Graduate Texts in Mathematics 175

Springer
New York
Berlin
Heidelberg
Barcelona
Budapest
Hong Kong
London
Milan
Paris
Santa Clara
Singapore
Tokyo

Graduate Texts in Mathematics

1 TAKEUTI/ZARING. Introduction to Axiomatic Set Theory. 2nd ed.
2 OXTOBY. Measure and Category. 2nd ed.
3 SCHAEFER. Topological Vector Spaces.
4 HILTON/STAMMBACH. A Course in Homological Algebra. 2nd ed.
5 MAC LANE. Categories for the Working Mathematician.
6 HUGHES/PIPER. Projective Planes.
7 SERRE. A Course in Arithmetic.
8 TAKEUTI/ZARING. Axiomatic Set Theory.
9 HUMPHREYS. Introduction to Lie Algebras and Representation Theory.
10 COHEN. A Course in Simple Homotopy Theory.
11 CONWAY. Functions of One Complex Variable I. 2nd ed.
12 BEALS. Advanced Mathematical Analysis.
13 ANDERSON/FULLER. Rings and Categories of Modules. 2nd ed.
14 GOLUBITSKY/GUILLEMIN. Stable Mappings and Their Singularities.
15 BERBERIAN. Lectures in Functional Analysis and Operator Theory.
16 WINTER. The Structure of Fields.
17 ROSENBLATT. Random Processes. 2nd ed.
18 HALMOS. Measure Theory.
19 HALMOS. A Hilbert Space Problem Book. 2nd ed.
20 HUSEMOLLER. Fibre Bundles. 3rd ed.
21 HUMPHREYS. Linear Algebraic Groups.
22 BARNES/MACK. An Algebraic Introduction to Mathematical Logic.
23 GREUB. Linear Algebra. 4th ed.
24 HOLMES. Geometric Functional Analysis and Its Applications.
25 HEWITT/STROMBERG. Real and Abstract Analysis.
26 MANES. Algebraic Theories.
27 KELLEY. General Topology.
28 ZARISKI/SAMUEL. Commutative Algebra. Vol.I.
29 ZARISKI/SAMUEL. Commutative Algebra. Vol.II.
30 JACOBSON. Lectures in Abstract Algebra I. Basic Concepts.
31 JACOBSON. Lectures in Abstract Algebra II. Linear Algebra.
32 JACOBSON. Lectures in Abstract Algebra III. Theory of Fields and Galois Theory.
33 HIRSCH. Differential Topology.
34 SPITZER. Principles of Random Walk. 2nd ed.
35 WERMER. Banach Algebras and Several Complex Variables. 2nd ed.
36 KELLEY/NAMIOKA et al. Linear Topological Spaces.
37 MONK. Mathematical Logic.
38 GRAUERT/FRITZSCHE. Several Complex Variables.
39 ARVESON. An Invitation to C^*-Algebras.
40 KEMENY/SNELL/KNAPP. Denumerable Markov Chains. 2nd ed.
41 APOSTOL. Modular Functions and Dirichlet Series in Number Theory. 2nd ed.
42 SERRE. Linear Representations of Finite Groups.
43 GILLMAN/JERISON. Rings of Continuous Functions.
44 KENDIG. Elementary Algebraic Geometry.
45 LOÈVE. Probability Theory I. 4th ed.
46 LOÈVE. Probability Theory II. 4th ed.
47 MOISE. Geometric Topology in Dimensions 2 and 3.
48 SACHS/WU. General Relativity for Mathematicians.
49 GRUENBERG/WEIR. Linear Geometry. 2nd ed.
50 EDWARDS. Fermat's Last Theorem.
51 KLINGENBERG. A Course in Differential Geometry.
52 HARTSHORNE. Algebraic Geometry.
53 MANIN. A Course in Mathematical Logic.
54 GRAVER/WATKINS. Combinatorics with Emphasis on the Theory of Graphs.
55 BROWN/PEARCY. Introduction to Operator Theory I: Elements of Functional Analysis.
56 MASSEY. Algebraic Topology: An Introduction.
57 CROWELL/FOX. Introduction to Knot Theory.
58 KOBLITZ. p-adic Numbers, p-adic Analysis, and Zeta Functions. 2nd ed.
59 LANG. Cyclotomic Fields.
60 ARNOLD. Mathematical Methods in Classical Mechanics. 2nd ed.

continued after index

W.B. Raymond Lickorish

An Introduction to Knot Theory

With 114 Illustrations

Springer

W.B. Raymond Lickorish
Professor of Geometric Topology, University of Cambridge,
and Fellow of Pembroke College, Cambridge
Department of Pure Mathematics and Mathematical Statistics
Cambridge CB2 1SB
England

Mathematics Subject Classification (1991): 57-01, 57M25, 16S34, 57M05

ISBN 0-387-98254-X Springer-Verlag New York Berlin Heidelberg SPIN 10628672

Preface

This account is an introduction to mathematical knot theory, the theory of knots and links of simple closed curves in three-dimensional space. Knots can be studied at many levels and from many points of view. They can be admired as artifacts of the decorative arts and crafts, or viewed as accessible intimations of a geometrical sophistication that may never be attained. The study of knots can be given some motivation in terms of applications in molecular biology or by reference to parallels in equilibrium statistical mechanics or quantum field theory. Here, however, knot theory is considered as part of geometric topology. Motivation for such a topological study of knots is meant to come from a curiosity to know how the geometry of three-dimensional space can be explored by knotting phenomena using precise mathematics. The aim will be to find invariants that distinguish knots, to investigate geometric properties of knots and to see something of the way they interact with more adventurous three-dimensional topology. The book is based on an expanded version of notes for a course for recent graduates in mathematics given at the University of Cambridge; it is intended for others with a similar level of mathematical understanding. In particular, a knowledge of the very basic ideas of the fundamental group and of a simple homology theory is assumed; it is, after all, more important to know about those topics than about the intricacies of knot theory.

There are other works on knot theory written at this level; indeed most of them are listed in the bibliography. However, the quantity of what may reasonably be termed mathematical knot theory has expanded enormously in recent years. Much of the newly discovered material is not particularly difficult and has a right to be included in an introduction. This makes some of the excellent established treatises seem a little dated. However, concentrating entirely on developments of the past decade gives a most misleading view of the subject. An attempt is made here to outline some of the highlights from throughout the twentieth century, with a little bias towards recent discoveries.

The present size of the subject means that a choice of topics must be made for inclusion in any first course or book of reasonable length. Such selection must be subjective. An attempt has been made here to give the flavour and the results from three or four main techniques and not to become unduly enmeshed in any of them.

Firstly, there is the three-manifold method of manipulating surfaces, using the pattern of simple closed curves in which two surfaces intersect. This leads to the theorem concerning the unique factorisation of knots into primes and to the theory concerning the primeness of alternating diagrams. Combinatorics applied to knot and link diagrams lead (by way of the Kauffman bracket) to the Jones polynomial, an invariant that is good, but not infallible, at distinguishing different knots and links. This invariant also has applications to the way diagrams of certain knots might be drawn. Next, techniques of elementary homology theory are used on the infinite cyclic cover of the complement of a link to lead to the "abelian" invariants, in particular to the well known Alexander polynomial. That is reinforced by the association of that polynomial invariant with the Conway polynomial, as well as by a study of the fundamental group of a link's complement. The use of (framed) links to describe, by means of "surgery", any closed orientable three-manifold is explored. Together with the skein theory of the Kauffman bracket, this idea leads to some "quantum" invariants for three-manifolds. A technique, belonging to a more general theory of three-manifolds, that will *not* be described is that of the W. Haken's classification of knots. That technique gives a theoretical algorithm which always decides if two knots are or are not the same. It is *almost* impossible to use it, but it is good to know it exists [42].

One can take the view that the object of mathematics is to *prove* that certain things are true. That object will here be pursued. A declaration that something is true, followed by copious calculations that produce no contradiction, should not completely satisfy the intellect. However, even neglecting all logical or philosophical objections to this quest, there are genuine practical difficulties in attempting to give a totally self-contained introduction to knot theory. To avoid pathological possibilities, in which diagrams of links might have infinitely many crossings, it is necessary to impose a piecewise linear or differential restriction on links. Then all manoeuvres must preserve such structures, and the technicalities of a piecewise linear or differential theory are needed. One needs, for example, to know that any two-dimensional sphere, smoothly or piecewise linearly embedded in Euclidean three-space, bounds a smooth or piecewise linear ball. This is the Schönflies theorem; the existence of wild horned spheres shows it is not true without the technical restrictions. What is needed, then, is a full development of the theory of piecewise linear or differential manifolds at least up to dimension three. Laudable though such an account might be, experience suggests that it is initially counter-productive in the study of knot theory. Conversely, experience of knot theory can produce the incentive to understand these geometric foundations at a later time. Thus some basic (intuitively likely) results of piecewise linear theory will sometimes be quoted, sometimes with a sketch of how they are proved. Perhaps here piecewise linear theory has an advantage over differential theory, because up to dimension three, simplexes are readily visualisable; but differential theory, if known, will answer just as well. That apologia underpins the start of the theory. Significant direct quotations of results have however also been made in the discussion of the fundamental group of a link complement. That topic has been treated extensively elsewhere, so the remarks here are intended to be but something of a little survey.

Also quoted is R. C. Kirby's theorem concerning moves between surgery links for a three-manifold. Furthermore, at the end of a section extensions of a theory just considered are sometimes outlined without detailed proof. Otherwise it is intended that everything should be proved!

W. B. Raymond Lickorish

Contents

1

A Beginning for Knot Theory

The mathematical theory of knots is intended to be a precise investigation into the way that 1-dimensional "string" can lie in ordinary 3-dimensional space. A glance at the diagrams on the pages that follow indicates the sort of complication that is envisaged. Because the theory is intended to correspond to reality, it is important that initial definitions, whilst being precise, exclude unwanted pathology both in the things being studied and in the properties they might have. On the other hand, obsessive concentration on basic geometric technology can deter progress. It can initially be but tasted if it seem onerous. At its foundations, knot theory will here be considered as a branch of topology. It is, at least initially, not a very sophisticated application of topology, but it benefits from topological language and provides some very accessible illustrations of the use of the fundamental group and of homology groups.

As is customary, $\mathbb{R}^n$ will denote n-dimensional Euclidean space and S^n will be the n-dimensional sphere. Thus S^n is the unit sphere in $\mathbb{R}^{n+1}$, but it can be regarded as being $\mathbb{R}^n$ together with an extra point at infinity. There is a linear or affine structure on $\mathbb{R}^n$; it contains lines and planes and r-simplexes (r-dimensional analogues of intervals, triangles and tetrahedra). S^n can also be regarded as the boundary of a standard $(n + 1)$-simplex, so that S^n is then triangulated with the structure of a simplicial complex bounding a triangulated $(n + 1)$-ball B^{n+1}. Sometimes it seems more natural to describe B^{n+1} as a disc; it is then denoted D^{n+1}.

Definition 1.1. A link L of m components is a subset of S^3, or of $\mathbb{R}^3$, that consists of m disjoint, piecewise linear, simple closed curves. A link of one component is a knot.

The piecewise linear condition means that the curves composing L are each made up of a finite number of straight line segments placed end to end, "straight" being in the linear structure of $\mathbb{R}^3 \subset \mathbb{R}^3 \cup \infty = S^3$ or, alternatively, in the structure of one of the 3-simplexes that make up S^3 in a triangulation. In practice, when drawing diagrams of knots or links it is assumed that there are so very many straight line segments that the curves appear pretty well rounded. This insistence on having a *finite number* of straight line segments prevents a link from having an infinite number of kinks, getting ever smaller as they converge to a point (those links are called "wild"). An alternative way of avoiding wildness is to require that L be a smooth 1-dimensional submanifold of the smooth 3-manifold S^3. That leads to an equivalent theory, but in these low dimensions simplexes are often easier to manipulate than are sophisticated theorems of differential manifolds. Thus a piecewise linear condition applies to practically everything discussed here, but it will be given as little emphasis as possible.

Definition 1.2. Links L_1 and L_2 in S^3 are equivalent if there is an orientation-preserving piecewise linear homeomorphism $h : S^3 \to S^3$ such that $h(L_1) = (L_2)$.

Here the piecewise linear condition means that after subdividing the simplexes in each copy of S^3 into possibly very many smaller simplexes, h maps simplexes to simplexes in a linear way. Soon, equivalent links will be regarded as being the same link; in practice this causes no confusion. If the links are oriented or their components are ordered, h may be required to preserve such attributes. It is a basic theorem of piecewise linear topology that such an h is *isotopic* to the identity. This means there exist $h_t : S^3 \to S^3$ for $t \in [0, 1]$ so that $h_0 = 1$ and $h_1 = h$ and $(x, t) \mapsto (h_t x, t)$ is a piecewise linear homeomorphism of $S^3 \times [0, 1]$ to itself. Thus certainly the *whole* of S^3 can be continuously distorted, using the homeomorphism h_t at time t, to move L_1 to L_2. An inept attempt to define equivalence in terms of moving one *subset* until it becomes the other could misguidedly permit knots to be pulled tighter and tighter until any complication disappears at a single point. If L_1 and L_2 are equivalent, their complements in S^3 are, of course, homeomorphic 3-dimensional manifolds. Thus it is reasonable to try to distinguish links by applying any topological invariant (for example, the fundamental group) to such complements. Similarly, any facet of the extensive theory of 3-dimensional manifolds can be applied to link complements; the theory of knots and links forms a fundamental source of examples in 3-manifold theory. It has recently been proved, at some length [37], that two *knots* with homeomorphic oriented complements are equivalent; that is not true, in general, for links of more than one component (a fairly easy exercise).

An elementary method of changing a link L in $\mathbb{R}^3$ to an equivalent link is to find a planar triangle in $\mathbb{R}^3$ that intersects L in exactly one edge of the triangle, delete that edge from L, and replace it by the other two edges of the triangle. See Figure 1.1. It can be shown that if two links are equivalent, they differ by a finite sequence of such moves or the inverses of such moves (replace two edges of a triangle by the other one). This result will be assumed; any proof would have to penetrate the technicalities of piecewise linear theory (a proof can be found in [17]).

Using such (possibly very small) moves, L can easily be changed so that it is in general position with respect to the standard projection $p : \mathbb{R}^3 \to \mathbb{R}^2$. Here this means that each line segment of L projects to a line segment in $\mathbb{R}^2$, that the

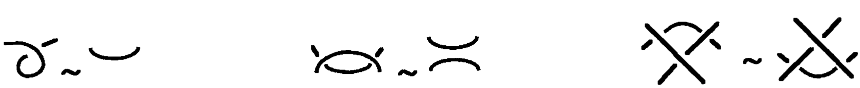
Figure 1.1

projections of two such segments intersect in at most one point which for disjoint segments is not an end point, and that no point belongs to the projections of three segments. Given such a situation, the image of L in $\mathbb{R}^2$ together with "over and under" information at the crossings is called a *link diagram* of L. Of course, a crossing is a point of intersection of the projections of two line segments of L; the "over and under" information refers to the relative heights above $\mathbb{R}^2$ of the two inverse images of a crossing. This information is always indicated in pictures by breaks in the under-passing segments.

If L_1 and L_2 are equivalent, they are related by a sequence of triangle moves as described above. After moving all the vertices of all the triangles by a very small amount, it can be assumed that the projections of no three of the vertices lie on a line in $\mathbb{R}^2$ and the projections of no three edges pass through a single point. Then each triangle projects to a triangle, and one can analyse the effect on link diagrams of each triangle move. One of the more interesting possibilities is shown in Figure 1.2.

Figure 1.2

With a little careful thought, it follows that any two diagrams of equivalent links L_1 and L_2 are related by a sequence of *Reidemeister moves* and an orientation-preserving homeomorphism of the plane. The Reidemeister moves are of three types, shown below in Figure 1.3; each replaces a simple configuration of arcs and crossings in a disc by another configuration. A move of Type I inserts or deletes a "kink" in the diagram; moves of Type III preserve the number of crossings. Any homeomorphism of the plane must, of course, preserve all crossing information.



Type I Type II Type III

Figure 1.3

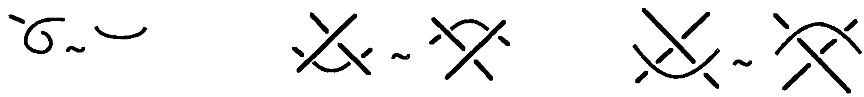


Figure 1.4

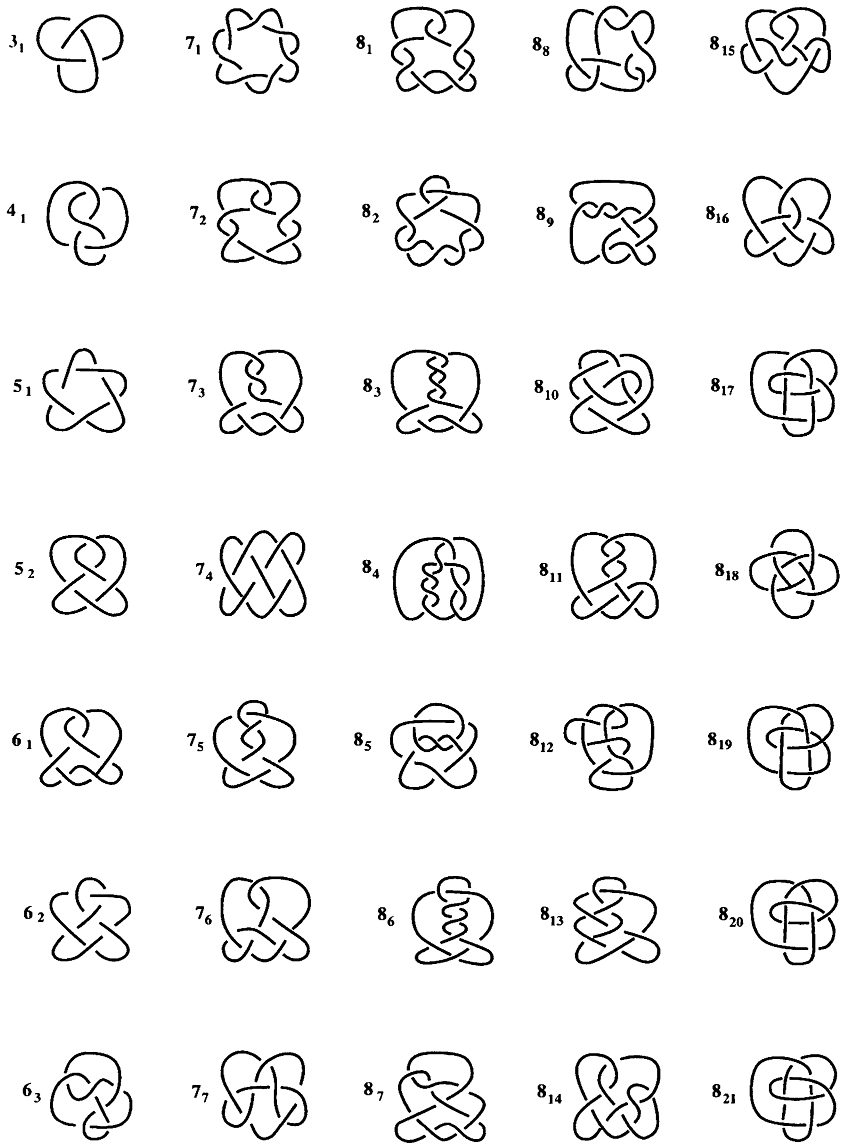
Figure 1.5

The "moves" shown in Figure 1.4 can be seen (exercise) to be consequences of the three types of Reidemeister move.

