotopy groups vanish.

Proof. It is clear that X is path connected. It remains to show that all higher homotopy groups vanish.

First we prove that $\pi_2(X)$ is trivial. This follows from the *sphere theorem* and *Schönflies theorem.* If X has non-trivial second homotopy group, by the sphere theorem, it admits an embedding of $\mathbb{S}^2$ representing a non-trivial element of $\pi_2(X)$. By the Schönflies theorem, the image of $\mathbb{S}^2$ cuts X into two components. Since the knot K is connected, it can only lie in one component, so $\mathbb{S}^2$ must be null-homotopic in the other component. This is a contradiction.

Next we prove that $\pi_n(X)$ is trivial for $n \geq 3$. We proceed as follows:

1. Lift higher homotopy groups to the universal cover.

2. Show that every homology group of the universal cover vanishes.

3. Conclude from Hurewicz's theorem that all homotopy groups of the universal cover vanish, and deduce that $\pi_n(X)$ is trivial for all $n \geq 3$.

1. Let $\widetilde{X}$ denote the universal cover of X. Then for $n \geq 2$, $\pi_n(X)$ is trivial if and only if $\pi_n(\widetilde{X})$ is trivial. This can be seen from the homotopy long exact sequence.

2. It's clear that $H_2(\widetilde{X})$ vanishes from the previous result. It's also clear that $H_n(\widetilde{X})$ vanishes for $n \geq 4$, since $\widetilde{X}$ is a 3-manifold. Since $\widetilde{X}$ is a universal cover of a connected space, it is connected and simply connected, so $H_0(X)$ and $H_1(X)$ are also trivial. Finally note that the universal cover $\widetilde{X}$ is non-compact, since $\pi_1(X)$ is non-trivial. By Poincaré duality, the top homology of any non-compact connected manifold vanishes, so $H_3(\widetilde{X})$ is also trivial. In summary we've shown that all homology groups of $\widetilde{X}$ vanish.

3. By Hurewicz's theorem, the first non-zero homology group and first non-zero homotopy group occur in the same dimension. Since all homology groups of $\widetilde{X}$ vanish, so must all homotopy groups. Recall that for $n \geq 2$, $\pi_n(X)$ is trivial if and only if $\pi_n(\widetilde{X})$. Therefore for $n \geq 2$, $\pi_n(X)$ is trivial.

This completes the proof that the only non-trivial homotopy group of X is $\pi_1(X) = G$, and X is also connected. Therefore X is a $K(G, 1)$. □

The main result of this section can be interpreted as *the homotopy type of* $\mathbb{S}^3 - K$ *is determined by* $\pi_1(\mathbb{S}^3 - K)$. However, different knots can have the same knot groups.

Proposition 9.1.6. The invariant $\pi_1(\mathbb{S}^3 - K)$ of a knot K does not distinguish all knots.

Example. If K_1 and K_2 are knots, then $K_1 + K_2$ and $K_1 + r\overline{K_2}$ have homotopic knot complements. However, knots are in general not equivalent. For example, the square knot and granny knot.

However, the following theorem holds:

Theorem 9.1.7. *Let K_1, K_2 be knots, and μ_i, λ_i meridians and longitudes of K_i. Suppose there is an isomorphism between $\pi_1(\mathbb{S}^3 - K_1)$ and $\pi_2(\mathbb{S}^3 - K_2)$ which sends $[\mu_1]$ to $[\mu_2]$ and $[\lambda_1]$ to $[\lambda_2]$. Then K_1 and K_2 are equivalent.*

Moreover, the following two theorems tell us that for prime knots, we can drop the condition on meridians and longitudes.

Theorem 9.1.8. *If K_1 and K_2 are prime knots with isomorphic knot groups, then their knot complements are homeomorphic.*

Theorem 9.1.9. *If K_1 and K_2 are unoriented knots in $\mathbb{S}^3$ and there is an orientation preserving homeomorphism between their complements, then K_1 and K_2 are equivalent as unoriented knots.*

9.2 Alexander polynomial revisited

Recall the following process for determining the Alexander polynomial of a knot via the knot group:

1. Let K be a knot, and $G = \pi_1(\mathbb{S}^3 - K)$. Let $P = \langle x_1, \ldots, x_n \mid r_1, \ldots, r_m \rangle$ be a finite presentation of G. (For example, obtained as the Wirtinger presentation of K.)

2. Calculate the *Jacobian*
$$J(P)_{ij} = \partial_{x_j} r_i$$
of the presentation, where ∂_{x_j} is the Fox derivative.

3. Replacing each x_i with t, we obtain a matrix $J_T(P)$ with entries in $\mathbb{Z}[t]$. In fact, $J_T(P)$ is an Alexander matrix in the sense that the rth elementary ideals of $J_T(P)$ are the rth Alexander ideals.