If the point at infinity is added to $\mathbb{R}^2$, so that all moves and diagrams are now regarded as being in S^2, then the "moves" of Figure 1.5 are combinations of Reidemeister moves of types two and three only (an easy exercise). Diagrams related by moves of Type II and Type III only are sometimes said to be *regularly isotopic*. It will always be assumed that S^3 and $\mathbb{R}^3$ are oriented. The components of an n-component link can be oriented in 2^n ways, and a choice of orientation, indicated by arrows on a diagram, is extra information that may or may not be given. If K is an oriented knot, the *reverse* of K—denoted $\mathrm{r}K$—is the same knot as a set but with the other orientation. Often K and $\mathrm{r}K$ are equivalent. If L is a link in S^3 and $\rho : S^3 \to S^3$ is an orientation-reversing piecewise linear homeomorphism, then $\rho(L)$ is a link called the *obverse* or *reflection* of L. Up to equivalence of $\rho(L)$, the choice of ρ is immaterial; $\rho(L)$ is denoted $\overline{L}$. Regarding S^3 as $\mathbb{R}^3 \cup \infty$, one can take ρ to be the map $(x, y, z) \mapsto (x, y, -z)$, and then it is clear that a diagram for $\overline{L}$ is the same as one for L but with all the over-passes changed to under-passes. As will later become clear, sometimes L and $\overline{L}$ are equivalent, sometimes they are not. There do exist oriented knots (the knot named 9_{32} is an example) for which K, $\mathrm{r}K$, $\overline{K}$ and $\overline{\mathrm{r}K}$ are four distinct oriented knots.

A knot K is said to be the *unknot* if it bounds an embedded piecewise linear disc in S^3. Triangle moves across the 2-simplexes of a triangulation of such a disc show that the unknot is equivalent to the boundary of a single 2-simplex linearly embedded in S^3, and hence it has (as expected) a diagram with no crossing at all. Two oriented knots K_1 and K_2 can be added together to form their sum $K_1 + K_2$ by a method that corresponds to the intuitive idea of tying one and then the other in the same piece of string; see Figure 1.6. More precisely, regard K_1 and K_2 as being in distinct copies of S^3, remove from each S^3 a (small) ball that meets the given knot in an unknotted spanning arc (one where the ball-arc pair is piecewise linearly homeomorphic to the product of an interval with a disc-point pair), and then identify together the resulting boundary spheres, and their intersections with the knots, so that all orientations match up. Some basic piecewise linear theory

TABLE 1.1. The Knot Table to Eight Crossings



shows that balls meeting the knots in unknotted spanning arcs are essentially unique, so that the addition of oriented knots is (up to equivalence, of course) well defined. It is immediate that this addition is commutative, and it is easily seen to be associative. The unknot is a zero for this addition, but it will be seen a little later that no knot other than the unknot has an additive inverse.

Figure 1.6

Definition 1.3. A knot K is a prime knot if it is not the unknot, and $K = K_1 + K_2$ implies that K_1 or K_2 is the unknot.

(Whereas "irreducible" might be a better term than "prime", this is traditional terminology, and it transpires that prime knots do have the usual algebraic property of primeness.)

Fairly simple knots can be defined by drawing diagrams, and to refuse to do this would be pedantic in the extreme. The *crossing number* of a knot is the minimal number of crossings needed for a diagram of the knot. Table 1.1 is a table of diagrams of all knots with crossing number at most 8. There are 35 such knots. Following traditional expediency, the unknot is omitted, only prime knots are included and *all* orientations are neglected (so that each diagram represents one, two or four oriented knots in oriented S^3 by means of the above operations r and ρ). A notation such as "8_5" beside a diagram simply means that it shows the fifth knot with crossing number 8 in a traditional ordering (begun in the nineteenth century by P. G. Tait [118] and C. N. Little [92]). Such terminology and tables of diagrams exist for knots up to eleven crossings. It is easy to tabulate knot diagrams and, for low numbers of crossings, to be confident that a list is complete; the difficulty comes in proving that the entries are prime and that the tabulation contains no duplicates. This is accomplished by associating to a knot some "invariant"—a well-defined mathematical entity such as a a number, a polynomial, or a group—and proving the invariants are distinct. Many such invariants are discussed later. Recent calculations by M. B. Thistlethwaite have produced the data in Table 1.2 for the number of prime knots (with the above conventions that neglect orientation) for crossing number up to 15. The table has been checked by J. Hoste and J. Weeks using totally independent methods from those of Thistlethwaite.

TABLE 1.2.

Crossing number	3	4	5	6	7	8	9	10	11	12	13	14	15
Number of knots	1	1	2	3	7	21	49	165	552	2176	9988	46972	253293

The naming of knots by means of traditional ordering is overwhelmed by the quantity of twelve-crossing knots. C. H. Dowker and Thistlethwaite [26] have adapted Tait's knot notation to produce a coding for knots that is suitable for a computer. The method is as follows: Follow along a knot diagram from some base point, allocating in order the integers $1, 2, 3, \ldots$ to the crossings as they are reached. Each crossing receives two numbers, one from the over-pass strand, one from the under-pass. At each crossing one of the numbers will be even and the other odd. Thus an n-crossing diagram with a base point produces a pairing between the first n odd numbers and the first n even numbers. An even number is then decorated with a minus sign if the corresponding strand is an under-pass; if it is an over-pass, it is undecorated. If the knot is prime, its diagram can easily be reconstructed uniquely (neglecting orientations) from that pairing with signs. Thus, specifying the signed even numbers in the order in which they correspond to the odd numbers $1, 3, 5, \ldots, 2n - 1$ specifies the knot up to reflection. Of course, there is no unique such specification, but for a given n, there can be only finitely many such ways of describing a knot. Selecting the lowest possible n and the first description in a lexicographical ordering of the strings of even numbers does give a canonical name for the (unoriented, prime) knot from which the knot can be constructed. For example, the first four knots in the tables are given by the notations

$$4\,6\,2, \quad 4\,6\,8\,2, \quad 4\,8\,10\,2\,6, \quad 6\,8\,10\,2\,4.$$

The crossing number is an easily defined example of the idea of a knot invariant. Knots with different crossing numbers cannot be equivalent. However, because it is defined in terms of a minimum taken over the infinity of possible diagrams of a knot, the crossing number is in general very difficult to calculate and use. The *unknotting number* $u(K)$ of a knot K is likewise a popular but intractable invariant; it will be mentioned in Chapter 7. By definition, $u(K)$ is the minimum number of crossing changes (from "over" to "under" or *vice versa*) needed to change K to the unknot, where the minimum is taken over all possible sets of crossing changes in all possible diagrams of K. However, if intuitively K is thought of as a curve moving around in S^3, then $u(K)$ is the minimum number of times that K must pass through itself to achieve the unknot. This obvious measure of a knot's complexity is often hard to determine and use. In fact, knowledge of the unknotting number of a knot might better be thought of as an end product of knot theory. If it has been shown that K is not the unknot, but that one crossing change on some diagram of K does give the unknot, then of course $u(K) = 1$. Thus, for example, it will soon be clear that $u(3_1) = u(4_1) = 1$. However, at the time of writing, $u(8_{10})$ is unknown (it is either 1 or 2). A discussion of the problem of finding unknotting numbers and of many, many other problems in knot theory can be found in [67].

A glance at Table 1.1 shows that all the knots up to 8_{18} have the property that in the displayed diagrams, the "over" or "under" nature of the crossings alternates as one travels along the knot. A knot is called *alternating* if it has such a diagram; alternating knots do seem to have particularly pleasant properties. It will later be seen that knots 8_{19}, 8_{20} and 8_{21} are not alternating. The apparent preponderance of alternating knots is simply a phenomenon of low crossing numbers. Looking at the given table, it is easy to imagine how various of its knots can be generalised to form infinite sets of knots by inserting extra crossings in a variety of ways. Further, note that for either orientation, $r(4_1) = 4_1 = \overline{4_1}$ and $r(3_1) = 3_1$; later it will be seen that $3_1 \neq \overline{3_1}$. Also $8_{17} = r\overline{8_{17}}$, but it is known that $8_{17} \neq r(8_{17})$. A proof of this last result is not easy; it follows from F. Bonahon's "equivariant characteristic variety theorem" [14], and it was also proved by A. Kawauchi [63]; another proof is in [40]. The first examples of knots that differ from their reverses were those of H. F. Trotter [125], which will be discussed in Chapter 11.

It is usually much more relevant to consider various classes of knots and links that have been found to be interesting, rather than to seek some list of all possible knots. An example, which later will be featured often, is that of pretzel knots and links. The *pretzel link* $P(a_1, a_2, \ldots, a_n)$ is shown in Figure 1.7. Here the a_i are integers indicating the number of crossings in the various "tassels" of the diagram. If a_i is positive, the crossings are in the sense shown (the complete "tassel" has a right-hand twist); if a_i is negative, the crossings are in the opposite sense. As n varies and different values are chosen for the a_i, this gives an infinite collection of links. Indeed, counting link components shows that it gives infinitely many links, but various invariants will later be used to distinguish pretzel knots.

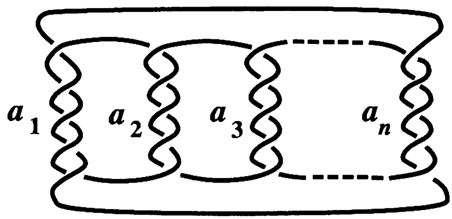


Figure 1.7

The upper two diagrams of Figure 1.8 show *rational* (or *2-bridge*) knots or links, denoted $C(a_1, a_2, \ldots, a_n)$. Such a link has no more than two components. The diagrams differ slightly in the way the various strands are joined at the right-hand edge of the diagram; the first method is for odd n, the second for even n. Again the a_i are integers, the sense of the crossings being as in the first diagram when all a_i are positive (so that then the upper "tassels" twist to the left and the lower ones to the right). For example, the second diagram shows $C(4, 2, 3, -3)$. This notation, devised by J. H. Conway [20], is chosen so that the link can be termed the "(p, q) rational link" where the rational number q/p has the repeated fraction expansion

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

It turns out that different ways of expressing q/p as such a repeated fraction always give the same link (though a link can correspond to distinct rationals). For

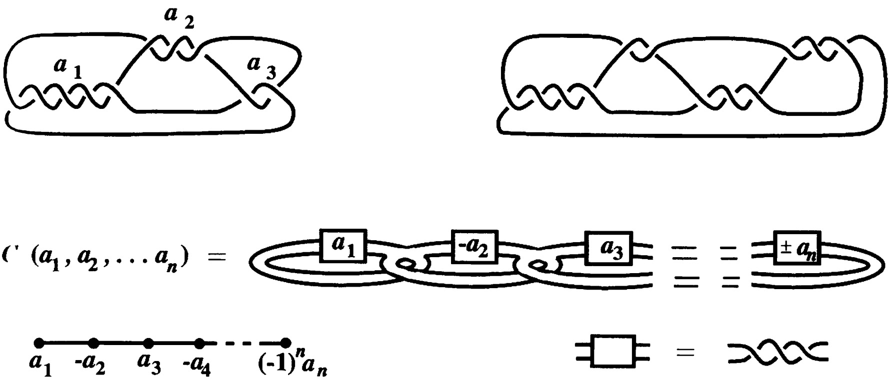


Figure 1.8

a (p, q) rational knot, $|p|$ is an invariant of the knot—namely, its determinant (see Chapter 9). An important property of a rational link is that it can be formed by gluing together two trivial 2-string tangles. Such a tangle is a 3-ball containing two standard (unknotted, unlinked) disjoint spanning arcs. Each arc meets the boundary of its ball at just its end points. The gluing process identifies together the boundaries of the balls to obtain S^3, and to produce the link, it identifies the four ends of the arcs in one ball with the ends of those in the other. This can be seen by considering a vertical line through one of the diagrams in Figure 1.8. The line meets the link in four points. The diagram to one side of the line represents two arcs in a ball and, forgetting the configuration on the other side of the line, the arcs untwist.

The remainder of Figure 1.8 shows how $C(a_1, a_2, \ldots, a_n)$ can be regarded as the boundary of n twisted bands "plumbed" together. If the a_i in the expression for q/p as a repeated fraction are all even, then the union of these bands is an orientable surface. The recipe for this plumbing can be encoded in a simple linear graph, as shown, in which each vertex represents a twisted band and each edge a plumbing. The boundary of a collection of bands plumbed according to the recipe of a tree (a connected graph with no closed loop) is called an *arborescent link.* (Conway called such a link "algebraic".) If the tree has only one vertex incident to more than two edges, the resulting link is a "*Montesinos link*"; the pretzel links are simple examples. Arborescent links have been classified by Bonahon and L. C. Siebenmann [15].

The ideas of braids and the braid group give a useful way of describing knots and links. A *braid of n strings* is n oriented arcs traversing a box steadily from the left to the right. The box will be depicted as a square or rectangle, and the arcs will join n *standard fixed* points on the left edge to n such points on the right edge. Over-passes are indicated in the usual way. The arcs are required to meet each vertical line that meets the rectangle in precisely n points (the arcs can never turn back in their progress from left to right). Two braids are the same if they are ambient isotopic (that is, the strings can be "moved" from one position to the other) *while keeping their end points fixed.* The standard generating element σ_i is shown in Figure 1.9 , as is the way of defining a product of braids by placing one after another. Given any braid b, its ends on the right edge may be joined to those on the left edge, in the standard way shown, to produce the *closed braid* $\widehat{b}$ that represents a link in S^3. Any braid can be written as a product of the σ_i and their inverses (σ_i^{-1} is σ_i with the crossing switched), and it is a result discovered by J. W. Alexander that any oriented link is the closure of some braid for some n. There are moves (the Markov moves; see Chapter 16) that explain when two braids have the same closure. More details can be found in [9] or [7]. The n-string braids form a group B_n with respect to the above product; it has a presentation

$$\langle\, \sigma_1, \sigma_2, \ldots, \sigma_{n-1}; \quad \sigma_i\sigma_j = \sigma_j\sigma_i \ \text{ if } |i-j| \geq 2\,, \quad \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}\,\rangle.$$

Figure 1.9 shows the braid $\sigma_1\sigma_2\ldots\sigma_{n-1}$. If $b = (\sigma_1\sigma_2\ldots\sigma_{n-1})^m$, then $\widehat{b}$ is called the (n, m) *torus link.* It is a knot if n and m are coprime. This link can be drawn on the standard (unknotted) torus in $\mathbb{R}^3$ (just consider the $n-1$ parallel strings of $\sigma_1\sigma_2\ldots\sigma_{n-1}$ as being on the bottom of the torus, and the other string as looping over the top of the torus).

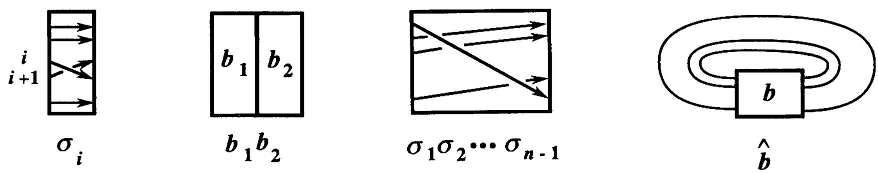


Figure 1.9

There are many methods of constructing complicated knots in easy stages. A common process is that of the construction of a *satellite* knot. Start with a knot K in a solid torus T. This is called a *pattern.* Let $e : T \to S^3$ be an embedding so that eT is a regular neighbourhood of a knot C in S^3. Then eK is called a *satellite* of C, and C is sometimes called a *companion* of eK. The process is illustrated in Figure 1.10, where a satellite of the trefoil knot 3_1 is constructed. Note that if $K \subset T$ and C are given, there are still different possibilities for the satellite, for T can be twisted as it embeds around C. A simple example of the construction is provided by the sum $K_1 + K_2$ of two knots; the sum is a satellite of K_1 and of K_2. If K is a (p, q) torus knot on the boundary of T, then eK is called the (p, q) *cable*

Figure 1.10

Figure 1.11

knot about C provided e maps a longitude of T to a longitude of C (see Definition 1.6).

A crossing in a diagram of an oriented link can be allocated a sign; the crossing is said to be positive or negative, or to have sign $+1$ or -1. The standard convention is shown in Figure 1.11. The convention uses orientations of both strands appearing at the crossing and also the orientation of space. A positive crossing shows one strand (either one) passing the other in the manner of a "right-hand screw". Note that, for a *knot*, the sign of a crossing does not depend on the knot orientation chosen, for reversing orientations of both strands at a crossing leaves the sign unchanged.

Definition 1.4. Suppose that L is a two-component oriented link with components L_1 and L_2. The linking number $\text{lk}(L_1, L_2)$ of L_1 and L_2 is half the sum of the signs, in a diagram for L, of the crossings at which one strand is from L_1 and the other is from L_2.

Note at once that this is well defined, for any two diagrams for L are related by a sequence of Reidemeister moves, and it is easy to see that the above definition is not changed by such a move (a move of Type I causes no trouble, as it features strands from only one component). The linking number is thus an invariant of oriented two-component links. To be equivalent, two such links must certainly have the same linking number. The definition given of linking number is symmetric:

$$\text{lk}(L_1, L_2) = \text{lk}(L_2, L_1).$$

This definition of linking number is convenient for many purposes, but it should not obscure the fact that linking numbers embody some elementary homology theory. Suppose that K is a knot in S^3. Then K has a regular neighbourhood N that is a solid torus. (This is easy to believe, but, technically, the regular neighbourhood is the simplicial neighbourhood of K in the second derived subdivision of a triangulation of S^3 in which K is a subcomplex.) The *exterior* X of K is the closure of $S^3 - N$. Thus X is a connected 3-manifold, with boundary ∂X that is a torus. This X has the same homotopy type as $S^3 - K$, $X \cap N = \partial X = \partial N$ and $X \cup N = S^3$. (Note the custom of using "∂" to denote the boundary of an object.)