4. In particular, the Alexander polynomial is the generator of the smallest principal ideal containing the minors of the largest square blocks of $J_T(P)$.

Recall that the rth elementary ideals of $J_T(P)$ do not depend on the choice of presentation P. This shows that all of the information of the Alexander polynomial is contained in the knot group. Of course, this is expected, since the knot group determines the homotopy type of the knot complement, and the Alexander polynomial is computed via the homology of (a cover of) the knot complement.

Next we describe how this can be extended to links, in a way that actually improves the Alexander polynomial.

1. Let L be a link, and $G = \pi_1(\mathbb{S}^3 - L)$. Let

$$P = \langle x_1^1, \ldots, x_{n_1}^1, x_1^2, \ldots, x_{n_c}^c \mid r_1, \ldots, r_m \rangle$$

be a finite presentation of G, Obtained as the Wirtinger presentation of L. Then each generator belongs to one of the c components of L, as labelled in the superscript.

2. Consider a surjection $\ell : \{1, \ldots, c\} \to \{1, \ldots, \nu\}$. This defines a "colouring" of the components, and gives rise to an Alexander polynomial in ν variables.

3. Compute the Jacobian $J(P)$ of the presentation P using Fox derivatives. Replace each instance of x_k^i in $J(P)$ with $t_{\ell(i)}$. This gives a matrix $J_T(P)$ with entries in $\mathbb{Z}[t_1, \ldots, t_\nu]$.

4. An Alexander polynomial in ν variables is then obtained by computing the greatest common divisor of the first minors of $J_T(P)$.

Remark. Rather than computing all of the minors of $J_T(P)$, one can simply compute the minor obtained by deleting any row and the jth column, and then divide the result by $t_{\ell(j)} - 1$. This works because of the redundancy in the Wirtinger presentation of a link group.

Theorem 9.2.1. *If p, q, r are odd integers with $|p|, |q|, |r|$ distinct and greater than 1, then $P(p, q, r)$ is not equivalent to its reverse.*

We do not give a proof here, but this result can be obtained by studying the knot groups. More precisely, one can write down the Wirtinger presentation for knot group, and if the knot is reversible, there must exist an automorphism of the knot group sending meridians to inverse meridians and words to their inverses. It can be shown that this is not possible.

9.3 Exercises

Exercise 9.3.1. (Exercise to myself) Compute a three variable Alexander polynomial for Borromean rings, as well as the one-variable Alexander polynomial, and make some comparisons.

Solution: First we compute a 3-variable Alexander polynomial for the Borromean rings. We begin this by writing down a Wirtinger presentation for its link complement:

$$\pi_1(\mathbb{S}^3 - L) = \langle x_1, x_2, y_1, y_2, z_1, z_3 \mid x_1y_1x_1^{-1}y_2^{-1}, y_2x_2y_1^{-1}x_2^{-1}, z_1x_1z_1^{-1}x_2^{-1}, \ldots, z_2y_2z_1^{-1}y_2^{-1}\rangle.$$

Using Fox calculus, the Jacobian of this presentation is

$$J = \begin{pmatrix} 1-x_1y_1x_1^{-1} & 0 & z_1 & -x_2z_2x_1^{-1} & 0 & 0 \\ 0 & y_2-y_2x_2y_1^{-1}x_2^{-1} & 1 & 0 & 0 & \\ x_1 & -y_2x_2y_1^{-1} & 0 & 0 & 1-y_1z_1y_1^{-1} & 0 \\ -x_1y_1x_1^{-1}y_2^{-1} & 1 & 0 & 0 & 0 & z_2-z_2y_2z_1^{-1}y_2^{-1} \\ 0 & 0 & 1-z_1x_1z_1^{-1} & 0 & y_1 & -z_2y_2z_1^{-1} \\ 0 & 0 & 0 & x_2-x_2z_2x_1^{-1}z_2^{-1} & -y_1z_1y_1^{-1}z_2^{-1} & 1 \end{pmatrix}.$$

Mapping $x_1, x_2 \mapsto t_1$, $y_1, y_2 \mapsto t_2$, and $z_1, z_2 \mapsto t_3$, we have

$$J_T = \begin{pmatrix} 1-t_2 & 0 & t_3 & -t_3 & 0 & 0 \\ 0 & t_2-1 & 1 & 0 & 0 & \\ t_1 & -t_1 & 0 & 0 & 1-t_3 & 0 \\ -1 & 1 & 0 & 0 & 0 & t_3-1 \\ 0 & 0 & 1-t_1 & 0 & t_2 & -t_2 \\ 0 & 0 & 0 & t_1-1 & -1 & 1 \end{pmatrix}.$$

The minors of the above matrix are

$$\pm(t_1-1)^2(t_2-1)(t_3-1), \pm(t_1-1)(t_2-1)^2(t_3-1), \pm(t_1-1)(t_2-1)(t_3-1)^2.$$

Therefore the three-variable Alexander polynomial is $(t_1-1)(t_2-1)(t_3-1)$ (up to multiplication by a unit in $\mathbb{Z}[t_i^{\pm 1}]$).