Theorem 1.5. *Let K be an oriented knot in (oriented) S^3, and let X be its exterior. Then $H_1(X)$ is canonically isomorphic to the integers $\mathbb{Z}$, generated by the class of* *a simple closed curve μ in ∂N that bounds a disc in N meeting K at one point. If C is an oriented simple closed curve in X, then the homology class $[C] \in H_1(X)$ is* $\text{lk}(C, K)$. *Further,* $H_3(X) = H_2(X) = 0$.

PROOF. This result is true in any reasonable homology theory with integer coefficients; indeed, it follows at once from the relatively sophisticated theorem of Alexander duality. The following proof uses the Mayer–Vietoris theorem, which relates the homology of two spaces to that of their union and intersection. As it has been assumed that all links are piecewise linearly embedded, it is convenient to think of simplicial homology and to suppose that X and N are sub-complexes of some triangulation of S^3. Consider then the following Mayer–Vietoris exact sequence for X and the solid torus N that intersect in their common torus boundary:

$$\begin{array}{ccccccc} & & H_3(X) \oplus H_3(N) & \longrightarrow & H_3(S^3) & \longrightarrow & \cdots \\ \cdots \longrightarrow & H_2(X \cap N) \longrightarrow & H_2(X) \oplus H_2(N) & \longrightarrow & H_2(S^3) & \longrightarrow & \cdots \\ \cdots \longrightarrow & H_1(X \cap N) \longrightarrow & H_1(X) \oplus H_1(N) & \longrightarrow & H_1(S^3) & \longrightarrow & \cdots. \end{array}$$

Now, $H_3(X) \oplus H_3(N) = 0$. This is because any connected triangulated 3-manifold with non-empty boundary deformation retracts to some 2-dimensional subcomplex (just "remove" 3-simplexes one by one, starting at the boundary), and hence it has zero 3-dimensional homology. The homology of the torus, the solid torus and the 3-sphere are all known as part of any elementary homology theory, so in the above it is only $H_2(X)$ and $H_1(X)$ that are not known.

The groups $H_3(S^3)$ and $H_2(X \cap N)$ are both copies of $\mathbb{Z}$. Recall that the Mayer–Vietoris sequence comes from the corresponding short exact sequence of chain complexes. A generator of $H_3(S^3)$ is represented by the chain consisting of the sum of all the 3-simplexes of S^3 coherently oriented. This pulls back to the sum of the 3-simplexes in X plus those in N. That maps by the boundary (chain) map to the sum of the 2-simplexes in ∂X plus those in ∂N, and this in turn pulls back to the sum of the (coherently oriented) 2-simplexes in $X \cap N$; this represents a generator of $H_2(X \cap N)$. Thus inspection of the map in the sequence between $H_3(S^3)$ and $H_2(X \cap N)$ shows that a generator is sent to a generator, and hence the map is an isomorphism. As $H_2(S^3) = 0$, the exactness implies that $H_2(X) \oplus H_2(N) = 0$.

As $H_2(S^3) = 0$ and $H_1(S^3) = 0$, the map from $H_1(X \cap N) = \mathbb{Z} \oplus \mathbb{Z}$ to $H_1(X) \oplus H_1(N)$ is an isomorphism. As $H_1(N) = \mathbb{Z}$, this implies that $H_1(X) = \mathbb{Z}$. This isomorphism $H_1(X \cap N) \to H_1(X) \oplus H_1(N)$ is induced by the inclusion maps of $X \cap N$ into each of X and N. Suppose that μ is a non-separating simple closed curve in $X \cap N$ that bounds a disc in the solid torus N, oriented so that μ encircles K with a right-hand screw. Then μ represents an element that is indivisible (that is, it is not the multiple of another element by a non-unit integer) in $H_1(X \cap N)$; of course, μ represents zero in $H_1(N)$. Thus under the above isomorphism, $[\mu] \mapsto (1, 0) \in \mathbb{Z} \oplus \mathbb{Z} = H_1(X) \oplus H_1(N)$, for the image must still be indivisible, and this can be taken to define the choice of identification of $H_1(X)$ with $\mathbb{Z}$. Examination of the definition of linking numbers in terms of signs of crossings shows that C is homologous in X to $\text{lk}(C, K)[\mu]$. □

Note that, with the notation of the above proof, a unique element of $H_1(X \cap N)$ must map to $(0, 1)$, where the $1 \in H_1(N)$ is represented by the oriented curve K. As $(0, 1)$ is indivisible, this class is represented by a simple closed curve λ in $X \cap N$. This gives substance to the following definition:

Definition 1.6. Let K be an oriented knot in (oriented) S^3 with solid torus neighbourhood N. A meridian μ of K is a non-separating simple closed curve in ∂N that bounds a disc in N. A longitude λ of K is a simple closed curve in ∂N that is homologous to K in N and null-homologous in the exterior of K.

Note that λ and μ, the longitude and meridian, both have standard orientations coming from orientations of K and S^3, they are well defined up to homotopy in ∂N and their homology classes form a base for $H_1(\partial N)$. The above ideas can easily be extended to the following result for links of several components.

Theorem 1.7. *Let L be an oriented link of n components in (oriented) S^3 and let X be its exterior. Then $H_2(X) = \bigoplus_{n-1} \mathbb{Z}$. Further, $H_1(X)$ is canonically isomorphic to $\bigoplus_n \mathbb{Z}$ generated by the homology classes of the meridians $\{\mu_i\}$ of the individual components of L.*

PROOF. The proof of this is just an adaptation of that of the previous theorem. Here N is now a disjoint union of n solid tori. The map $H_3(S^3) \to H_2(X \cap N)$ is the map $\mathbb{Z} \to \bigoplus_n \mathbb{Z}$ that sends 1 to $(1, 1, \ldots, 1)$, implying that $H_2(X) = \bigoplus_{n-1} \mathbb{Z}$. Now $H_1(N \cap X) = \bigoplus_{2n} \mathbb{Z}$ and $H_1(N) = \bigoplus_n \mathbb{Z}$, and the map $H_1(N \cap X) \to H_1(N) \oplus H_1(X)$ is still an isomorphism, so $H_1(X) = \bigoplus_n \mathbb{Z}$. The argument about the generators is as before. □

If C is an oriented simple closed curve in the exterior of the oriented link L, the *linking number* of C and L is defined by $\operatorname{lk}(C, L) = \sum_i \operatorname{lk}(C, L_i)$ where the L_i are the components of L. By Theorem 1.7, $\operatorname{lk}(C, L)$ is the image of $[C] \in H_1(X) \equiv \bigoplus_n \mathbb{Z}$ under the projection onto $\mathbb{Z}$ that maps each generator to 1.

Exercises

1. Show that the knot 4_1 is equivalent to its reverse and to its reflection.
2. A diagram of an oriented knot is shown on a screen by means of an overhead projector. What knot appears on the screen if the transparency is turned over?
3. From the theory of the Reidemeister moves, prove that two diagrams in S^2 of the same oriented knot in S^3 are equivalent, by Reidemeister moves of only Types II and III, if and only if the the sum of the signs of the crossings is the same for the two diagrams.
4. Attempt a classificaton of links of two components up to six crossings, noting any pairs of links in your table that you have not yet proved to be distinct.

5. Show that any diagram of a knot K can be changed to a diagram of the unknot by changing some of the crossings from "over" to "under". How many changes are necessary?

6. Prove that the (p, q) torus knot, where p and q are coprime, is equivalent to the (q, p) torus knot. How does it relate to the $(p, -q)$ and $(-p, -q)$ torus knots?

7. Find descriptions of the knot 8_9 in the Dowker–Thistlethwaite notation, in the Conway notation as a 2-bridge knot $C(a_1, a_2, a_3, a_4)$ and also as a closed braid $\widehat{b}$.

8. Prove that any 2-bridge knot is an alternating knot.

9. A knot diagram is said to be *three-colourable* if each segment of the diagram (from one under-pass to the next) can be coloured red, blue or green so that all three colours are used and at each crossing either one colour or all three colours appear. Show that three-colourability is unchanged by Reidemeister moves. Deduce that the knot 3_1 is indeed distinct from the unknot and that 3_1 and 4_1 are distinct. Generalise this idea to *n-colourability* by labelling segments with integers so that at every crossing, the over-pass is labelled with the average, modulo n, of the labels of the two segments on either side.

10. Can n-colourability distinguish the Kinoshita–Terasaka knot (Figure 3.3) from the unknot?

11. Let X_1 and X_2 be the exteriors of two non-trivial knots K_1 and K_2. Determine how a homeomorphism $h : \partial X_1 \to \partial X_2$ can be chosen so that the 3-manifold $X_1 \cup_h X_2$ has the same homology groups as S^3.

12. Let M be a homology 3-sphere, that is, a 3-manifold with the same homology groups as S^3. Show that the linking number of a link of two disjoint oriented simple closed curves in M can be defined in a way that gives the standard linking number when $M = S^3$.

2

Seifert Surfaces and Knot Factorisation

It will now be shown that any link in S^3 can be regarded as the boundary of some surface embedded in S^3. Such surfaces can be used to study the link in different ways. Here they are used to show that knots can be factorised into a sum of prime knots. Later they will feature in the theory and calculation of the Alexander polynomial.

Definition 2.1. A Seifert surface for an oriented link L in S^3 is a connected compact oriented surface contained in S^3 that has L as its oriented boundary.

Examples of such surfaces are shown in Figure 2.1 and have been mentioned in Chapter 1 for two-bridge knots. Of course, any embedding into S^3 of a compact connected oriented surface with non-empty boundary provides an example of a link equipped with a Seifert surface. A surface is non-orientable if and only if it contains a Möbius band. *Some* surface can be constructed with a given link as its boundary in the following way: Colour black or white, in chessboard fashion, the regions of S^2 that form the complement of a diagram of the link. Consider all the regions of one colour joined by "half-twisted" strips at the crossings. This is a surface with the link as boundary, and it may well be orientable. However, it may quite well be non-orientable for either one or both of the two colours. The usual diagram of the knot 4_1 has both such surfaces non-orientable. Thus, although this method may provide an excellent Seifert surface, a general method, such as that of Seifert which follows, is needed.

Figure 2.1

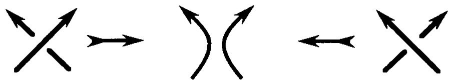


Figure 2.2

Theorem 2.2. *Any oriented link in S^3 has a Seifert surface.*

PROOF. Let D be an oriented diagram for the oriented link L and let $\hat{D}$ be D modified as shown in Figure 2.2. $\hat{D}$ is the same as D except in a small neighbourhood of each crossing where the crossing has been removed in the only way compatible with the orientation. This $\hat{D}$ is just a disjoint union of oriented simple closed curves in S^2. Thus $\hat{D}$ is the boundary of the union of some disjoint discs all on one side of (above) S^2. Join these discs together with half-twisted strips at the crossings. This forms an oriented surface with L as boundary; each disc gets an orientation from the orientation of $\hat{D}$, and the strips faithfully relay this orientation. If this surface is not connected, connect components together by removing small discs and inserting long, thin tubes. □

In the above proof, $\hat{D}$ was a collection of disjoint simple closed curves constructed from D. These curves are called the *Seifert circuits* of D. The Seifert circuits of the knot 8_{20} are shown in Figure 2.3. A Seifert surface for this knot is then constructed by adding three discs above the page and eight half-twisted strips near the crossings to join the discs together.

Figure 2.3

The proof of Theorem 2.2 gives a way of constructing a Seifert surface from a diagram of the link. The surface that results may however not be the easiest for any specific use. A surface coming from the chessboard colouring technique, or from some partial use of it, may well seem more agreeable. The diagram of Figure 2.4 shows how, at least intuitively, a knot can have two very different Seifert surfaces; the two thin circles can be joined by a tube after following along the narrow ("knotted") strip or after swallowing that part of the picture.

Definition 2.3. The genus $g(K)$ of a knot K is defined by

$$g(K) = \min.\{\text{genus}(F) : F \text{ is a Seifert surface for } K\}.$$

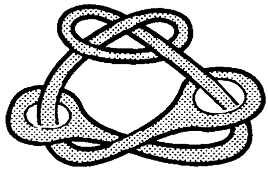


Figure 2.4

Here F has one boundary component, so as an abstract surface it is a disc with a number of "hollow handles" added. That number is its genus. More precisely, the genus of F is $\frac{1}{2}(1 - \chi(F))$, where $\chi(F)$ is the Euler characteristic of F. The Euler characteristic in turn can be defined as the number of vertices minus the number of edges plus the number of triangles in any triangulation of F. It does not seem to be common to discuss the genus of a link, but there is no difficulty in extending the definition.

Note that it follows at once that K is the unknot if and only if it has genus 0. Also, if K has a Seifert surface of genus 1 and K is known not to be the unknot, then $g(K) = 1$. The proof of Theorem 2.2 constructs a Seifert surface F for K from a diagram D of K. If D has n crossings and s Seifert circuits, then $\chi(F) = s - n$, so that $g(K) \leq \frac{1}{2}(n - s + 1)$.

It has already been noted that though it is easy to define numerical knot and link invariants by minimising some geometric phenomenon associated with it, often such invariants are very hard to calculate and difficult to use. The genus of a knot, however, has a utility that arises from the following result of [115], which states that knot genus is additive.

Theorem 2.4. *For any two knots K_1 and K_2,*

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

PROOF. Firstly, suppose that K_1 and K_2, together with minimal genus Seifert surfaces F_1 and F_2, are situated far apart in S^3. Each F_i is a connected surface with non-empty boundary, so elementary homology theory shows that $F_1 \cup F_2$ does not separate S^3. Thus one can choose an arc α from a point in K_1 to a point in K_2 that meets $F_1 \cup F_2$ at no other point and that intersects once a 2-sphere separating K_1 from K_2. The union of $F_1 \cup F_2$ with a "thin" strip around α (twisted to match orientations) gives a Seifert surface for $K_1 + K_2$ that has genus the sum of the genera of F_1 and F_2. Thus

$$g(K_1 + K_2) \leq g(K_1) + g(K_2)\,.$$

Now suppose that F is a minimal genus Seifert surface for $K_1 + K_2$. Let Σ be a 2-sphere, intersecting $K_1 + K_2$ transversely at two points, of the sort that occurs in the definition of $K_1 + K_2$. Thus Σ separates $K_1 + K_2$ into two arcs α_1 and α_2, and if β is any arc in Σ joining the two points of $\Sigma \cap (K_1 + K_2)$, then $\alpha_1 \cup \beta$

and $\alpha_2 \cup \beta$ are copies of K_1 and K_2. Now F and Σ are surfaces in S^3. Here it is being assumed throughout that all such inclusions are piecewise linear (as usual, "smooth" is just as good). Thus each can be regarded as a sub-complex of some triangulation of S^3, and Σ can be moved (by a general position argument, moving "one vertex at a time") to a position in which it is transverse to the whole of F. (The local situation is then modelled on the intersection of two planes, or half-planes, placed in general position in 3-dimensional Euclidean space.) Thus, without loss of generality, it may be assumed that $F \cap \Sigma$ is a 1-dimensional manifold which must be a finite collection of simple closed curves and one arc β joining the points of $\Sigma \cap (K_1 + K_2)$. Each of these simple closed curves separates Σ into two discs (using the 2-dimensional Schönflies theorem), only one of which contains β. Let C be a simple closed curve of $F \cap \Sigma$ that is *innermost* on $\Sigma - \beta$. This means that C bounds in Σ a disc D, the interior of which misses F. Now use D to do *surgery* on F in the following way: Create a new surface $\widehat{F}$ from F by deleting from F a small annular neighbourhood of C and replacing it by two discs, each a "parallel" copy of D, one on either side of D. If C did not separate F, this $\widehat{F}$ would be a Seifert surface for $K_1 + K_2$ of genus lower than that of F (since the surgery has the effect of removing a hollow handle). As that is not possible, C separates F, and so $\widehat{F}$ is disconnected. Consider the component of $\widehat{F}$ that contains $K_1 + K_2$. This is a surface of the same genus as F but which meets Σ in fewer simple closed curves (C, at least, has been eliminated). Repetition of this process yields a Seifert surface F' for $K_1 + K_2$, of the same genus as F, that intersects Σ only in β. Thus Σ separates F' into two pieces which are Seifert surfaces for K_1 and K_2. Hence

$$g(K_1) + g(K_2) \leq g(K_1 + K_2),$$

which, together with the preceding inequality, proves the result. □

Corollary 2.5. *No (non-trivial) knot has an additive inverse. That is, if $K_1 + K_2$ is the unknot, then each of K_1 and K_2 is unknotted.*

Corollary 2.6. *If K is a non-trivial knot and $\sum_1^n K$ denotes the sum of n copies of K, then if $n \neq m$ it follows that $\sum_1^n K \neq \sum_1^m K$. There are, then, certainly infinitely many distinct knots.*

Corollary 2.7. *A knot of genus 1 is prime.*

Corollary 2.8. *A knot can be expressed as a finite sum of prime knots.*

PROOF. If a knot is not prime, it can be expressed as the sum of two knots of smaller genus. Now use induction on the genus. □

It will be worthwhile recalling now the following basic Schönflies theorem, already mentioned in the introduction. Essentially, it states that S^2 cannot knot in S^3.

Theorem 2.9. Schönflies Theorem. *Let $e : S^2 \longrightarrow S^3$ be any piecewise linear embedding. Then $S^3 - eS^2$ has two components, the closure of each of which is a piecewise linear ball.*

No proof will be given here for this fundamental, non-trivial result (for a proof see [81]). The piecewise linear condition has to be inserted, as there exist the famous "wild horned spheres" that are are examples of topological embeddings $e : S^2 \longrightarrow S^3$ for which the complementary components are not even simply connected.

The next result considers the different ways in which a knot might be expressed as the sum of other knots. It is the basic result needed to show that the expression of a knot as a sum of prime knots is essentially unique. The technique of its proof again consists of minimising the intersection of surfaces in S^3 that meet transversely in simple closed curves, but the procedure here is more sophisticated than in the proof of Theorem 2.4. In the proof, use will be made of the idea of a *ball-arc pair.* Such a pair is just a 3-ball containing an arc which meets the ball's boundary at just its two end points. The pair is unknotted if it is pairwise homeomorphic to $(D \times I, \star \times I)$, where $\star$ is a point in the interior of the disc D and I is a closed interval.

Theorem 2.10. *Suppose that a knot K can be expressed as $K = P + Q$, where P is a prime knot, and that K can also be expressed as $K = K_1 + K_2$. Then either*
(a) $K_1 = P + K_1'$ *for some* K_1', *and* $Q = K_1' + K_2$, *or*
(b) $K_2 = P + K_2'$ *for some* K_2', *and* $Q = K_1 + K_2'$.