Next we compute the 1-variable Alexander polynomial for Borromean rings, using the incidence matrix approach. An incidence matrix for the Borromean rings (with two columns corresponding to adjacent regions deleted) is the following:

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ -1 & -1 & -t & 0 & 0 & 0 \\ t & 1 & 0 & 0 & t & 0 \\ -t & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & t & 1 & t \\ 0 & 0 & -1 & -1 & 0 & -t \end{pmatrix}.$$

The Alexander polynomial is therefore $\det A = (t-1)^4$. This is as expected: for link with at least two components, the multivariable and single-variable Alexander polynomials are related by one additional factor of $(t-1)$. Explicitly,

$$\Delta_L(t) = \Delta_L(t, \ldots, t)(t-1),$$

where $\Delta_L(t_1, \ldots, t_n)$ is a multivariable Alexander polynomial. △

Exercise 9.3.2. (Exercise to myself) Does the sphere theorem hold in higher dimensions?

Solution: Consider the manifold $X_n = \mathbb{S}^{n-2} \times \mathbb{S}^2$ for $n \geq 5$. We prove that this manifold is a counter example to the higher dimensional analogue of the sphere theorem.

First we note that $\pi_{n-1}(X_n) \neq 0$. This is because

$$\pi_{n-1}(X_n) \cong \pi_{n-1}(\mathbb{S}^{n-2}) \times \pi_{n-1}(\mathbb{S}^2) \neq 0.$$

The second inequality comes from the fact that $\pi_{n-1}(\mathbb{S}^{n-2})$ is non-trivial (by calculations of stable homotopy groups).

Next we show that any embedding of $\mathbb{S}^{n-1}$ in X_n represents the trivial element in $\pi_{n-1}(X_n)$. We make use of the following result: *if the first homology of an n-manifold has no 2-torsion, then any compact connected submanifold of codimension 1 is separating.* The homology groups of X_n are as follows:

$$H_i(X_n) = \begin{cases} \mathbb{Z} & i \in \{0, 2, n-2, n\} \\ 0 & \text{otherwise.} \end{cases}$$

Since the first homology vanishes, any compact connected codimension 1 submanifold separates X_n. In particular, any embedding of $\mathbb{S}^{n-1}$ separates X_n.

By the definition of a connected sum, this embedding of $\mathbb{S}^{n-1}$ realises X_n as a connected sum of compact oriented n-manifolds A and B. Then for $i \in \{0, n\}$, $H_i(A) = H_i(B) = \mathbb{Z}$. For $0 < i < n$, it follows from the Mayer-Vietoris sequence that

$$H_i(X_n) = H_i(A) \oplus H_i(B).$$

Since $H_2(X_n) = \mathbb{Z}$, without loss of generality, $H_2(A) = 0$. By Poincaré duality and $H_{n-2}(X_n) = \mathbb{Z}$, it follows that $H_{n-2}(B) = \mathbb{Z}$ and $H_{n-2}(A) = 0$. Therefore A is a homology sphere. Finally because $\pi_1(X_n)$ is trivial, A is a simply connected homology sphere. By Whitehead's theorem A is then a homotopy sphere, and by the Poincaré conjecture, it is homeomorphic to $\mathbb{S}^n$. Therefore $\mathbb{S}^{n-1}$ contracts to a point on the "A side" of its embedding in X_n. This completes the counter example.

Note that our proof doesn't hold for $\mathbb{S}^2 \times \mathbb{S}^2$ without some more work. However, in that case it can be resolved by using the intersection form, and we obtain the same conclusion. △

Exercise 9.3.3. (Lickorish 11.8) Show that the genus of the (p, q) torus knot is $\frac{(p-1)(q-1)}{2}$.

Solution: We first show that the genus is at least $(p-1)(q-1)/2$. This follows from a calculation of the Alexander polynomial. Using Fox calculus, we see that

$$\Delta_{T_{p,q}} = \frac{(1-t)(1-t^{pq})}{(1-t^p)(1-t^q)}.$$

Therefore the breadth of the Alexander polynomial is $pq + 1 - p - q = (p-1)(q-1)$. But recall the bound

$$2g(T_{p,q}) \geq \text{breadth } \Delta_{T_{p,q}}.$$

Therefore the genus of $T_{p,q}$ is at least $(p-1)(q-1)/2$.