PROOF. Let Σ be a 2-sphere in S^3, meeting K transversely at two points, that demonstrates K as the sum $K_1 + K_2$. The factorisation $K = P + Q$ implies that there is a 3-ball B contained in S^3 such that $B \cap K$ is an arc α (with K intersecting ∂B transversely at the two points $\partial\alpha$) so that the ball-arc pair (B, α) becomes, on gluing a trivial ball-arc pair to its boundary, the pair (S^3, P). As in the proof of Theorem 2.4, it may be assumed, after small movements of Σ, that Σ intersects ∂B transversely in a union of simple closed curves disjoint from K. The immediate aim will be to reduce $\Sigma \cap \partial B$. Note that if this intersection is empty, then B is contained in one of the two components of $S^3 - \Sigma$, and the result follows at once.

As $\Sigma \cap K$ is two points, any oriented simple closed curve in $\Sigma - K$ has linking number zero or ± 1 with K. Amongst the components of $\Sigma \cap \partial B$ that have zero linking number with K select a component that is innermost on Σ (with $\Sigma \cap K$ considered "outside"). This component bounds a disc $D \subset \Sigma$, with $D \cap \partial B = \partial D$. Now ∂D bounds a disc $D' \subset \partial B$ with $D' \cap K = \emptyset$ (by linking numbers), though $D' \cap \Sigma$ may have many components (see Figure 2.5). By the Schönflies theorem, the sphere $D \cup D'$ bounds a ball. "Moving" D' across this ball to just the other side of D changes B to a new position, with $\Sigma \cap \partial B$ now having fewer components than before. As the new position of B differs from the old by the addition or subtraction of a ball disjoint from K, the new (B, α) pair corresponds to P exactly as before. After repetition of this procedure, it may be assumed that each component of $\Sigma \cap \partial B$ has linking number ± 1 with K. (Thus, on each of the spheres Σ and ∂B, the components of $\Sigma \cap \partial B$ look like lines of latitude encircling, as the two poles, the two intersection points with K.)

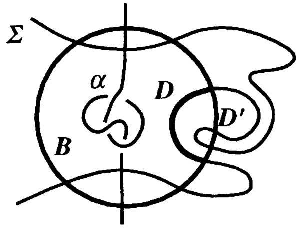


Figure 2.5

If now $\Sigma \cap B$ has a component that is a disc D, then $D \cap K$ is one point, and as P is prime, one side of D in B is a trivial ball-arc pair (see Figure 2.5). Removing from B (a regular neighbourhood of) this trivial pair produces a new B with the same properties as before but having fewer components of $\Sigma \cap B$. Thus it may be assumed that every component of $\Sigma \cap B$ is an annulus.

Let A be an annulus component of $\Sigma \cap B$. Then ∂A bounds an annulus A' in ∂B and A may be chosen (furthest from α) so that $A' \cap \Sigma = \partial A'$. Let M be the part of B bounded by the torus $A \cup A'$ and otherwise disjoint from $\Sigma \cup \partial B$. Let Δ be the closure of one of the components of $\partial B - A'$. Then Δ is a disc, with $\partial \Delta$ one of the components of A', and $\Delta \cap K$ equal to a single point (though $\Delta \cap \Sigma$ may have many components). This is illustrated schematically in Figure 2.6. Let $N(\Delta)$ be a small regular neighbourhood of Δ in the closure of $B - M$. This should be thought of as a thickening of Δ into $B - M$. The pair $(N(\Delta), N(\Delta) \cap \alpha)$ is a trivial ball-arc pair. However, $M \cup N(\Delta)$ is a ball, because its boundary is a sphere, and the fact that P is prime implies that the ball-arc pair $(M \cup N(\Delta), N(\Delta) \cap \alpha)$ is either trivial or a copy of the pair (B, α). If it is trivial (that is, when M is a solid torus), B may be changed, as before, by removing (a neighbourhood of) this pair to give a new B with fewer components of $\Sigma \cap B$. Otherwise, M is a copy of B less a neighbourhood of α, and that is just the exterior of the knot P; $\partial \Delta$ corresponds to a meridian of P. The closure of one of the complementary domains of Σ in S^3,

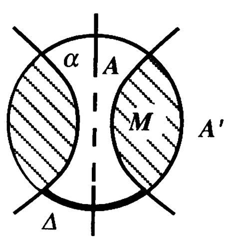


Figure 2.6

way that corresponding to K_1, contains M, and $M \cap \Sigma = A$. The meridian $\partial\Delta$ bounds a disc in $\Sigma - A$ that meets K at one point. This means that P is a summand of K_1 as required, so $K_1 = P + K_1'$ for some K_1'.

In this last circumstance, remove the interior of M and replace it with a solid torus $S^1 \times D^2$. Glue the boundary of the solid torus to ∂M, and ensure that the boundary of any meridional disc of $S^1 \times D^2$ is identified with a curve on ∂M that cuts $\partial\Delta$ at one point. Then $(S^1 \times D^2) \cup N(\Delta)$ is a ball, so B has been changed to become a new ball B', and (B', α) is a trivial ball-arc pair. The closure of $S^3 - B$ is unchanged; it is still a ball, so S^3 is changed to a new copy of S^3. In that new copy, the knot has become Q and, viewed as being decomposed by Σ, it has become $K_1' + K_2$. Thus $Q = K_1' + K_2$. □

Corollary 2.11. *Suppose that P is a prime knot and that $P + Q = K_1 + K_2$. Suppose also that $P = K_1$. Then $Q = K_2$.*

PROOF. By Theorem 2.10, there are two possibilities. The first is that for some K_1', $P + K_1' = K_1 = P$ and $Q = K_1' + K_2$. But then the genus of K_1' must be zero, so K_1' is the unknot and so $Q = K_2$. The second possibility is that for some K_2', $P + K_2' = K_2$ and $Q = K_2' + K_1$. But then $Q = K_2' + P = K_2$. □

Theorem 2.12. *Up to ordering of summands, there is a unique expression for a knot K as a finite sum of prime knots.*

PROOF. Suppose $K = P_1 + P_2 + \cdots + P_m = Q_1 + Q_2 + \cdots + Q_n$, where the P_i and Q_i are all prime. By the theorem, P_1 is a summand of Q_1 or of $Q_2 + Q_3 + \cdots + Q_n$, and if the latter, then it is a summand of one of the Q_j for $j \geq 2$, by induction on n. Of course if P_1 is a summand of Q_j, then $P_1 = Q_j$. By the corollary, P_1 and Q_j may then be cancelled from both sides of the equation, and the result follows by induction on m. Note that this induction starts when $m = 0$. Then $n = 0$ because the unknot cannot be expressed as a sum of non-trivial knots (again by consideration of genus). □

The theorems of this chapter are intended to make it reasonable to restrict attention to prime knots in most circumstances. Certainly that is the tradition when considering knot tabulation.

Exercises

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, would cut the torus that contains the knot.
2. For a 2-bridge knot K there is a 2-sphere separating S^3 into two balls, each of which intersects K in two standard arcs. By considering how this sphere might intersect a 2-sphere meeting the knot at two points, prove that a non-trivial 2-bridge knot is prime.

3. The bridge number of a knot K in S^3 is the least integer n for which there is an S^2 separating S^3 into two balls, each meeting K in n standard (unknotted and unlinked) spanning arcs. Show that the sum of two 2-bridge knots is a 3-bridge knot.

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

5. Prove that any knot may be changed to the unknot by a sequence of moves, each of which changes four arcs contained in a ball from one of the following configurations to the other.

[Think of the knot as the boundary of a non-orientable surface.]

6. Let F be the Seifert surface for a knot constructed by means of the Seifert method (Theorem 2.2). Let N be a regular neighbourhood of F. Show that the closure of $S^3 - N$ is a handlebody (that is, it is homeomorphic to a regular neighbourhood of a connected graph in S^3) homeomorphic to N.

7. Show, as outlined below, that a knot K with exterior X has a Seifert surface. Construct $f : X \to S^1$ as follows: First define $f|\partial X$ so that f maps a longitude to a single point and, when restricted to a meridian, f is a homeomorphism. Such an f can be extended over the 1-skeleton $T^{(1)}$ of some triangulation T of X so that if C is an oriented simple closed curve in $T^{(1)}$, then $\text{lk}(C, K) = [fC] \in H_1(S^1)$. Finally extend f over the 2-skeleton, then over the 3-skeleton (using the fact that any map $S^2 \to S^1$ extends over the 3-ball). Assuming f is simplicial with respect to some triangulations of X and S^1 (subdivisions of T and of a standard triangulation of S^1), consider $f^{-1}(x)$ where x is a point that is not a vertex in S^1.

8. Suppose that a knot A were to have an additive inverse B so that $A + B$ is the unknot. Let K be the simple closed curve in S^3 described as an infinite sum $A + B + A + B + \cdots$ where each summand is in a ball, the balls becoming successively smaller and converging to a single point. This K will not be piecewise linear. By considering the infinite sum as both $(A + B) + (A + B) + \cdots$ and $A + (B + A) + (B + A) + \cdots$, show that there is a homeomorphism (probably not piecewise linear) of S^3 to itself sending A to the unknot.

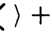
9. Suppose that addition of links is defined by just removing an unknotted ball-arc pair from each and identifying the resultant boundaries. Show that this is *not* a well-defined operation and that $L_1 + L_2 = L_1 + L_3$ does not necessarily imply that $L_2 = L_3$.

3

The Jones Polynomial

The theory of the polynomial invented by V. F. R. Jones gives a way of associating to every knot and link a Laurent polynomial with integer coefficients (that is, a finite polynomial expression that can include negative as well as positive powers of the indeterminate). The association of polynomial to link will be made by using a link diagram. The whole theory rests upon the fact that if the diagram is changed by a Reidemeister move, the polynomial stays the same. The polynomial for the link is then defined independently of the choice of diagram. Thus, if two links can be shown, by means of specific calculation from diagrams, to have distinct polynomials, then they are indeed distinct links. This is a relatively easy way of distinguishing knots with diagrams of few crossings. Table 3.1 displays the Jones polynomials for the knots of at most eight crossings shown in Chapter 1. Those polynomials are, by easy inspection, all distinct, so the corresponding knots are all distinct. As will be observed, the Jones polynomial is good, but not infallible, at distinguishing knots. However, that is not its most exciting achievement. Other invariants have, particularly with the aid of computers, always managed to distinguish any interesting pair of knots. Some of those invariants will be encountered in later chapters. The Jones polynomial, however, has been used to prove pleasing new results concerning the possible diagrams that certain knots can possess (see Chapter 5). In addition, the Jones polynomial has been much generalised; it has been developed into a theory, allied in some sense to quantum theory, giving invariants for 3-dimensional manifolds (see Chapter 13) and has been the genesis of a remarkable resurgence of interest in knot theory in all its forms. It is amazing that so simple, powerful and provocative a theory remained unknown until 1984, [53]. Because of the ease with which it can be developed, understood and used, the Jones polynomial has a place very near to the beginning of any exposition of knot theory. The simplest way to define it is by using a slightly different polynomial: the bracket polynomial discovered by L. H. Kauffman [59].

Definition 3.1. The Kauffman bracket is a function from unoriented link diagrams in the oriented plane (or, better, in S^2) to Laurent polynomials with integer coefficients in an indeterminate A. It maps a diagram D to $\langle D \rangle \in \mathbb{Z}[A^{-1}, A]$ and is characterised by

(i) $\langle \bigcirc \rangle = 1$,
(ii) $\langle D \sqcup \bigcirc \rangle = (-A^{-2} - A^{2})\langle D \rangle$,
(iii) $\langle$  $\rangle = A\langle$  $\rangle + A^{-1}\langle$  $\rangle$.

In this definition, $\bigcirc$ is the diagram of the unknot with no crossing, and $D \sqcup \bigcirc$ is a diagram consisting of the diagram D together with an extra closed curve $\bigcirc$ that contains no crossing at all, not with itself nor with D. In (iii) the formula refers to three link diagrams that are exactly the same except near a point where they differ in the way indicated. The bracket polynomial of a diagram with n crossings can be calculated by expressing it as a linear sum of 2^n diagrams with no crossing, using (iii), and noting that any diagram with c components and no crossing has, by (i) and (ii), $(-A^{-2} - A^{2})^{c-1}$ for its polynomial. In doing this, (iii) must be used on the crossings in some order, but it is easy to see (by transposing adjacent crossings in the order) that another choice of order does not effect the outcome. This means that the bracket polynomial is defined for link diagrams in the plane, and that it satisfies (i), (ii) and (iii). (If ever the empty diagram is required, it must be given the "polynomial" $(-A^{-2} - A^{2})^{-1}$.) If a diagram is changed in some way, then perhaps the polynomial changes, though the method of calculation makes it clear that changing a diagram by means of an orientation-preserving homeomorphism of the whole plane has no effect on the polynomial. The effect on $\langle D \rangle$ of a Reidemeister move on D will now be investigated.

Lemma 3.2. *If a diagram is changed by a Type I Reidemeister move, its bracket polynomial changes in the following way:*

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

PROOF.

$\langle$  $\rangle = A\langle$  $\rangle + A^{-1}\langle$  $\rangle$
$= (A(-A^{-2} - A^{2}) + A^{-1})\langle$  $\rangle$.

That produces the first equation; the second follows in the same way. □

Note that if in (iii) the crossing on the left-hand side were changed, then the right-hand side would be the same except for the interchange of A and A^{-1}. This follows from an application of (iii) rotated through $\pi/2$. This means that if $\overline{D}$ is the reflection of D—that is, D with the overs and unders of all of its crossings changed—then $\langle \overline{D} \rangle = \overline{\langle D \rangle}$, where the over-bar on the right denotes the effect of the involution on $\mathbb{Z}[A^{-1}, A]$ induced by exchanging A and A^{-1}. The two equations of Lemma 3.2 are related by this observation. This lemma is used several times in the following examples, which calculate the bracket polynomial of a diagram of a simple two-component link and then of a diagram of a trefoil knot.

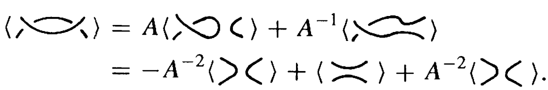
$\langle$  $\rangle = A\langle$  $\rangle + A^{-1}\langle$  $\rangle$
$(\ A^{-1} \quad A^{-1})$.

$$= A(-A^4 - A^{-4}) + A^{-7}$$
$$= (A^{-7} - A^{-3} - A^5).$$

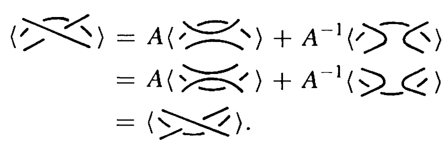
Lemma 3.3. *If a diagram D is changed by a Type II or Type III Reidemeister move, then* $\langle D \rangle$ *does not change. That is,*

Hence $\langle D \rangle$ *is invariant under regular isotopy of D.*

PROOF. (i)



(ii)



Here the second line follows from the first by using (i) twice. □

Definition 3.4. The writhe $w(D)$ of a diagram D of an *oriented* link is the sum of the signs of the crossings of D, where each crossing has sign $+1$ or -1 as defined (by convention) in Figure 1.11.

Note that this definition of $w(D)$ uses the orientation of the plane and that of the link. Note, too, that $w(D)$ does not change if D is changed under a Type II or Type III Reidemeister move. However, $w(D)$ does change by $+1$ or -1 if D is changed by a Type I Reidemeister move. It is thought that nineteenth-century knot tabulators believed that the writhe of a diagram was a knot invariant, at least when no reduction in the number of crossings by a Type I move was possible in a diagram. That lead to the famous error of the inclusion, in the early knot tables, of both a knot *and* its reflection, listed as 10_{161} and 10_{162} (an error detected by K. Perko in the 1970's). See Figure 3.1. The writhes of the diagrams are -8 and 10, respectively; yet, modulo reflection, these diagrams represent the same knot.

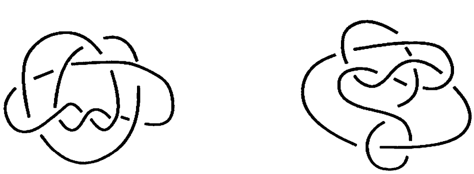


Figure 3.1

The writhe of an oriented link diagram and the bracket polynomial of the diagram with orientation neglected are, then, both invariant under Reidemeister moves of Types II and III, and both behave in a predictable way under Type I moves. This leads to the following result, which is essentially a statement of the existence of the Jones invariant.

Theorem 3.5. *Let D be a diagram of an oriented link L. Then the expression*

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

is an invariant of the oriented link L.

PROOF. It follows from Lemma 3.3 that the given expression is unchanged by Reidemeister moves of Types II and III; Lemma 3.2 and the above remarks on $w(D)$ show it is unchanged by a Type I move. As any two diagrams of two equivalent links are related by a sequence of such moves, the result follows at once. □

Definition 3.6. The Jones polynomial $V(L)$ of an oriented link L is the Laurent polynomial in $t^{1/2}$, with integer coefficients, defined by

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

where D is any oriented diagram for L.

Here $t^{1/2}$ is just an indeterminate the square of which is t. In fact, links with an odd number of components, including knots, have polynomials consisting of only integer powers of t. It is easy to show, by induction on the number of crossings in a diagram, that the given expression does indeed belong to $\mathbb{Z}[t^{-1/2}, t^{1/2}]$. Note that by Theorem 3.5, the Jones polynomial invariant is well defined and that $V(\text{unknot}) = 1$. At the time of writing, it is unknown whether there is a non-trivial knot K with $V(K) = 1$ and finding such a K, or proving none exists, is thought to be an important problem. The following table gives the Jones polynomial of knots with diagrams of at most eight crossings. It does not take very long to calculate such a table directly from the definition. It is clear that if the orientation of *every* component of a link is changed, then the sign of each crossing does not change. Thus the Jones polynomial of a *knot* does not depend upon the orientation chosen for the knot. It is easy to check that if the oriented link L^* is obtained from the oriented link L by reversing the orientation of one component K, then $V(L^*) = t^{-3\operatorname{lk}(K, L-K)}V(L)$. Thus the Jones polynomial depends on orientations in a very elementary way. Displayed in Table 3.1 are the coefficients of the Jones polynomials of the knots shown in Chapter 1. A bold entry in the table is a coefficient of t^0. For example,

$$V(6_1) = t^{-4} - t^{-3} + t^{-2} - 2t^{-1} + 2 - t + t^2.$$

The bracket polynomial of a diagram can be regarded as an invariant of *framed* unoriented links. For the moment, regard a framed link as a link L with an integer

TABLE 3.1. Jones Polynomial Table

3_1	-1	1	0	1	**0**						
4_1	1	-1	**1**	-1	1						
5_1	-1	1	-1	1	0	1	0	**0**			
5_2	-1	1	-1	2	-1	1	**0**				
6_1	1	-1	1	-2	**2**	-1	1				
6_2	1	-2	2	-2	2	**-1**	1				
6_3	-1	2	-2	**3**	-2	2	-1				
7_1	-1	1	-1	1	-1	1	0	1	0	0	**0**
7_2	-1	1	-1	2	-2	2	-1	1	**0**		
7_3	**0**	0	1	-1	2	-2	3	-2	1	-1	
7_4	**0**	1	-2	3	-2	3	-2	1	-1		
7_5	-1	2	-3	3	-3	3	-1	1	0	**0**	
7_6	-1	2	-3	4	-3	3	**-2**	1			
7_7	-1	3	-3	**4**	-4	3	-2	1			
8_1	1	-1	1	-2	2	-2	**2**	-1	1		
8_2	1	-2	2	-3	3	-2	2	-1	**1**		
8_3	1	-1	2	-3	**3**	-3	2	-1	1		
8_4	1	-2	3	-3	3	**-3**	2	-1	1		
8_5	**1**	-1	3	-3	3	-4	3	-2	1		
8_6	1	-2	3	-4	4	-4	3	**-1**	1		
8_7	-1	2	**-2**	4	-4	4	-3	2	-1		
8_8	-1	2	-3	**5**	-4	4	-3	2	-1		
8_9	1	-2	3	-4	**5**	-4	3	-2	1		
8_{10}	-1	2	**-3**	5	-4	5	-4	2	-1		
8_{11}	1	-2	3	-5	5	-4	4	**-2**	1		
8_{12}	1	-2	4	-5	**5**	-5	4	-2	1		
8_{13}	-1	2	-3	5	-5	**5**	-4	3	-1		
8_{14}	1	-3	4	-5	6	-5	4	**-2**	1		
8_{15}	1	-3	4	-6	6	-5	5	-2	1	0	**0**
8_{16}	-1	3	-5	6	-6	6	**-4**	3	-1		
8_{17}	1	-3	5	-6	**7**	-6	5	-3	1		
8_{18}	1	-4	6	-7	**9**	-7	6	-4	1		
8_{19}	**0**	0	0	1	0	1	0	0	-1		
8_{20}	-1	1	-1	2	-1	**2**	-1				
8_{21}	1	-2	2	-3	3	-2	2	**0**			

assigned to each component. Let D be a diagram for L with the property that for each component K of L, the part of D corresponding to K has as its writhe the integer assigned to K. Then $\langle D \rangle$ is an invariant of the framed link. Note that any diagram for L can be adjusted by moves of Type I (or its reflection) to achieve any given framing.

The Jones polynomial is characterised by the following proposition, which follows easily from the above definition (though historically it preceded that definition).

Proposition 3.7. *The Jones polynomial invariant is a function*

$$V : \{\text{Oriented links in } S^3\} \longrightarrow \mathbb{Z}[t^{-1/2}, t^{1/2}]$$

such that

(i) $V(\text{unknot}) = 1$,

(ii) *whenever three oriented links L_+, L_- and L_0 are the same, except in the neighbourhood of a point where they are as shown in Figure 3.2, then*

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

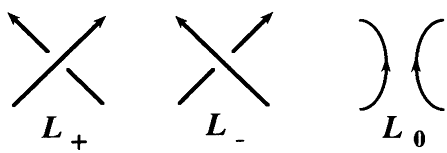


Figure 3.2

PROOF.

$$\langle \times \rangle = A\langle\,)\,(\,\rangle + A^{-1}\langle \asymp \rangle,$$
$$\langle \times \rangle = A^{-1}\langle\,)\,(\,\rangle + A\langle \asymp \rangle.$$

Multiplying the first equation by A, the second by A^{-1}, and subtracting gives

$$A\langle \times \rangle - A^{-1}\langle \times \rangle = (A^2 - A^{-2})\langle\,)\,(\,\rangle.$$

Thus, for the oriented links with diagrams as shown, using the fact that in those diagrams $w(L_+) - 1 = w(L_0) = w(L_-) + 1$, it follows that

$$-A^4V(L_+) + A^{-4}V(L_-) = (A^2 - A^{-2})V(L_0).$$

The substitution $t^{1/2} = A^{-2}$ gives the required answer. □

Working from Proposition 3.7, a straightforward exercise shows that if L' is L together with an additional trivial (unknotted, unlinking) component, then its Jones polynomial is given by $V(L') = (-t^{-1/2} - t^{1/2})V(L)$. Proposition 3.7 characterises the invariant in that using it allows the Jones polynomial of any oriented link to be calculated. This follows from the fact that any link can be changed to an unlink of c unknots (for which the Jones polynomial is $(-t^{-1/2} - t^{1/2})^{c-1}$) by changing crossings in some diagram; formula (ii) of Proposition 3.7 relates the polynomials before and after such a change with the that of a link diagram with fewer crossings (which has a known polynomial by induction).

The Jones polynomial of the sum of two knots is just the product of their Jones polynomials, that is,

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

This follows at once by considering a calculation of the polynomial of $K_1 + K_2$ and operating firstly on the crossings of just one summand. The same formula is true for links, but the sum of two links is not well defined; the result depends on which two components are fused together in the summing operation. That fact can easily be used, in a straightforward exercise, to produce two distinct links with the same Jones polynomial.

If an oriented link has a diagram D, its reflection has $\overline{D}$ as a diagram; of course, $w(D) = -w(\overline{D})$. As $\langle\overline{D}\rangle = \overline{\langle D\rangle}$, this means that if $\overline{L}$ is the reflection of the oriented link L, then $V(\overline{L})$ is obtained from L by interchanging $t^{-1/2}$ and $t^{1/2}$. The bracket polynomial of a diagram, of writhe equal to 3, for the right-handed trefoil knot 3_1 has already been calculated, and that at once determines that $-t^4 + t^3 + t$ is the Jones polynomial of the right-hand trefoil knot. Thus its reflection, the left-hand trefoil knot, has Jones polynomial $-t^{-4} + t^{-3} + t^{-1}$, and as this is a different polynomial, the two trefoil knots are distinct knots (that is, the trefoil knot is not *amphicheiral*). The figure-eight knot 4_1 is seen, by simple experiment, to be the same knot as its reflection; a glance at Table 3.1 verifies that its Jones polynomial is indeed symmetric between t and t^{-1}.

Figure 3.3 shows two distinct knots with the same Jones polynomial. The knot on the left is the Kinoshita–Terasaka knot, and that on the right is the Conway knot. That the knots are distinct can be shown by analysing their knot groups [110] or by determining their genera [32]. These two knots are related by the process called *mutation*. (Conway was the first to use this term.) That means that there is a ball in S^3 whose boundary meets one of the knots at four points. If this ball, with its intersection with the knot, is removed from S^3, rotated through angle π about an axis (in such a way as to preserve the four points), and then replaced, then the result is the other knot. In the diagrams, the boundary of the ball is indicated by

Figure 3.3

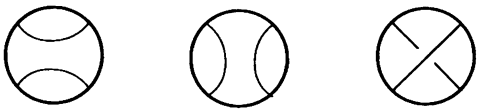


Figure 3.4

a dotted circle; the three possible axes of rotation are an axis perpendicular to the plane of the diagram, a north–south axis and an east–west axis (though the latter produces no change in the example depicted). In the case of oriented knots, it may be necessary to change all the orientations within the ball in addition to rotating it, so that the result should be consistently oriented. Now the Jones polynomial can be calculated using Proposition 3.7. Use this first on the crossings within the ball, changing and destroying crossings and removing unlinking unknots, until the Jones polynomial of the knot (or link) is a linear sum of Jones polynomials of links that, within the ball, are all of one of the three forms of Figure 3.4.

As each of these three configurations within the ball is unchanged by any of the three rotations, the same calculation ensues whether or not the ball is rotated. In fact, as oriented links are here being considered, only two of these three diagrams can occur; which two depends on the way the arrows are deployed.

Pretzel links offer another easy example of mutation. There is a mutation on the pretzel link $P(a_1, a_2, \ldots, a_n)$ of Figure 1.7 that interchanges a_i and a_{i+1}. Thus the Jones polynomial of $P(a_1, a_2, \ldots, a_n)$ is not changed when the $\{a_i\}$ are permuted in any way.

It should be noted that the length of a calculation of the Jones polynomial of a link made directly from the definition depends exponentially on the number of crossings in a diagram. Thus it is impractical when the number of crossings is not small. There is however a calculation for the (p, q) torus knot given in Theorem 14.13.

Exercises

1. Find the Jones polynomial of the $(2, q)$-torus knot.

2. Calculate the Jones polynomial of the 2-bridge knot given in Conway notation by $C(a, b)$, where a and b are positive integers.

3. Show that the Jones polynomial of an oriented link L takes the value $(-2)^{\#L-1}$ when $t = 1$, where $\#L$ is the number of components of L.

4. What is the value of the Jones polynomial of an oriented link L (i) when $t^{1/2} = e^{2\pi i/3}$ and (ii) when $t^{1/2} = e^{\pi i/3}$?

5. Calculate $V(5_1)$ using only the characterisation of the Jones polynomial given in Proposition 3.7

6. Prove that the knots 8_8 and 10_{129}, as shown in Figure 16.1, have the same Jones polynomial.

7. By considering the closure of braids of the form $\sigma_1^n \sigma_2 \sigma_1^{-2} \sigma_2$, find two links with distinct Jones polynomials but with homeomorphic exteriors.

8. Suppose that K_1 and K_2 are knots and that $K_1 \sqcup K_2$ is the "distant union" of K_1 and K_2, namely the two component link consisting of a copy of K_1 and a copy of K_2 separated by a 2-sphere. Show that $V(K_1 \sqcup K_2) = (-t^{-1/2} - t^{1/2})V(K_1)V(K_2)$.

9. Determine which knots with crossing number at most 8, other than 8_{17}, are amphicheiral (equivalent to their reflections). [In fact, $8_{17} \neq \overline{8_{17}}$.]

10. Verify the discovery of Perko that the knots illustrated in Figure 3.1 differ simply by reflection.

11. By considering the intersection between a disc spanning the unknot U and a 2-sphere meeting U at four points, show that U is the only knot related to U by mutation.

4

Geometry of Alternating Links

An alternating diagram for a link is, as explained in Chapter 1, one in which the over or under nature of the crossings alternates along every link-component in the diagram; the crossings always go "... over, under, over, under,... " when considered from any starting point. A link is said to be *alternating* if it possesses such a diagram. It has long been realised that alternating diagrams for a knot or link are particularly agreeable. However, the question posed by R. H. Fox—"What is an alternating knot?"—by which he was asking for some topological characterisation of alternating knots without mention of diagrams, is still unanswered. In later chapters the way in which the alternating property interacts with polynomial invariants will be discussed. In what follows here, some of the geometric properties of alternating links, discovered by W. Menasco [94], will be considered. The results are paraphrased by saying that an alternating link is split if and only if it is obviously split and prime if and only if it is obviously prime. Here "obviously" means that the property can at once be observed in the alternating diagram. This then establishes a ready supply of prime knots. Much of the ensuing discussion will concern 2-spheres embedded in S^3. It is to be assumed, as usual, that all such embeddings are piecewise linear (that is, simplicial with respect to some subdivisions of the basic triangulations).

Definition 4.1. A link $L \subset S^3$, having at least two components, is a split link if there is a 2-sphere in $S^3 - L$ separating S^3 into two balls, each of which contains a component of L. A link diagram D in S^2 is a split diagram if there is a simple closed curve in $S^2 - D$ separating S^2 into two discs each containing part of D.

Theorem 4.2. *Suppose a link L has an alternating diagram D. Then L is a split link if and only if D is a split diagram.*

The proof of this will be one of the two main aims of this chapter. The next definition generalises Definition 1.3 to links (rather than knots) and expresses primeness in a slightly different way. It also extends the idea of primeness to diagrams.

Definition 4.3. A link $L \subset S^3$, other than the unknot, is prime if every 2-sphere in S^3 that intersects L transversely at two points bounds, on one side of it, a ball that intersects L in precisely one unknotted arc. A diagram $D \subset S^2$, of a link other than the unknot, is a prime diagram if any simple closed curve in S^2 that meets D transversely at two points bounds, on one side of it, a disc that intersects D in a diagram U of the unknotted ball-arc pair. D is strongly prime if such a U is always the trivial zero-crossing diagram.

Note that the only prime split link is the trivial link of two components. In Chapter 5 it will be seen that it is straightforward to determine whether an alternating diagram represents the unknot, and so, given the alternating condition, references to the unknot in the above definition cause no problem. The second main result of the chapter is as follows:

Theorem 4.4. *Suppose L is a link that has an alternating diagram D. Then L is a prime link if and only if D is a prime diagram.*

This result shows at once that the alternating diagrams in the knot tables do indeed represent prime knots, for it is easy to check that those diagrams are prime. The proofs of these results depend upon a procedure for moving surfaces contained in the complement of a link, or transverse to it, to a standard position with reference to a diagram. This procedure, now to be described, is very general and does not use the alternating condition. Proofs of the stated theorems follow from the application of that condition to standard position surfaces. The description does require some notation and terminology as follows.

As usual, if $D \subset S^2 \subset S^3$ is a diagram for a link L, D is a collection of curves with self-intersections in the sphere S^2, together with over or under information at these intersections. The link L will be taken to be equal to D except near the crossings and, near any crossing, to be on a small sphere centred on the crossing. These small spheres are the boundaries of small balls called *bubbles*. The over-passing arcs are on the "upper" (or "Northern") halves of the small spheres, the under-passing arcs on the "lower" halves, S^2 being regarded as separating the small spheres into "upper" hemispheres on one side of S^2 and "lower" hemispheres on the other side. This is shown in Figure 4.1 on the left. Let S_+ and S_- be the two 2-spheres created from S^2 by removing the intersection of S^2 with all the bubbles and replacing those discs by the upper hemispheres or the lower hemispheres,

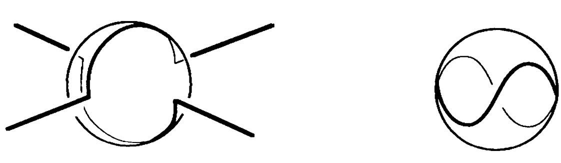


Figure 4.1

respectively, of the bubbles' boundaries. Let B_+ and B_- be the balls bounded by S_+ and S_-, so that B_+, B_- and the bubbles have disjoint interiors.

Let F be a surface in S^3 that is transverse to L. By a general position isotopy in S^3, F can be moved to a new position in which it is still transverse to L (it will meet L at points of $L \cap S^2$), and is transverse to S_+, S_- and to the north–south axes of all the bubbles. This means that F can be taken to meet each of S_+ and S_- in the union of disjoint simple closed curves and to meet each bubble in disjoint saddles. (Maybe F is not transverse to S^2.) Each saddle is just a disc spanning the bubble; its boundary intersects S^2 in four points that divide it into four arcs, two arcs in S_+ and two in S_- (see Figure 4.1). The surface F, in such general position, will be said to be in *standard position* with respect to the above data if, in addition, three conditions hold:

(A) Each of $F \cap B_+$ and $F \cap B_-$ is a disjoint union of discs.
(B) No component of $F \cap S_+$ or $F \cap S_-$ meets any bubble in more than one arc.
(C) Each component of both $F \cap S_+$ and $F \cap S_-$ meets some saddle or meets L.

Lemma 4.5. *Let D be a non-split diagram for L. Suppose that F is a 2-sphere with the property that it separates the components of L; then F can be replaced by another 2-sphere with the same property that is in standard position.*

PROOF. (a) Suppose that C is amongst the n components of $F \cap S_+$ that do not bound disc components of $F \cap B_+$. Choose C to be innermost on S_+ amongst such components. Then C is the boundary of a disc Δ in S_+, and any component of $F \cap S_+$ contained in the interior of Δ does bound a disc of $F \cap B_+$. Thus if Δ' denotes a copy of Δ displaced into B_+, Δ' can be chosen so that $\Delta' \cap F = \partial\Delta'$, $\partial\Delta'$ being a copy of C displaced along F into B_+. Now $\partial\Delta'$ separates the sphere F into two discs E_1 and E_2. Then $\Delta' \cup E_1$ or $\Delta' \cup E_2$ separates the components of L (because F did so). Let this new sphere be F'. Then $F' \cap S_+$ has fewer than n components not bounding discs in $F' \cap B_+$, for either C is no longer part of that intersection or, if C is still present, C now bounds a disc. Furthermore, $(F' \cap B_-) \subset (F \cap B_-)$. Thus, by repeating this, it may be assumed that F satisfies condition (A).

(b) Let H be the upper hemisphere of the boundary of a bubble. H is a disc in S_+ that meets L in one over-pass arc and meets F in disjoint arcs all parallel to the over-pass. Let δ be a diameter of H that intersects each of these arcs transversely at one point. The components of $F \cap S_+$ are disjoint simple closed curves on the sphere S_+. If δ meets one of these components at more than one point, then δ must meet some such component at two points of $\delta \cap F \cap S_+$ that are consecutive along δ. (This follows by considering the "innermost" component that δ meets.) Thus, if some component C of $F \cap S_+$ meets the bubble in more than one arc, C can be chosen so that C meets δ at adjacent points p_1 and p_2 of $\delta \cap F$; see Figure 4.2. If p_1 and p_2 on δ are on opposite sides of the over-pass, then they are both on opposite sides of *the same* saddle. The simple closed curve γ that consists of an arc from p_1 to p_2 across the saddle, followed by an arc from p_2 to p_1 in the disc in B_+ bounded by C, is homotopic in $S^3 - L$ to the meridian loop around the over-pass arc. On the

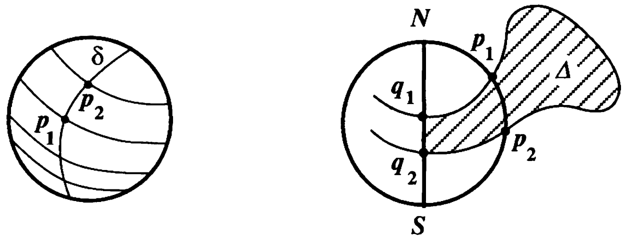


Figure 4.2

other hand γ, being contained in the sphere F, is null-homotopic in $S^3 - L$. This implies, erroneously by Theorem 1.7, that the meridian is null-homotopic. Thus p_1 and p_2 are on the same side of the over-pass and hence are points on adjacent saddles. Let q_1 and q_2 be the points where these two saddles intersect the north–south axis of the bubble. Consider the simple closed curve that consists of an arc in the first saddle from q_1 to p_1, an arc from p_1 to p_2 in the disc in B_+ bounded by C, an arc in the second saddle from p_2 to q_2 and then back to q_1 along the axis (see Figure 4.2). This curve bounds a disc Δ that can be chosen, using condition (A), to meet F only in the above composition of arcs from q_1 to q_2 and to be disjoint from L. Now move F by an isotopy that pushes F across Δ to a new position in which the intersection points q_1 and q_2 with the axis have been removed. Hence F can be changed to a new position with two fewer saddles. The previous process for ensuring that F satisfies condition (A) can then be repeated (it certainly does not increase the number of saddles). Repetition ensures that conditions (A) and (B) are satisfied.