Conversely, to show that the genus is at most $(p-1)(q-1)/2$, we exhibit a Seifert surface with this genus. Draw the standard diagram of the torus knot $T_{p,q}$, and orient each of the p strands in the same direction. Then the Seifert algorithm gives a Seifert surface with:

$$f = p \text{ disks}, c = q(p-1) \text{ crossings}, n = 1 \text{ boundary component}.$$

Therefore the genus of the corresponding surface is

$$(2 + c - f - n)/2 = (p-1)(q-1)/2.$$

$\triangle$

Chapter 10

Applying knots to 3-manifolds

One of the powers of knot theory is that it sheds a lot of light on the study 3-manifolds. In this chapter we explore applications of knot theory to 3-manifolds. First we show that knots genuinely give rise to all possible 3-manifolds (via surgery along embedded knots in $\mathbb{S}^3$). Next we use this idea to construct invariants of 3-manifolds from the Jones polynomial.

10.1 All 3-manifolds arise from knots

To study 3-manifolds, we begin by studying automorphisms of surfaces. The following lemma shows that studying these automorphisms should shed light on 3-manifold topology:

Lemma 10.1.1. Suppose U, V are 3-manifolds with homeomorphic boundaries, and that $h_0, h_1 : \partial U \to \partial V$ are *isotopic* homeomorphisms. Then $U \sqcup_{h_0} V$ and $U \sqcup_{h_1} V$ are homeomorphic.

Two homeomorphisms are said to be *isotopic* if they are homotopic, and the homotopy is a homeomorphism at each t. The collection of all isotopy classes of automorphisms of a manifold forms a group under composition, called the *mapping class group*. We write $\mathrm{Mod}(M)$ to denote the isotopy classes of orientation preserving automorphisms of M, and $\mathrm{Mod}^{\pm}(M)$ to denote the isotopy classes of all automorphisms of M.

Definition 10.1.2. A *Dehn twist* is an automorphism of a surface F isotopic to the following map T:

- Let $A \subset F$ be an embedded annulus $\mathbb{S}^1 \times [0, 1]$.
- Define $T : F \to F$ to be the identity on $\Sigma - A$.
- Define T by $T(e^{i\theta}, t) = (e^{i(\theta - 2\pi t)}, t)$ on A.

An important result from the theory of mapping class groups of surfaces is that the mapping class group is finitely generated by Dehn twists.

Theorem 10.1.3 (Dehn-Lickorish theorem). *For $g \geq 0$, the mapping class group* $\mathrm{Mod}(S_g)$ *is generated by finitely many Dehn twists about non-separating simple closed curves.*

Here S_g denotes the closed surface of genus g. Excellent exposition on the mapping class group (and the above result) is available in Farb and Margalit [FM12]. In fact, the following improvement can be made:

Theorem 10.1.4 (Wajnryb). *For $g \geq 3$, the mapping class group of S_g (or S_g with one boundary component) is finitely presented by $2g+1$ generators (corresponding to Dehn twists). The relations are described in [FM12].*

For subsequent purposes, we don't need such powerful results. The following result is proven in a more elementary manner in Lickorish, but we prove it using the above theorem because it's so much shorter to write out!

Proposition 10.1.5. Let F be a surface with boundary. Let $p_1, \ldots, p_n$ be disjoint simple closed curves in the interior of F, the union of which doesn't separate F. Let $q_1, \ldots, q_n$ be another such family of disjoint simple closed curves. Then there is an automorphism h of F generated by Dehn twists such that $hp_i = q_i$ for each i.

Proof. We do not need to prove that h is generated by Dehn twists, since this follows from the Dehn-Lickorish theorem. It remains to prove that any h sending p_i to q_i exists. This follows from the *change of coordinates principle* which is also described in [FM12].

By the classification of surfaces, F cut along $\{p_i\}$ and F cut along $\{q_i\}$ (which we denote by F_p and F_q) are homeomorphic. In fact, since the boundary components corresponding to the p_i are pairwise disjoint, they can be inductively mapped to the boundaries corresponding to the q_i by orientation preserving homeomorphisms. Moreover, these can be chosen to descend to an automorphism of F sending each p_i to q_i. $\square$

This proposition applies to a surgery classification of 3-manifolds.

Definition 10.1.6. A *handlebody of genus g* is an orientable 3-manifold that is a 3-ball with g 1-handles attached.

A *handle* is visually similar to a connected sum with a solid torus. Formally, an r handle is defined as follows:

- Let M be an n-manifold. Let $e : \partial D^r \times D^{n-r} \to \partial M$ be an embedding. This is called the *framing* of the attaching sphere $\partial D^r \times 0$.
- *M with an r-handle attached* is the space

$$M \sqcup_f (D^r \times D^{n-r}).$$

A result from Morse theory is that every 3-manifold can be decomposed into a pair of handlebodies, called a Heegaard splitting.