(c) Finally, suppose that a component C of $F \cap S_+$ meets no saddle at all. Thus $C \subset S^2 - D$, and C bounds a disc in $F \cap B_+$ and a disc in $F \cap B_-$, the union of these discs being F. As C does not separate D, this union of two discs cannot separate L. Thus condition (C) is satisfied. □

(An alternative method for (c) is more useful in more general circumstances. As D is not a split diagram, C bounds a disc Δ' in $S^2 - D$ which is contained in $S_- \cap S_+$. Replace the disc of $F \cap B_+$ bounded by C with Δ', and then displace Δ' a little into B_-. Repetition of this process changes F, reducing the number of components of $F \cap S_+$ and $F \cap S_-$, until condition (C) is satisfied.)

Lemma 4.6. *Suppose that L, with diagram D, is not a split link. Suppose that F is a 2-sphere meeting L transversely at two points, with the property that F separates S^3 into two 3-balls, neither of which intersects L in a trivial ball-arc pair. Then F can be replaced by another 2-sphere, with the same property, that is in standard position.*

PROOF. The proof of this lemma follows closely that of the preceding one. In (a), the boundary of the disc Δ' cannot separate, on F, the two points of $L \cap F$, or else a meridian of L would be null-homotopic in $S^3 - L$. So, $\partial\Delta'$ bounds a disc E in $F - L$, and $\Delta' \cup E$ bounds (by the Schönflies theorem) a 3-ball that is disjoint from L (as L is not split). This ball can be used to change F by an isotopy that has the effect of replacing E with Δ'.

In (b), for the case when p_1 and p_2 are on the same side of the over-pass, the reasoning is the same as before. When they are on opposite sides, consider the simple closed curve γ constructed as before. This γ bounds a disc Γ that meets L at one point, with $\Gamma \cap F = \gamma$. Now, γ must separate on F the points of $F \cap L$, or a meridian is null-homotopic. F can now be replaced by the union of Γ and one of the components of $F - \gamma$. It is straightforward to check (using the fact that additive inverses to knots do not exist) that a correct choice of component preserves the property that (the new) F does not bound a trivial ball-arc pair. This replacement reduces the number of saddles required, and so repeating the process finitely many times achieves conditions (A) and (B).

The final part of the proof, to achieve condition (C), is exactly as before. □

Now, using all the preceding notation, suppose that F is a surface in standard position, and that the diagram D, used to specify the concept of standard position, is *alternating*. Consider a component C, temporarily oriented, of $F \cap S_+$. Suppose, when C enters a certain region of $(S^2 \cap S_+) - D$, it has a saddle to its left; then it can only leave that region with a saddle on its right, or at a point of $F \cap L$. This follows from the alternating property; see Figure 4.3. Thus, proceeding along C, saddles occur on the ... left, right, left, right ..., *except* that points of $F \cap L$ can substitute for some of these saddles.

PROOF OF THEOREM 4.2. Clearly, if D is a split diagram, then L is a split link. So, suppose L is split and D is non-split. By Lemma 4.5, there is a 2-sphere F in standard position that separates the components of L. Suppose C is an innermost component of $F \cap S_+$, so that C bounds a disc Δ in S_+ with $\Delta \cap F = C$. By condition (C), C meets a saddle and there changes from one region of $(S^2 \cap S_+) - D$ to another. (Consideration of the chessboard colouring of these regions shows at once that C must meet, in total, an even number of saddles in order to return to

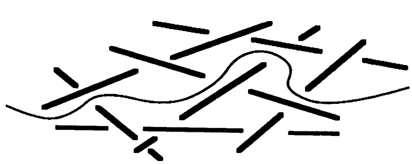


Figure 4.3

the original region.) Thus the alternating condition implies that C has at least one saddle to its left and one to its right. The arc of such a saddle, on the side of the saddle opposite to C, is part of some other component of $F \cap S_+$ (by condition (B)). As there is a saddle on either side of C, some component of $F \cap S_+$ is in the interior of Δ, and this contradicts the choice of C. Hence there is no component at all of $F \cap S_+$, and similarly $F \cap S_-$ is also empty. Thus $F \subset B_+$ or $F \subset B_-$, and in either case F does not separate L. □

PROOF OF THEOREM 4.4. Suppose that the link L, with alternating diagram D, is not prime. If L is a split link, then by Theorem 4.2 D is a split diagram, and it is easy to see that D is not prime. Thus it may be assumed that L is not split. There is a 2-sphere F in S^3 that intersects L transversely at two points, separates S^3 into two balls that do not meet L in just an unknotted arc, and is (by Lemma 4.6) in standard position. Let C be an innermost component of $F \cap S_+$ on S_+. As in the preceding proof, C must have an even number (at least two) of places where it either meets L or is incident on a saddle. If there are two or more consecutive saddles, the alternating property implies (as in the proof of Theorem 4.2) that C cannot be innermost. There are only two intersections with L available. Thus either (i) C contains both such intersections and two saddle-arcs separating them, or (ii) C contains one intersection and one saddle-arc, or (iii) C contains just the two intersection points with L and no saddle-arc. If $F \cap S_+$ has more than one component, it has at least *two* innermost components; each meets L, as has just been observed. Case (i) cannot occur because there must be components of $F \cap S_+$ other than C to account for the arcs on the other sides of the saddles, but no more points of $F \cap L$ are available for another innermost arc. The situation of case (ii) is shown on the left of Figure 4.4; the thicker arcs are parts of L, and the ellipse represents C. The corresponding part of the configuration in $F \cap S_-$ is shown on the right, where it is seen that a contradiction to condition (B) arises. Thus case (iii) is the only possibility, and $F \cap S_+ = F \cap S_- = F \cap S^2$, this being one simple closed curve intersecting D at two points only. Because F separates L into non-trivial summands, this means that D is not a prime diagram. □

Observe that *in toto*, the method of the above proofs is first to use the hypotheses about F to put F into standard position and then to use the observation, implied by the alternating nature of D, that left saddles and right saddles alternate along

Figure 4.4

a component of $F \cap S_+$ (though a point of $F \cap L$ may replace such a saddle) to complete the argument. This method has been extended [94] with only a little extra ingenuity to produce the following results. Detailed proofs will not be given here; to produce them by extending the proofs of Theorems 4.2 and 4.4 is little more than an exercise. First a general definition from the theory of 3-manifolds is required.

Definition 4.7. Suppose F is a surface, other than a 2-sphere, contained in a 3-manifold M. Then F is incompressible in M if any disc $\Delta \subset M$ that spans F in M (that is, $\Delta \cap F = \partial\Delta$) has the property that $\partial\Delta$ bounds a disc in F. A 2-sphere is incompressible in M if it does not bound a 3-ball contained in M.

This means that F has no "significant" spanning disc at all.

Proposition 4.8. *Suppose L is a non-split, prime, alternating link and F is a closed incompressible surface in $S^3 - L$. Then there exists a disc Δ spanning F in S^3 that meets L transversely at precisely one point.*

Corollary 4.9. *Suppose L is a non-split, prime, alternating link. Any incompressible torus T contained in $S^3 - L$ is parallel to the boundary of a solid torus neighbourhood of one of the components of L.*

A torus with that final property is called a *peripheral torus* of L. Note that in using this result, the non-split and prime conditions can easily be verified from the preceding theorems. The theory developed by W. P. Thurston [121], on the existence of hyperbolic structures on 3-manifolds, requires that no non-peripheral incompressible tori should be present. That there should be no "essential" annuli is also required. This theory, applied to the result of the corollary above, then shows that the complement of any non-split, prime, alternating link, other than a twist link (see Figure 4.5), has a complete hyperbolic structure of finite volume.

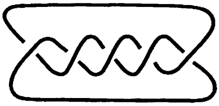


Figure 4.5

Definition 4.10. A Conway sphere for a link L in S^3 is a 2-sphere Σ in S^3 that meets L transversely at four points such that (i) $\Sigma - L$ is incompressible in $S^3 - L$ and (ii) any 2-sphere in $S^3 - \Sigma$ meeting L transversely at two points bounds a ball in $S^3 - \Sigma$ meeting L in just an unknotted arc.

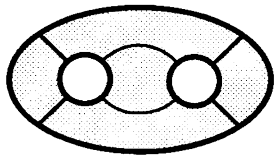


Figure 4.6

Note that the first condition implies that a disc spanning Σ in $S^3 - L$ cannot separate the part of L that lies on the same side of Σ as the disc. Discussion of Conway spheres is the essence of the *characteristic variety theory* for links due to Bonahon and Siebenmann ([14], [15]). They show that for any knot that is not a satellite, there is a well-defined maximal collection of Conway spheres that divides the knot into an arborescent part and a part in which any Conway sphere is pairwise parallel to a boundary component. The arborescent part consists of some copies of the 3-ball with two holes containing six arcs as in Figure 4.6, and some trivial 2-string tangles, glued together along some of their boundary (S^2, 4 point) pairs. The following result means that it is easy to spot Conway spheres from alternating link diagrams; it can be used to show that alternating knots near the beginning of the knot table certainly have no such spheres.

Proposition 4.11. *Suppose L is a non-split, prime link with alternating diagram D. If L has a Conway sphere, then it has a Conway sphere Σ such that $\Sigma \cap S_+$ is either* (i) *one curve containing all four points of $\Sigma \cap L$ and meeting no saddle, as on the left of Figure 4.7, or* (ii) *two curves, each containing two of the points of $\Sigma \cap L$ separated by two saddle-arcs, as on the right of Figure 4.7.*

Note that in either case $\Sigma \cap S_-$ is of the same form as $\Sigma \cap S_+$. In case (ii), the Conway sphere has two minima, two saddles and two maxima. Some recent extensions of Menasco's method can be found in [41] and [2].

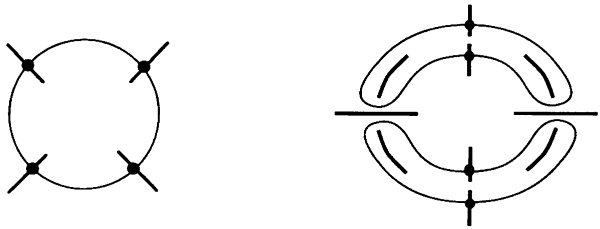
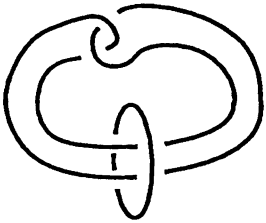


Figure 4.7

Exercises

1. Prove that a two-component link L that consists of a non-trivial knot K and a longitude of K is never a split link.
2. Prove, using the Jones polynomial, that the Whitehead link shown below is not a split link.



3. Find a prime diagram of a non-prime knot. Find a non-split diagram of a split link.
4. Show that a non-prime minimal crossing diagram of an alternating knot need not be an alternating diagram.
5. Let K_1 and K_2 be (possibly non-prime) knots. If $K_1 + K_2$ is alternating, show that K_1 and K_2 are both alternating.
6. Prove Proposition 4.8 and Corollary 4.9.

5

The Jones Polynomial of an Alternating Link

This chapter contains some of the most impressive applications of the Jones polynomial. They give solutions to two problems encountered by P. G. Tait in the nineteenth century. It is shown that an alternating knot diagram, when "reduced" in a rather elementary way, has the minimal number of crossings and that its writhe is an invariant of the knot. The crossing number of some other types of knot is also determined.

Let D be an n-crossing link diagram with its crossings labelled $1, 2, 3, \ldots, n$. A *state* for D is a function $s : \{1, 2, 3, \ldots, n\} \rightarrow \{-1, 1\}$. Of course, there are 2^n such states. Given D and a state s for D, let sD be a diagram constructed from D by replacing each crossing by two segments that do not cross. There are two possible ways of doing this. At the i^{th} crossing one way (the positive way) is used if $s(i) = 1$, and the other way (the negative way) it used if $s(i) = -1$. This is illustrated in Figure 5.1.

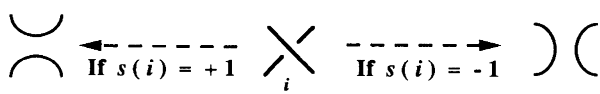


Figure 5.1

The diagram sD, having no crossing at all, is just a set of disjoint simple closed curves. Let there be $|sD|$ such curves. With this notation it is easy to write down a one-line formula for $\langle D \rangle$, the Kauffman bracket of D, as a summation over all possible 2^n states. The proof of this formula, which follows in Proposition 5.1, is simply that it immediately satisfies the criteria of Definition 3.1.

Proposition 5.1. *If D is a link diagram with n crossings, the Kauffman bracket of D is given by*

$$\langle D \rangle = \sum_s \left(A^{\sum_{i=1}^{n} s(i)} (-A^{-2} - A^2)^{|sD|-1} \right),$$

where the summation is over all functions $s : \{1, 2, 3, \ldots, n\} \rightarrow \{-1, 1\}$.

Now, let s_+ and s_- be the two constant states, so that for every i, $s_+(i) = 1$ and $s_-(i) = -1$. Of course, s_+ is the only state s for which $\sum_{i=1}^{n} s(i) = n$, and s_- is the only one for which $\sum_{i=1}^{n} s(i) = -n$.

Definition 5.2. The diagram D is plus-adequate if $|s_+ D| > |sD|$ for all s with $\sum_{i=1}^{n} s(i) = n - 2$ and is minus-adequate if $|s_- D| > |sD|$ for all s with $\sum_{i=1}^{n} s(i) = 2 - n$. If both conditions hold, D is called adequate.

Although this looks complicated, it is in fact easy to test whether a diagram be adequate: Change D to $s_+ D$ by replacing all the crossings in the positive manner described above, and inspect the diagram $s_+ D$. If the two segments of $s_+ D$ that replace a crossing of D never belong to the *same* component of $s_+ D$, then D is plus-adequate. So, just examine each component of $s_+ D$ to see if it ever abuts itself at a former crossing. The same procedure applied to $s_- D$ detects minus-adequacy. The prime example of this is the following result.

Proposition 5.3. *A reduced alternating link diagram is adequate.*

Here, "reduced" means that there is no crossing of the form featured in Figure 5.2 or its reflection (in which the squares labelled X and Y contain the whole diagram away from the crossing). Such a crossing is called a *nugatory* or *removable* crossing. It is a crossing at which one region of the complement of the diagram in the plane features twice, appearing near the crossing in a pair of diagonally opposite quadrants. (In practice such a crossing could be removed by rotating half of the link.)

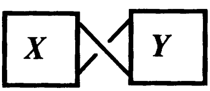


Figure 5.2

PROOF OF PROPOSITION 5.3. Let the complementary planar regions of the diagram be coloured black and white in a chessboard fashion. The alternating condition implies that the components of $s_+ D$ are the boundaries of the regions of one of the colours (the black ones, say) with the corners rounded off. Similarly, the components of $s_- D$ bound the white regions. The lack of removable crossings implies at once that D is adequate, for no region abuts itself. □

A specific non-alternating example is provided by the standard diagram of many pretzel knots. Figure 1.7 shows a diagram of the pretzel knot $P(a_1, a_2, \ldots, a_n)$. Recall that the crossings are all of the sense indicated when a_i is positive and in the other sense when a_i is negative. If $p_1, p_2, \ldots, p_r$ are all positive integers and $q_1, q_2, \ldots, q_s$ are all negative, then $P(p_1, p_2, \ldots, p_r, q_1, q_2, \ldots, q_s)$ is adequate provided that $p_i \geq 2$ and $q_i \leq -2$ for each i, and $r \geq 2$ and $s \geq 2$. Adequacy follows by simple inspection.

If P is any Laurent polynomial in some indeterminate, the maximum and minimum powers of the indeterminate that occur in P will be denoted $M(P)$ and $m(P)$. In what comes next the aim is to determine $M\langle D\rangle$ and $m\langle D\rangle$, the maximum and minimum powers of A that occur in the bracket (Laurent) polynomial of a diagram D.

Lemma 5.4. *Let D be a link diagram with n crossings. Then*
(i) $M\langle D\rangle \le n + 2|s_+D| - 2$, *with equality if D is plus-adequate, and*
(ii) $m\langle D\rangle \ge -n - 2|s_-D| + 2$, *with equality if D is minus-adequate.*

PROOF. (This is due, essentially, to Kauffman.) For any state s for D let

$$\langle D|s\rangle = A^{\sum_{i=1}^{n} s(i)}(-A^{-2} - A^2)^{|sD|-1},$$

so that $\langle D\rangle = \sum_s \langle D|s\rangle$. As $\sum_{i=1}^{n} s_+(i) = n$, it follows that $M\langle D|s_+\rangle = n + 2|s_+D| - 2$. Now any state s can be achieved by starting with s_+ and changing, one at a time, the value of s_+ on selected integers that label the crossings. In other words, there exist states $s_0, s_1, s_2, \ldots, s_k$ with $s_0 = s_+, s_k = s$ and $s_{r-1}(i) = s_r(i)$ for all $i \in \{1, 2, \ldots n\}$ except for a single integer i_r for which $s_{r-1}(i_r) = 1$ and $s_r(i_r) = -1$. Then $\sum_{i=1}^{n} s_r(i) = n - 2r$ and, because $s_{r-1}D$ and s_rD are the same diagram except near one crossing of D, $|s_rD| = |s_{r-1}D| \pm 1$. Hence $M\langle D|s_{r-1}\rangle - M\langle D|s_r\rangle$ is 0 or 4. Thus $M\langle D|s_r\rangle \le M\langle D|s_{r-1}\rangle$, and so, for all s, it follows that

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

If D is plus-adequate, it is immediate that $|s_1D| = |s_+D| - 1$, so that $M\langle D|s_r\rangle$ decreases at the first step, when r changes from 0 to 1, and never rises thereafter. Thus $M\langle D|s\rangle < n + 2|s_+D| - 2$ when $s \ne s_+$. Hence, in summing to achieve $\langle D\rangle$, the maximal degree term of $\langle D|s_+\rangle$ is never cancelled by a term from $\langle D|s\rangle$ for any s. The second statement of the lemma is really just the reflection of the first; its proof can be achieved by applying the above to $\overline{D}$. □

Corollary 5.5. *If D is an adequate diagram, then*

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

In order to interpret the last result, information is needed on $|s_+D|$ and $|s_-D|$. This is provided in the next two lemmas. Note that a diagram of a link is said to be a *connected* diagram if it is a connected subset of the plane (when drawn with no gaps for the under-passes); that is, it is not a split diagram in the sense of Definition 4.1.

Lemma 5.6. *Let D be a connected link diagram with n crossings. Then*

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

PROOF. Use induction on n. The result is clearly true when $n = 0$; suppose it to be true for diagrams with $n - 1$ crossings. Select a crossing of D. For at least one of the two ways of replacing the crossing with two segments that do not cross, the resulting diagram D' is connected. Suppose, with no loss of generality, that this is achieved by the positive way. Then $s_+D = s_+D'$ and $|s_-D| = |s_-D'| \pm 1$. Thus, using the induction hypothesis,

$$|s_+D| + |s_-D| = |s_+D'| + |s_-D'| \pm 1 \leq (n-1) + 2 \pm 1 \leq n + 2 . \qquad \square$$

Lemma 5.7. *Let D be a connected n-crossing diagram.*
(i) *If D is alternating, then* $|s_+D| + |s_-D| = n + 2$.
(ii) *If D is non-alternating and strongly prime (see Definition 4.3), then*

$$|s_+D| + |s_-D| < n + 2.$$

PROOF. When D is alternating, $|s_+D| + |s_-D|$ is the number of planar regions in the complement of D (as $|s_+D|$ is the number of black regions, $|s_-D|$ the number of white regions in a chessboard colouring). However, D is a four-valent planar graph, so consideration of the Euler number of the sphere shows that the number of regions is $n+2$ (for the number of edges is $2n$). Hence $|s_+D| + |s_-D| = n+2$.

Now suppose that D is non-alternating and strongly prime. Use induction on n. The induction starts easily when $n = 2$ with the two-crossing non-alternating diagram of two unlinked components. Thus, suppose $n \geq 3$. As D is non-alternating, it has two consecutive crossings that are both over-crossings or both under-crossings. Let c be a third crossing. As before, c can be removed in a positive or negative way. As D is strongly prime, the diagram resulting from either way will be connected. Consider the chessboard shading of the complementary regions of D and the graph Γ formed by taking a vertex for each black region and, for every crossing, an edge joining the vertices of the black regions that abut at that crossing. Strong primeness means that removal of any vertex does not separate Γ. The two ways of removing c correspond in Γ to removing, or shrinking to a point, the edge corresponding to c to produce a graph Γ'. If deleting the interior of an edge e of Γ produces a separating vertex v, then shrinking it does not produce a separating vertex (because v must be in any component of the complement of a neighbourhood of e in Γ). Thus one way of removing c gives a diagram D' that is strongly prime. Now D' is non-alternating because it has the same two consecutive similar crossings as had D. Thus the induction hypothesis can be applied to D' to give $|s_+D'| + |s_-D'| < n + 1$, and, as in the previous proof, this at once gives the required result. $\square$

The next result, the work of Kauffman, K. Murasugi and Thistlethwaite, is one of the main triumphs of the Jones polynomial. Its consequences have already been advertised here. As explained below in the corollary, it implies that a reduced alternating diagram of a knot is a diagram with the minimal number of crossings for that knot. This was inherently a conjecture of Tait's when he was compiling the first knot tables [118]. Firstly a simple definition is needed.

Definition 5.8. Suppose V is a Laurent polynomial in the indeterminate t. The breadth $B(V)$ of V is the difference between the maximal degree of t and the minimal degree of t that occur in V. (Thus $B(V) = M(V) - m(V)$.)

Theorem 5.9. *Let D be a connected, n-crossing diagram of an oriented link L with Jones polynomial $V(L)$. Then*
 (i) $B(V(L)) \leq n$;
 (ii) *if D is alternating and reduced, then* $B(V(L)) = n$;
 (iii) *if D is non-alternating and a prime diagram, then* $B(V(L)) < n$.

PROOF. Recall that under the substitution $t = A^{-4}$ the Jones polynomial is given by $V(L) = (-A)^{-3w(D)}\langle D\rangle$, so that $4B(V(L)) = B\langle D\rangle = M\langle D\rangle - m\langle D\rangle$ (where $M\langle D\rangle$ and $m\langle D\rangle$ refer to powers of A). Hence, by Lemmas 5.4 and 5.6,

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

But if D is alternating and reduced, then it is adequate, and the inequalities of Lemma 5.4 are then equalities. Then the first part of Lemma 5.7 implies that $4B(V(L)) = 4n$. When D is prime and non-alternating, any diagram summand that is a non-trivial diagram of the unknot makes no contribution to the Jones polynomial but does contribute to the number of crossings. Thus, without loss of generality, it may be assumed that D is strongly prime. Then the strict inequality of Lemma 5.7 produces the required result. □

Corollary 5.10. *If a link L has a connected, reduced, alternating diagram of n crossings, then it has no diagram of less than n crossings; any non-alternating prime diagram for L has more than n crossings.*

PROOF. The existence of the reduced alternating diagram for L implies, using Theorem 5.9 (ii), that $B(V(L)) = n$. If L has another diagram of m crossings, then Theorem 5.9 (i) implies that $n = B(V(L)) \leq m$. If this second diagram is non-alternating, then, by Theorem 5.9 (iii), $n = B(V(L)) < m$. □

Note that, from Table 3.1 the eight-crossing knots 8_{19}, 8_{20} and 8_{21} have Jones polynomials of breadth less than eight. Thus, by the above, if they were to have alternating diagrams, those diagrams would have less than eight crossings. However, knots with crossing number 7 or less have been classified earlier in the table, and no knot appears with the same polynomial as 8_{19}, 8_{20} or 8_{21}. Thus those three knots have no alternating diagrams at all. They are non-alternating knots.

The idea of taking parallels of diagrams provides another source of adequate diagrams, as will now be explained. The idea was used in [116], as detailed in the next theorem, to give a quick proof of a result of Thistlethwaite [119] establishing the invariance of the writhe of reduced alternating diagrams of a knot. Thus, if early compilers of knot tables believed writhe to be an invariant, they were correct within the domain of alternating diagrams.

Figure 5.3

Definition 5.11. If D is a link diagram, let its r-parallel D^r be the diagram in which each link-component of D is replaced by r copies, all parallel in the plane, each copy repeating the "over" and "under" behaviour of the original link-component.

Figure 5.3 shows a diagram and its 2-parallel.

Lemma 5.12. *If D is plus-adequate, then D^r is plus-adequate; if D is minus-adequate, then D^r is minus-adequate.*

PROOF. The result is immediate, because $s_+(D^r) = (s_+D)^r$; see Figure 5.4. If D is plus-adequate, no component of $s_+(D^r)$ abuts itself at a former crossing, as it runs parallel to a component of s_+D which, itself, has that property. □

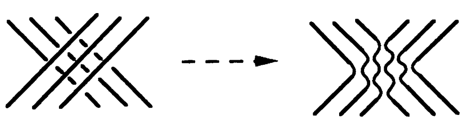


Figure 5.4

Theorem 5.13. *Let D and E be diagrams, with n_D and n_E crossings respectively, for the same oriented link L. Suppose that D is plus-adequate; then*

$$n_D - w(D) \leq n_E - w(E).$$

PROOF. Let $\{L_i\}$ be the components of L, and let D_i and E_i be the subdiagrams of D and E corresponding to L_i. Choose non-negative integers μ_i and ν_i such that for each i, $w(D_i) + \mu_i = w(E_i) + \nu_i$. Change D to D_* by changing each D_i to D_{*i} by adding to D_i a total of μ_i positive kinks. Similarly, change E to E_* by adding ν_i positive kinks to E_i for each i. Note that D_* is still plus-adequate, $w(D_{*i}) = w(E_{*i})$, and $w(D_*) = w(E_*)$, because the sum of the signs of crossings of distinct components is determined by the linking numbers of components of L. Now D^r_* and E^r_* are diagrams of the same link, namely L with each L_i replaced by r copies with mutual linking number $w(D_{*i})$. Thus they have the same Jones polynomial. But they have the same writhe (namely, $r^2 w(D_*)$), and so $\langle D^r_* \rangle = \langle E^r_* \rangle$. Now by

Lemma 5.4,

$$M\langle E_*^r\rangle \le (n_E + \sum_i \nu_i)r^2 + 2(|s_+E| + \sum_i \nu_i)r - 2,$$

$$M\langle D_*^r\rangle = (n_D + \sum_i \mu_i)r^2 + 2(|s_+D| + \sum_i \mu_i)r - 2,$$

the equality occurring since D_*^r is plus-adequate. This is true for all r, so, comparing coefficients of r^2,

$$n_D + \sum_i \mu_i \le n_E + \sum_i \nu_i,$$

so that $n_D - \sum_i w(D_i) \le n_E - \sum_i w(E_i)$. Hence, once again using the fact that the sum of the signs of crossings of distinct components is determined by linking numbers of L, $n_D - w(D) \le n_E - w(E)$. □

Corollary 5.14. *Let D and E be as above.*

(i) *The number of negative crossings of D is less than or equal to the number of negative crossings of E.*
(ii) *The number of positive crossings in a minus-adequate diagram is minimal.*
(iii) *An adequate diagram has the minimal number of crossings.*
(iv) *Two adequate diagrams of the same link (e.g. reduced alternating diagrams) have the same writhe.*

The corollary is just restating the theorem in different ways. An example of the use of the corollary is the two famous diagrams (the Perko pair), originally labelled 10_{161} and 10_{162}, shown in Figure 3.1. The diagrams 10_{161} and $\overline{10}_{162}$ represent the same knot. Observe that $w(10_{161}) = -8$ and $w(\overline{10}_{162}) = -10$. Inspection of the diagrams shows that $\overline{10}_{162}$ is minus-adequate, the minimal number possible of positive crossings being zero. However, 10_{161} is plus-adequate, and so any diagram must have at least nine negative crossings. As 10_{161} has no diagram of less than ten crossings (from the classification tables), it is impossible to display the minimal number of positive crossings and the minimal number of negative crossings on the same diagram, and the two minima are achieved by the two given diagrams.

The above theory gives, then, the recent solutions to two of the three "conjectures" formulated by Tait a century ago—namely, that reduced alternating diagrams minimise crossing number, and that two such diagrams of the same link have the same writhe. The third of these "conjectures"—that two reduced alternating diagrams of the same link are related by a sequence of "flyping" operations—has also recently been proved [95]. Such a "flyping" operation is shown in Figure 5.5.

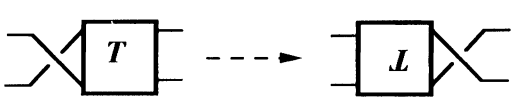


Figure 5.5

Exercises

1. Give an example of an n crossing diagram D for which $M\langle D\rangle - m\langle D\rangle = 0$.
2. Let K be a prime alternating knot. Show that any adequate diagram of K must be alternating.
3. Let $c(K)$ denote the crossing number of a knot K. If K_1 and K_2 are alternating knots prove that $c(K_1) + c(K_2) = c(K_1 + K_2)$. [Such an equality is not known to be true for arbitrary knots.]
4. A knot K has a reduced alternating diagram with n crossings where n is odd. Show that K is not equivalent to its reflection $\overline{K}$. Can $K + K$ be equivalent to its reflection?
5. Let D be a reduced n-crossing diagram of a knot K and suppose $B(V(K)) = n$. If D is not alternating, in what sense can it be said to be nearly alternating?
6. Show that a Whitehead double (a satellite using the curve shown in Figure 6.5) of a non-trivial alternating knot never has trivial Jones polynomial.
7. Show that the diagram of the Kinoshita–Terasaka knot shown in Figure 3.3 is adequate. What is the breadth of the Jones polynomial of this knot. Consider the same questions about the Conway knot. Each knot diagram in Figure 3.3 can be regarded as obtained by "summing together" a pair of "tangle" diagrams of two linked arcs in a disc, each tangle meeting the boundary of the disc at four points. What properties of the tangle diagrams will ensure adequacy of the knot diagram?
8. Find two prime knots that are distinct, even when orientations are neglected, that have (minimal) crossing number 15.

6

The Alexander Polynomial

The Alexander polynomial of an oriented link is, like the Jones polynomial, a Laurent polynomial associated with the link in an invariant way. The two polynomials give different information about the geometric properties of knots and links. The Alexander polynomial will, for example, give a lower bound for the genus of a knot, but it is not as useful as the Jones polynomial for investigating the required number of crossings in a diagram. The Alexander polynomial will later, in Theorem 8.6, be described combinatorially in terms of diagrams in a way that parallels Proposition 3.7, but the real interest of this invariant is that, in contrast to the Jones polynomial, it has a long history [3] and is well understood in terms of elementary homology theory. The homology approach to the Alexander polynomial, which will now be explained, describes it as a certain invariant of a homology module. To appreciate this, a little information about presentation matrices of modules is needed. There follows, then, a basic discussion of this topic, aimed at obtaining results rapidly. It may be neglected by the *cognoscenti.*

Suppose that M is a module over a commutative ring R. It will be assumed that R has a 1 and that $1x = x$ for all $x \in M$. A module can be regarded, by the insecure, as a vector space over a ring rather than over a field. A module is *free* if any element in it can be uniquely expressed as a linear sum of elements in a *base*; the module of n-tuples of elements of R is the canonical example of a free R-module. A *finite presentation* for M is an exact sequence

$$F \xrightarrow{\alpha} E \xrightarrow{\phi} M \longrightarrow 0$$

where E and F are free R-modules with finite bases. If α is represented by the matrix A with respect to bases $e_1, e_2, \ldots, e_m$ and $f_1, f_2, \ldots, f_n$ of E and F (the notation being so that $\alpha f_i = \sum_j A_{ji} e_j$), then the matrix A, of m rows and n columns, is a *presentation matrix* for M. As ϕ is a surjection, the images of $e_1, e_2, \ldots, e_m$ can be thought of as generators for M, and the images of $f_1, f_2, \ldots, f_n$ as relations amongst those generators.

Theorem 6.1. *Any two presentation matrices A and A_1 for M differ by a sequence of matrix moves of the following forms and their inverses:*

(i) *Permutation of rows or columns;*

(ii) *Replacement of the matrix A by* $\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}$*;*

(iii) *Addition of an extra column of zeros to the matrix A;*

(iv) *Addition of a scalar multiple of a row (or column) to another row (or column).*

PROOF. Suppose that the matrices A and A_1 correspond, with respect to some bases, to the maps α and α_1 in the following presentations:

$$\begin{array}{ccccccc} F & \xrightarrow{\alpha} & E & \xrightarrow{\phi} & M & \longrightarrow & 0 \\ \downarrow \gamma & & \downarrow \beta & & \updownarrow 1 & & \\ F_1 & \xrightarrow{\alpha_1} & E_1 & \xrightarrow{\phi_1} & M & \longrightarrow & 0 \end{array}$$

The free base of E and the surjectivity of ϕ_1 can be used to construct a linear map $\beta : E \to E_1$ so that $\phi_1\beta = \phi$. Similarly, the freeness of F and exactness at E and E_1 produce a map $\gamma : F \to F_1$ such that $\beta\alpha = \alpha_1\gamma$. If then β and γ are represented by matrices B and C with respect to the given bases, then $BA = A_1C$. A completely symmetrical argument produces maps β_1 and γ_1 with matrices B_1 and C_1 such that $B_1A_1 = AC_1$. Letting "$\sim$" denote "equivalence by the above moves", the following is apparent.

$$\begin{aligned} A &\sim \begin{pmatrix} A & B_1 \\ 0 & I \end{pmatrix} && \text{(by (ii) and (iv))} \\ &\sim \begin{pmatrix} A & B_1 & B_1A_1 \\ 0 & I & A_1 \end{pmatrix} && \text{(by (iii) and (iv))} \\ &\sim \begin{pmatrix} A & B_1 & 0 \\ 0 & I & A_1 \end{pmatrix} && \text{(by (iv), as} AC_1 = B_1A_1) \\ &\sim \begin{pmatrix} A & B_1 & 0 & B_1B \\ 0 & I & A_1 & B \end{pmatrix} && \text{(by (iii) and (iv)).} \end{aligned}$$

Now, for any $e \in E$, $\phi\beta_1\beta e = \phi e$, so, by the exactness at E, the image of $(\beta_1\beta - 1_E)$ is contained in the image of α. Because E is free, there is a map $\delta : E \to F$ so that $\alpha\delta = \beta_1\beta - 1_E$. Thus, if D is the matrix representing δ, $AD = B_1B - I$. Hence, use of (iv) shows that

$$\begin{pmatrix} A & B_1 & 0 & B_1B \\ 0 & I & A_1 & B \end{pmatrix} \sim \begin{pmatrix} A & B_1 & 0 & I \\ 0 & I & A_1 & B \end{pmatrix}.$$

Hence

$$A \sim \begin{pmatrix} A & B_1 & 0 & I \\ 0 & I & A_1 & B \end{pmatrix} \sim \begin{pmatrix} A_1 & B & 0 & I \\ 0 & I & A & B_1 \end{pmatrix} \sim A_1,$$

where the second equivalence is by (i) and the third is by a repeat of the whole argument with the rôles of the two presentations interchanged. □

Definition 6.2. Suppose that M is a module over a commutative ring R, having an $m \times n$ presentation matrix A. The r^{th} elementary ideal $\mathcal{E}_r$ of M is the ideal of R generated by all the $(m - r + 1) \times (m - r + 1)$ minors of A.

Of course, an $(m - r + 1) \times (m - r + 1)$ minor is the determinant of the matrix that remains after the removal from A of $(r - 1)$ rows and $(n - m + r - 1)$ columns. The standard elementary properties of determinants, together with the above theorem, show that the elementary ideals are independent of the presentation matrix chosen to evaluate them. Note that $\mathcal{E}_{r-1} \subseteq \mathcal{E}_r$. By convention, $\mathcal{E}_r = R$ when $r > m$ and $\mathcal{E}_r = 0$ if $r \leq 0$. Note that if $n = m$, the matrix A is square. Then there is only one $m \times m$ minor, and $\mathcal{E}_1$ is the principal ideal of R generated by $\det A$. A standard example is gained by observing that a finite abelian group G is a $\mathbb{Z}$-module, it does have a square presentation matrix, and $\mathcal{E}_1$ is the ideal of $\mathbb{Z}$ generated by $|G|$, the order of the group G.

Returning to geometric things, consider the first homology group, with integer coefficients, of an orientable, compact, connected surface F with n boundary components. Any elementary homology theory—simplicial homology or singular homology, for example (or just basic intuition)—asserts that $H_1(F; \mathbb{Z}) = \oplus_{2g+n-1}\mathbb{Z}$ generated by $\{[f_i]\}$, where the f_i are the oriented simple closed curves shown in Figure 6.1. There follows now a consideration of what happens when F is embedded in S^3, probably with the "bands" of Figure 6.1 twisted, linked and knotted. The next result can be regarded as an instance of Alexander duality.