Theorem 10.1.7. *Let M be a closed connected orientable 3-manifold. Then M admits a* Heegaard splitting*. That is, $M = H_g \cup H'_g$, where H_g, H'_g are handlebodies of genus g, and $H_g \cap H'_g = \partial H_g = \partial H'_g = S_g$.*

The main theorem of this section is the following result:

Theorem 10.1.8. *Any closed connected orientable 3-manifold M is obtained from $\mathbb{S}^3$ by a finite sequence of 1-surgeries. (That is, finitely many disjoint copies of $\mathbb{S}^1 \times D^2$ are replaced with $D^2 \times \mathbb{S}^1$.)*

Remark. $\mathbb{S}^3$ bounds a 4-manifold, namely the 4-ball D^4. Replacing $\mathbb{S}^1 \times D^2$ with $D^2 \times \mathbb{S}^1$ corresponds to attaching a 2-handle $D^2 \times D^2$ to D^4. (The attaching sphere is $\mathbb{S}^1 \times 0$.) Therefore the 3-manifold bounds a 4-dimensional handlebody.

Proof. We give a proof sketch. Let M be a closed connected orientable 3-manifold. We prove the following result: *there exists a finite collection of disjoint solid tori $T'_1, \ldots, T'_N$ in M and $T_1, \ldots, T_N$ in $\mathbb{S}^3$ such that $M - \bigcup \operatorname{int} T'_i$ and $\mathbb{S}^3 - \bigcup \operatorname{int} T_i$ are homeomorphic.*

Fix a Heegaard splitting $H_g \cup H'_g = M$. The two handlebodies H_g, H'_g are glued along a homeomorphism $h : \partial H_g \to \partial H'_g$. Fix curves $p_1, \ldots, p_g$ in ∂H_g which are each "looped through a hole in H_g" so that they are non-separating disjoint simple closed curves each bounding a disk. Similarly fix $q_1, \ldots, q_g$ in $\partial H'_g$ which are each "looped around a hole in H'_g". By the earlier proposition relating to the change of coordinates principle, there is a homeomorphism $h' : \partial H_g \to \partial H'_g$ sending each p_i to q_i. Moreover, $H_g \sqcup_{h'} H'_g$ is the 3-sphere.

We now have two homeomorphisms, h and h', with which we can glue the surfaces. Define $h(p_i) = q'_i$ for each i. Then the q'_i are disjoint simple closed curves on H'_g. Again by the previous proposition, there is a automorphism $\Phi : H'_g \to H'_g$ sending each q'_i to q_i, and in particular Φ is generated by Dehn twists. Since Dehn twists are supported on annuli, by deleting sufficiently many solid tori from M and $\mathbb{S}^3$, Φ restricts to the identity map. Then the restrictions of h and h' to the corresponding domains and codomains are isotopic, so gluing surfaces along either homeomorphism gives homeomorphic 3-manifolds. □

The theorem can be improved. The copies of $\mathbb{S}^1 \times D^2$ that are replaced in $\mathbb{S}^3$ are neighbourhoods of links. To uniquely determine the the 3-manifold obtained from surgery along a link, we simply have to specify the link along with a *framing*. This is an integer, namely the self-linking number of each component given the parametrisations of $\mathbb{S}^1 \times D^2$.

By a result of Kirby, any two framed links describing the same 3-manifold are related by *Kirby moves* and Reidemeister moves. (Note that such framed links are called surgery diagrams.) The Kirby moves are of the following two types:

- Type 1: addition or removal of an unknotted component to the diagram, with framing ± 1.

- Type 2: replacing components C_1, C_2 with $C_1 + C_2', C_2$ (where C_2' is a push-off of C_2), and giving $C_1 + C_2'$ an updated framing.

A type 1 move is called a *blow up* or *blow down* and corresponds to a connected sum with $\pm\mathbb{CP}^2$ (in the underlying 4 manifold whose boundary is our 3-manifold of interest). A type 2 move corresponds to a *handle slide*, namely sliding the attaching sphere of one handle over another.

In the same way that ambient isotopy classes of links can be understood with link diagrams related by Reidemeister moves, we can understand closed connected orientable 3-manifolds by surgery diagrams related by Reidemeister moves and Kirby moves.

10.2 The origin of quantum $SU_q(2)$ invariants

In this section we give a non-rigorous introduction to quantum $SU_q(2)$ invariants of 3-manifolds. This is largely based on [Wit13]. In the next section, we will see how these arise from knot theory.

Quantum field theory can be formalised in terms of a *path integral.* The basic ingredient is an expression

$$A(\varphi_0, \varphi_1) = \int_{\Phi_0=\varphi_0,\Phi_1=\varphi_1} e^{iS[\Phi]}\, D[\Phi].$$

On the left, $A(\varphi_0, \varphi_1)$ is the probability amplitude for a state to evolve from φ_0 to φ_1. The right side is an integral taken over all field configurations with boundary data φ_0 and φ_1. Within the integrand, we have some measure $D[\Phi]$, and an action $S[\Phi]$. This is not well defined: the integral is taken over some typically infinite dimensional space, so there is no reason to expect a measure $D[\Phi]$ to exist.

Based on this formalism for amplitudes, also called quantum propagators, we can also compute correlation functions of observables. Specifically, if O_i are observables, then their correlation function is given by the path integral

$$\int e^{iS[\Phi]} \prod_i O_i(\Phi) D[\Phi].$$

A *quantum field theory* is therefore determined by the action functonal, and the space on which fields exist. (That is, the underlying manifold on which we consider bundles - fields are sections of these bundles.)

Example. A famous quantum field theory is *Chern-Simons theory.* The Chern-Simons functional is defined by

$$CS(A) = \int_M \mathrm{tr}(A \wedge dA + \frac{2}{3} A \wedge A \wedge A)$$

where A is a connection one-form defined on an oriented compact 3-manifold M. Witten showed that path integrals using the Chern-Simons functional can actually be computed (although lacking in mathematical rigour since we don't really know what the underlying measure is, as remarked earlier). Specifically, it was established that

$$V_K^{n-1}(q) = \int e^{ikCS(A)} \mathrm{tr}\rho_n(\mathrm{Hol}_K(A))\, dA,$$

where:

- $V_K^n(q)$ is the coloured Jones polynomial of K, evaluated at q. Thus V_K^1 is the usual Jones polynomial, and V_K^n is the Jones polynomial of an n-cable link.
- The variable q is equal to $e^{\frac{2\pi i}{n+k}}$. By fixing n and letting k vary, this gives an infinite number of points, and hence determines the n-coloured Jones polynomial.
- The integral is taken over *connection 1-forms on $SU(2)$-bundles over $\mathbb{S}^3$, modulo gauge.*
- As remarked earlier, $CS(A)$ is the Chern-Simons functional.
- The "observable" $\mathrm{tr}\rho_n(\mathrm{Hol}_K(A))$ consists of three parts.
  - given our choice of K, $\mathrm{Hol}_K(A)$ is the holonomy of the connection form A about the knot K.
  - $\rho_n : \mathrm{SU}(2) \to \mathrm{End}(V)$ is an irreducible representation of dimension n. (There is a unique such ρ for each n.)
  - The observable is then the trace of the matrix $\rho_n(\mathrm{Hol}_K(A))$.
- The measure is elusive!

On one hand, this doesn't formally make sense because it isn't clear how to integrate over a space of connection forms. However, this formalism gives rise to an invariant *of the sphere*; the base space over which we consider connection forms. We could just as well replace this with another oriented 3-manifold to obtain another function. That is, this construction gives a heuristic for an invariant of 3-manifolds which ties into knot theory. This is what we call the quantum $SU_q(2)$ invariant, or Witten-Reshetikhin–Turaev invariant. We now describe how to rigorously obtain the invariant from knot theory.

10.3 Quantum $SU_q(2)$ invariants from knot invariants

Since we can express closed connected orientable 3-manifolds by surgery diagrams, which are just link diagrams decorated with integers, it should be possible to extend link invariants to 3-manifold invariants. The extra condition is to verify invariance under Kirby moves.

Proposition 10.3.1. The equivalence class of the *linking matrix* of a surgery diagram is a 3-manifold invariant.

A linking matrix of M is the square symmetric matrix $A_{ij} = \mathrm{lk}(L_i, L_j)$, where L_i are the components of a surgery diagram of M. (The self-linking number $\mathrm{lk}(L_i, L_i)$ is the framing of each component.) To prove that the matrix is an invariant of M, we simply note that it is a *presentation matrix* for the first homology of M. (Therefore the *equivalence class* of the matrix in the sense of equivalences of presentation matrices) is an invariant of M.

Remark. This invariant is boring! The first homology is already understood in less round-about ways.

Next we aim to describe Witten's so-called *quantum $SU_q(2)$ invariants* via knot theory. The main machinery that must be developed is *linear skein theory*.

Definition 10.3.2. Let F be an oriented surface (possibly with boundary), and fix a complex number A. To define a *linear skein* of F, fix a finite number of points (possibly zero) on the boundary of F. A *link diagram* is defined to be a union of arcs and closed curves (with crossing data), so that any arcs necessarily have endpoints on specific points of ∂F. Conversely, every specified point must be the endpoint of an arc. Two diagrams are considered the same if they differ by a homeomorphism of F fixing ∂F pointwise.

The *linear skein* of F, denoted $\mathcal{S}(F)$, is the vector space over $\mathbb{C}$ of formal linear sums of unoriented link diagrams in F, modulo the relations

- $D \sqcup$ trivial closed curve $= (-A^2 - A^{-2})D$,
- $\diagup\!\!\!\!\diagdown = A \ )(+ A^{-1} \asymp$.

Remark. The definition above is clearly inspired by the Kauffman bracket, and will eventually provide the quantum invariant we wish to construct. Therefore our eventual invariant can be thought of as a modification of the Jones polynomial.

Example. $\mathcal{S}(\mathbb{R}^2) \cong \mathbb{C}$. To see this, fix any A, and consider

$$v = \sum_{i=1}^{n} \lambda_i D_i \in \mathcal{S}(\mathbb{R}^2).$$

By the second relation above, all crossings can be inductively removed, so that v is expressed as a sum of scalar multiples of unlinks. But any given unlink with m components is just a complex multiple of U by the first relation, where U is a diagram of the unknot. Therefore we can write $v = \lambda U$ for some $\lambda \in \mathbb{C}$. This shows that $\mathcal{S}(\mathbb{R}^2)$ is a one dimensional vector space over $\mathbb{C}$. Conversely, it is at least one dimensional, since U is non-zero. Therefore $\mathcal{S}(\mathbb{R}^2) \cong \mathbb{C}$ as required.

Example. $\mathcal{S}(\mathbb{S}^2) \cong \mathbb{C}$. An isomorphism between $\mathcal{S}(\mathbb{R}^2)$ and $\mathcal{S}(\mathbb{S}^2)$ is induced by the inclusion of $\mathbb{R}^2$ into $\mathbb{S}^2 = \mathbb{R}^2 \cup \{\infty\}$.

Example. $\mathcal{S}(\mathbb{S}^1 \times [0,1]) \cong \mathbb{C}[\alpha]$. Here $\mathbb{S}^1 \times [0,1]$ is a surface with boundary, but we don't distinguish any points in the boundary so that the diagrams consist of unions of closed curves. It can be shown that a basis of $\mathcal{S}(\mathbb{S}^1 \times [0,1])$ consists of curves encircling the cylinder. That is, $\{\alpha^0, \alpha^1, \alpha^2, \dots\}$, where α^i denotes the diagram with i simple closed curves encircling the cylinder.

There is a well defined product $\mathcal{S}(\mathbb{S}^1 \times [0,1]) \times \mathcal{S}(\mathbb{S}^1 \times [0,1]) \to \mathcal{S}(\mathbb{S}^1 \times [0,1])$ which can be thought of as gluing F to itself along boundary components. Then $\alpha^i \alpha^j = \alpha^{i+j}$, and $\mathcal{S}(\mathbb{S}^1 \times [0,1])$ equipped with this product becomes a commutative algebra. Moreover, it is then clear that $\mathcal{S}(\mathbb{S}^1 \times [0,1]) \cong \mathbb{C}[\alpha]$.

Example. $\mathcal{S}(D^2, 2n) \cong TL_n$. Here TL_n denotes the nth *Temperley-Lieb algebra*. A basis for $\mathcal{S}(D^2, 2n)$ consists of all diagrams with no crossings and no closed curves. (There are $\binom{2n}{n}/(n+1)$ such diagrams.) We denote D^2 with a square, with n distinguished points on the right edge and n on the left. Again the "gluing surfaces" operation gives a bilinear product (though not commutative). The identity element $\mathbf{1}$ is n parallel arcs from left to right. TL_n is generated by n elements $\mathbf{1}, e_1, \dots, e_n$ as an algebra. (Each e_i consists of the first $i-1$ and last $n-i-1$ points on each side of D^2 joined by horizontal arcs, with the points $i, i+1$ on each side joined by arcs with endpoints on the same edge.)

To better understand the Temperley-Lieb algebra, we introduce the *Jones-Wenzl idempotent* $f^{(n)} \in TL_n$. This is also called the *magic element*, and appears in many arguments regarding the Temperley-Lieb algebra.

Since $f^{(n)}$ belongs to $\mathcal{S}(D^2, 2n)$, it can be placed in the place. The n endpoints of arcs on either end of $f^{(n)}$ can be joined together in the canonical way by looping the endpoints over the "top" of the diagram (given a left and right of the diagram in D^2). This is the image of $f^{(n)}$ under a canonical linear map $TL_n \to \mathbb{C}$, and the image is denoted by Δ_n.

Definition 10.3.3. The *Jones-Wenzl idempotent* $f^{(n)} \in TL_n$ is characterised by the following properties, given that $A^4 \in \mathbb{C}$ is not a kth root of unity for $k \leq n$.

- $f^{(n)} e_i = 0 = e_i f^{(n)}$ for $i \in \{1, \dots, n-1\}$.
- $f^{(n)} - \mathbf{1}$ is generated by $\{e_1, \dots, e_{n-1}\}$.
- $f^{(n)} f^{(n)} = f^{(n)}$.
- $\Delta_n = (-1)^n (A^{2(n+1)} - A^{-2(n+1)})/(A^2 - A^{-2})$.

A proof is not given here, but we explore the last bullet point. Consider placing $(D^2, 2n)$ in an annulus and joining the n points on each side around the annulus to obtain a map $TL_k \to \mathcal{S}(\mathbb{S}^1 \times [0,1]) = \mathbb{C}[\alpha]$ for each $k \in \{0, \dots, n\}$. The image of $f^{(k)}$ under this map is some polynomial $S_k(\alpha)$ in α. By idempotence of f, one can show that this polynomial is given recursively by

$$S_{k+1}(\alpha) = \alpha S_k(\alpha) - S_{k-1}(\alpha).$$

The initial conditions are determined by $f^{(0)}$ being the empty diagram, and $f^{(1)} = \mathbf{1}$. Therefore $S_0(\alpha) = \alpha^0$ and $S_1(\alpha) = \alpha$. This is called the *Chebyshev polynomial*. By composing the map $TL_n \to \mathcal{S}(\mathbb{S}^1 \times [0,1])$ with the canonical map $\mathcal{S}(\mathbb{S}^1 \times [0,1]) \to \mathcal{S}(\mathbb{R})$, we obtain the map $f^{(n)} \to \Delta_n$ which can be verified to satisfy the desired relation above via the recurrence relation of S_k.

Definition 10.3.4. The *Chebyshev polynomial* S_n is defined by

$$S_{n+1}(x) = xS_n(x) - S_{n-1}(x), \quad S_0(x) = 1, S_1(x) = x.$$

This immediately satisfies two important properties:

$$S_n(x) = (-1)^n S_n(-x), \quad (t - t^{-1})S_n(t + t^{-1}) = t^{n+1} - t^{-(n+1)}.$$

Definition 10.3.5. For any $r \in \mathbb{Z}$, define $\omega \in \mathcal{S}(\mathbb{S}^1 \times [0,1])$ by

$$\omega = \sum_{n=0}^{r-2} \Delta_n S_n(\alpha).$$

This is one of the final ingredients for defining our quantum $\mathrm{SU}_q(2)$ invariants. Again we remark that $S_n(\alpha)$ is the image of $f^{(n)}$ under the canonical map $TL_n \to \mathbb{C}[\alpha]$, and Δ_n is the image of $f^{(n)}$ under the map $TL_n \to \mathbb{C}$.

Definition 10.3.6. Let D be a diagram of n ordered components. An n-ary multilinear map

$$\langle -, \ldots, - \rangle_D : \mathcal{S}(\mathbb{S}^1 \times [0,1])^n \to \mathcal{S}(\mathbb{R}^2)$$

is defined as follows:

- Consider an immersed annular neighbourhood A_i of each component c_i of D. Any self-intersections of A_i inherit the crossing data from c_i. Moreover, any A_i typically intersects some A_j, at which point they again inherit crossing data from D.
- Given an ith diagram in $\mathcal{S}(\mathbb{S}^1 \times [0,1])$, consider the image of this diagram under the above immersion. Record any new crossing data.

Example. Let U_+ and U_- be oriented planar figure-8 diagrams (with a positive crossing and negative crossing respectively). Then $\langle \alpha \rangle_{U_\pm}$ corresponds to taking a diagram α and applying a "global type 1 Reidemeister move".

Theorem 10.3.7. *Let M be a closed oriented 3-manifold with surgery diagram D. Let b_+ be the number of positive eigenvalues and b_- the number of negative eigenvalues of the linking matrix of D. If A is a primitive $4r$th root of unity (for $r \geq 3$), the expression*

$$\langle \omega, \ldots, \omega \rangle_D \langle \omega \rangle_{U_+}^{-b_+} \langle \omega \rangle_{U_-}^{-b_-}$$

is the quantum $SU(2)$ invariant *of M.*

Proof. We do not give a proof, but briefly describe the basic ideas. We require the above expression to be invariant under Reidemeister moves and Kirby moves of D. The first term $\langle \omega, \ldots, \omega \rangle_D$ is invariant under type 2 moves, i.e. handle slides. The additional two terms at the end add invariance under type 1 Kirby moves (blow ups and blow downs) (analogously to the way we multiply the Kauffman bracket by the writhe to obtain invariance under type 1 Reidemister moves). Reidemeister moves correspond to isotopies of D, under which the expression is also invariant. $\square$

References

[FM12] Benson Farb and Dan Margalit. *A Primer on Mapping Class Groups*. Princeton University Press, 2012.

[Lic97] W.B. Raymond Lickorish. *An Introduction to Knot Theory*. Graduate Texts in Mathematics. Springer-Verlag New York Berlin Heidelberg, 1997.

[Wit13] Edward Witten. *A New Look at the Jones Polynomial of a Knot*. `https://ncatlab.org/nlab/files/WittenOnJonesPolynomial2013.pdf`, 2013. Clay Conference, Oxford.