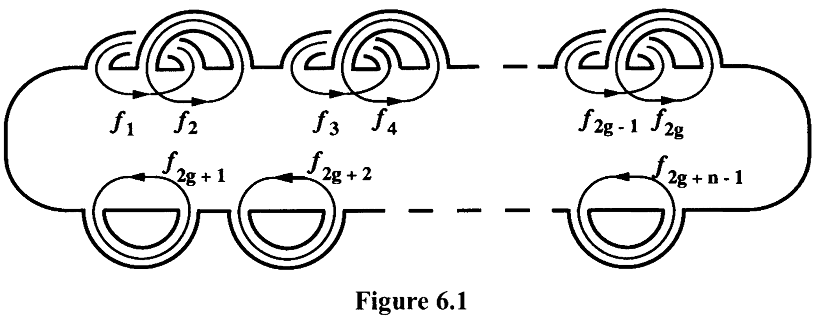


Figure 6.1

Proposition 6.3. *Suppose that F is a connected, compact, orientable surface with non-empty boundary, piecewise linearly contained in S^3. Then the homology groups $H_1(S^3 - F; \mathbb{Z})$ and $H_1(F; \mathbb{Z})$ are isomorphic, and there is a unique non-singular bilinear form*

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

with the property that $\beta([c], [d]) = \operatorname{lk}(c, d)$ for any oriented simple closed curves c and d in $S^3 - F$ and F respectively.

PROOF. The surface F is now embedded in S^3. As before, $H_1(F;\mathbb{Z}) = \bigoplus_{2g+n-1}\mathbb{Z}$ generated by $\{[f_i]\}$. Let V be a regular neighbourhood of F in S^3, so that V is just a 3-ball with $(2g+n-1)$ 1-handles attached. The inclusion of F in V is a homotopy equivalence, and $H_1(\partial V;\mathbb{Z}) = (\bigoplus_{2g+n-1}\mathbb{Z}) \oplus (\bigoplus_{2g+n-1}\mathbb{Z})$. For this, generators $\{[f_i'] : 1 \le i \le 2g+n-1\}$ and $\{[e_i] : 1 \le i \le 2g+n-1\}$ can be chosen so that each e_i is the boundary of a small disc in V that meets f_i at one point, and the inclusion $\partial V \subset V$ induces on homology a map sending $[f_i']$ to $[f_i]$ and $[e_i]$ to zero. Furthermore, the orientations of the $\{e_i\}$ can be chosen so that $\operatorname{lk}(e_i, f_j) = \delta_{ij}$ (the Krönecker delta). This all relates to the homology of the standard inclusion of F in a standard handlebody V; it is $S^3 - F$ that is of interest. Now, if V' is the closure of $S^3 - V$, then the inclusion of V' in $S^3 - F$ is a homotopy equivalence. The Mayer–Vietoris theorem for S^3 expressed as the union of V and V' asserts that the following sequence is exact:

$$H_2(S^3;\mathbb{Z}) \longrightarrow H_1(\partial V;\mathbb{Z}) \longrightarrow H_1(V;\mathbb{Z}) \oplus H_1(V';\mathbb{Z}) \longrightarrow H_1(S^3;\mathbb{Z})\,.$$

As the first and last groups in this sequence are zero, the map in the middle, induced by inclusion maps, is an isomorphism. Thus $H_1(V';\mathbb{Z})$ (which is isomorphic to $H_1(S^3 - F)$) is isomorphic to $\bigoplus_{2g+n-1}\mathbb{Z}$ and is generated by $\{[e_i] : 1 \le i \le 2g+n-1\}$. Now define

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

by $\beta([e_i],[f_j]) = \delta_{ij}$, and extend linearly. Suppose now that c and d are any oriented simple closed curves in $S^3 - F$ and F respectively, where $[c] = \sum_i \lambda_i[e_i]$ and $[d] = \sum_i \mu_i[f_i]$. Then $\beta([c],[d]) = \sum_i \lambda_i\mu_i$. However, $\operatorname{lk}(c, f_j) = [c] = \sum_i \lambda_i[e_i] \in H_1(S^3 - f_j;\mathbb{Z})$. Thus $\operatorname{lk}(c, f_j) = \lambda_j$. Similarly, $\operatorname{lk}(d, c) = \sum_i \mu_i[f_i] \in H_1(S^3 - c;\mathbb{Z})$, but this is $\sum_i \mu_i \operatorname{lk}(f_i, c)$, which by the above is $\sum_i \lambda_i\mu_i$. Hence, as required, $\beta([c],[d]) = \operatorname{lk}(c, d)$. □

Note that, whereas the above proof uses bases, β is characterised by linking numbers and is independent of bases. Note, too, that the bases used are mutually dual with respect to β in the sense that $\beta([e_i],[f_j]) = \delta_{ij}$, and so, using standard base changing arguments, corresponding to any base for $H_1(F;\mathbb{Z})$ there is a β-dual base for $H_1(S^3 - F;\mathbb{Z})$ and vice versa.

Now suppose that F is a Seifert surface for an oriented link L in S^3, so that $\partial F = L$. Let N be a regular neighbourhood of L, a disjoint union of solid tori that "fatten" the components of L. Let X be the closure of $S^3 - N$. Then $F \cap X$ is F with a (collar) neighbourhood of ∂F removed. Thus $F \cap X$ is just a copy of F and, just to simplify notation, it will be regarded as actually being F. This F has a regular neighbourhood $F \times [-1, 1]$ in X, with F identified with $F \times 0$ and the notation chosen so that the meridian of every component of L enters the neighbourhood at $F \times -1$ and leaves it at $F \times 1$. Let $i^\pm$ be the two embeddings $F \to S^3 - F$ defined by $i^\pm(x) = x \times \pm 1$ and, if c is an oriented simple closed curve in F, let $c^\pm = i^\pm c$.

Definition 6.4. Associated to the Seifert surface F for an oriented link L is the Seifert form

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

defined by $\alpha(x, y) = \beta((i^-)_\star x, y)$.

Note that, from Proposition 6.3 , α is defined and bilinear, and if a and b are simple closed oriented curves in F, then $\alpha([a],[b]) = \operatorname{lk}(a^-, b)$. Further, by sliding with respect to the second coordinate of $F \times [-1, 1]$, this is equal to $\operatorname{lk}(a, b^+)$.

Taking a basis $\{[f_i]\}$ for $H_1(F; \mathbb{Z})$ with a β-dual basis $\{[e_i]\}$ for $H_1(S^3 - F; \mathbb{Z})$ as before, α is represented by the *Seifert matrix* A, where

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

An immediate consequence is that in $H_1(S^3 - F; \mathbb{Z})$, $[f_i^-] = \sum_j A_{ij}[e_j]$ and $[f_j^+] = \sum_i A_{ij}[e_i]$.

Now let Y be the space X-cut-along-F. This means that Y is $X - F$ compactified, with *two* copies, F_- and F_+, of F replacing the removed copy of F (Y is homeomorphic to X less the open neighbourhood $F \times (-1, 1)$ of F). Of course, X can be recovered from Y by gluing F_+ and F_- together; thus $X = Y/\phi$, where ϕ is the natural homeomorphism $\phi : F_- \to F \to F_+$. Now take countably many copies of Y and glue them together to form a new space X_∞. More precisely, let $\{Y_i : i \in \mathbb{Z}\}$ be spaces homeomorphic to Y, and let $h_i : Y \to Y_i$ be a homeomorphism. Let X_∞ be the space formed from the disjoint union of all the Y_i by identifying $h_i F_-$ with $h_{i+1} F_+$ by means of the homeomorphism $h_{i+1}\phi h_i^{-1}$. The whole construction is illustrated in Figure 6.2, which shows X cut to form Y, then Y "uncurled", and then the copies Y_i of Y that are glued together to form X_∞.

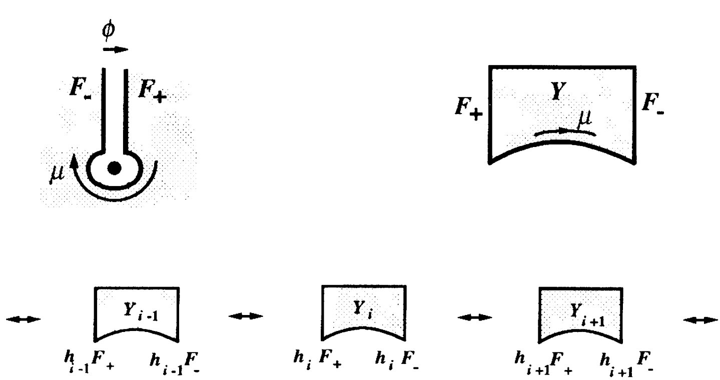


Figure 6.2

On X_∞ there is a natural self-homeomorphism $t : X_\infty \to X_\infty$ defined by $t|Y_i = h_{i+1}h_i^{-1}$. Clearly this is well defined; t is thought of as a translation of X_∞ by "one unit to the right". Hence the infinite cyclic group $\langle t \rangle$ generated by t acts on X_∞ as a group of homeomorphisms. Thus $\langle t \rangle$ also acts on $H_1(X_\infty; \mathbb{Z})$ (this action is really by means of the homology automorphism $t_\star$ induced by t, but the asterisk here is always suppressed). The ring $\mathbb{Z}$ acts on any abelian group, so the group-ring $\mathbb{Z}\langle t \rangle$ acts on $H_1(X_\infty; \mathbb{Z})$. Recall that for a group G written multiplicatively, the group-ring $\mathbb{Z}G$ consists of formal $\mathbb{Z}$-linear sums of elements of G. Addition in $\mathbb{Z}G$ comes from formal addition, and multiplication is induced by the multiplication in G and the distributive law. The ring $\mathbb{Z}\langle t \rangle$ is, then, just the ring $\mathbb{Z}[t^{-1}, t]$ of Laurent polynomials in t (that is, simply polynomials in t^{-1} and t with $\mathbb{Z}$ coefficients). The presence of this action means that $H_1(X_\infty; \mathbb{Z})$ is a module over the ring $\mathbb{Z}[t^{-1}, t]$. This terminology is used in the next fundamental theorem, which finds a presentation matrix for this module.

Theorem 6.5. *Let F be a Seifert surface for an oriented link L in S^3 and let A be a matrix, with respect to any basis of $H_1(F; \mathbb{Z})$, for the corresponding Seifert form. Then $tA - A^\tau$ is a matrix that presents the $\mathbb{Z}[t^{-1}, t]$-module $H_1(X_\infty; \mathbb{Z})$.*

PROOF. Express X_∞ as the union of subspaces Y' and Y'', where $Y' = \bigcup_i Y_{2i+1}$ and $Y'' = \bigcup_i Y_{2i}$. Each of these subspaces is the disjoint union of countably many copies of Y, and their intersection is the union of countably many copies of F. The homology of X_∞ will now be investigated, using the Mayer–Vietoris theorem, in terms of the homology of Y' and Y''. The Mayer–Vietoris long exact sequence of homology groups comes from a short exact sequence of chain complexes in a standard way. In this case the exact sequence of chain complexes is the following (where C_n is the n^{th} chain group):

$$0 \longrightarrow C_n(Y' \cap Y'') \xrightarrow{\alpha_n} C_n(Y') \oplus C_n(Y'') \xrightarrow{\beta_n} C_n(X_\infty) \longrightarrow 0.$$

Note that t interchanges Y' and Y'' so that the chain groups of these individual spaces are not modules over $\mathbb{Z}[t^{-1}, t]$; however, each term in the above sequence is such a module. To achieve an exact sequence of homology modules, α_n and β_n must be module maps with $\beta_n\alpha_n = 0$. This is achieved if β_n is defined by $\beta_n(a, b) = a + b$ and, for $x \in C_n(Y_{i-1} \cap Y_i)$, α_n is defined by $\alpha_n(x) = (-x, x) \in C_n(Y_{i-1}) \oplus C_n(Y_i)$. This short exact sequence of chain complexes of modules over $\mathbb{Z}[t^{-1}, t]$ gives rise, in the usual way, to the following long exact sequence of homology modules:

$$\to H_1(Y' \cap Y''; \mathbb{Z}) \xrightarrow{\alpha_\star} H_1(Y'; \mathbb{Z}) \oplus H_1(Y''; \mathbb{Z}) \xrightarrow{\beta_\star} H_1(X_\infty; \mathbb{Z}) \to$$

$$\to H_0(Y' \cap Y''; \mathbb{Z}) \xrightarrow{\alpha_\star} H_0(Y'; \mathbb{Z}) \oplus H_0(Y''; \mathbb{Z}).$$

Now F is, by definition of the term "Seifert surface", connected, so $H_0(F; \mathbb{Z}) = \mathbb{Z}$. But $Y' \cap Y''$ is countably many copies of F, each moved to the next by the homeomorphism t. Thus $H_0(Y' \cap Y''; \mathbb{Z})$ consists of one copy of $\mathbb{Z}$ for every power of t and so can be identified, as a module, with $\mathbb{Z}[t^{-1}, t] \otimes_{\mathbb{Z}} H_0(F; \mathbb{Z})$ (which is just a copy of $\mathbb{Z}[t^{-1}, t]$) with the generator of $H_0(Y_0 \cap Y_1; \mathbb{Z})$ corresponding to $1 \otimes 1$. However, $H_0(Y'; \mathbb{Z}) \oplus H_0(Y''; \mathbb{Z})$ is just the direct sum of countably many copies of $H_0(Y; \mathbb{Z})$, so this may be identified with $\mathbb{Z}[t^{-1}, t] \otimes_{\mathbb{Z}} H_0(Y; \mathbb{Z})$, with the generator of $H_0(Y_0; \mathbb{Z})$ corresponding to $1 \otimes 1$. Then $\alpha_\star(1 \otimes 1) = -(1 \otimes 1) + (t \otimes 1)$. This implies that on $H_0(Y' \cap Y''; \mathbb{Z})$, $\alpha_\star$ is injective, and hence $\beta_\star$ is a surjection.

Now apply to H_1 the same line of reasoning. $H_1(Y' \cap Y''; \mathbb{Z})$ can be identified with $\mathbb{Z}[t^{-1}, t] \otimes_{\mathbb{Z}} H_1(F; \mathbb{Z})$ so that $x \in H_1(Y_0 \cap Y_1; \mathbb{Z})$ corresponds to $1 \otimes x$. $H_1(Y'; \mathbb{Z}) \oplus H_1(Y''; \mathbb{Z})$ can be identified with $\mathbb{Z}[t^{-1}, t] \otimes_{\mathbb{Z}} H_1(Y; \mathbb{Z})$ so that $y \in H_1(Y_0)$ corresponds to $1 \otimes y$. Then, as a module, $H_1(Y' \cap Y''; \mathbb{Z})$ has a base $\{1 \otimes [f_i]\}$ and $H_1(Y'; \mathbb{Z}) \oplus H_1(Y''; \mathbb{Z})$ has a base $\{1 \otimes [e_i]\}$, where the e_i and f_i are the simple closed curves used in Proposition 6.3. Now the definition of $\alpha_\star$ shows that

$$\alpha_\star(1 \otimes [f_i]) = \sum_j (-A_{ij}(1 \otimes [e_j]) + A_{ji}(t \otimes [e_j]),$$

where A is the Seifert matrix with respect to the given bases. Hence, with respect to the module bases $\{1 \otimes [f_i]\}$ and $\{1 \otimes [e_i]\}$, $\alpha_\star$ is represented by matrix $tA - A^\tau$, and so, as $\beta_\star$ is surjective, this is a presentation matrix for the module $H_1(X_\infty; \mathbb{Z})$. □

It will be shown (fairly easily), in the following chapter on covering spaces that X_∞ and the action on it by $\langle t \rangle$ are well defined, given the oriented link L. Accept that fact for the time being. It implies at once that the $\mathbb{Z}[t^{-1}, t]$-module $H_1(X_\infty; \mathbb{Z})$ is an invariant of L. It is sometimes called the *Alexander module* of the oriented link. The actual module is cumbersome, but it has already been noted, as an immediate consequence of Theorem 6.1, that the elementary ideals of a module are invariants of that module.

Definition 6.6. The r^{th} Alexander ideal of an oriented link L is the r^{th} elementary ideal of the $\mathbb{Z}[t^{-1}, t]$ module $H_1(X_\infty; \mathbb{Z})$. The r^{th} Alexander polynomial of L is a generator of the smallest principal ideal of $\mathbb{Z}[t^{-1}, t]$ that contains the r^{th} Alexander ideal. The first Alexander polynomial is called the Alexander polynomial and is written $\Delta_L(t)$.

Note at once that a generator of a principal ideal is unique only up to multiplication by a unit (an invertible element) in the ring. Thus the Alexander polynomials, as defined above, are well defined only up to multiplication by $\pm t^{\pm n}$. Note, too, that the module $H_1(X_\infty; \mathbb{Z})$ does have a square presentation matrix, namely $tA - A^\tau$, where A is a Seifert matrix (by Theorem 6.5). Hence, the first elementary ideal is principal, and *the* Alexander polynomial of L is given by

$$\Delta_L(t) \doteq \det(tA - A^\tau),$$

where "$\doteq$" means "is equal to, up to multiplication by a unit".

EXAMPLE 6.7. The unknot has a disc D^2 for a Seifert surface. Cutting the exterior of the unknot along the disc gives $D^2 \times [-1, 1]$, and gluing countably many copies of this together produces $X_\infty = D^2 \times \mathbb{R}$. In this case, then, $H_1(X_\infty; \mathbb{Z}) = 0$, and this zero module is presented by the 1×1 unit matrix. Taking the determinant of this matrix (Theorem 6.5 is irrelevant here) shows that $\Delta_{\text{unknot}}(t) \doteq 1$.

EXAMPLE 6.8. Let K_n be the "twisted double" of the unknot, with orientation as shown in Figure 6.3. The lower part of the diagram has $2n - 1$ crossings in the sense shown when $2n - 1$ is positive; if $2n - 1$ is negative, the crossings there are in the opposite sense.

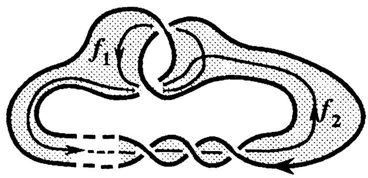


Figure 6.3

For the Seifert surface F take the surface shown, with generators for $H_1(F)$ represented by the oriented simple closed curves f_1 and f_2 as indicated. Recall that the Seifert matrix A is given by $A_{ij} = \text{lk}(f_i, f_j^+)$, where f_j^+ is a copy of f_j pushed off F into $S^3 - F$ in the direction defined by the oriented meridian of K_n. (The meridian encircles K_n in a "right-hand screw" direction.) Thus $A = \begin{pmatrix} 1 & 0 \\ -1 & n \end{pmatrix}$. Note that a diagonal entry $\text{lk}(f_i, f_i^+)$ is always the number of right-handed twists of an annular neighbourhood of f_i in F. It follows that

$$(tA - A^\tau) = \begin{pmatrix} t - 1 & 1 \\ -t & n(t - 1) \end{pmatrix};$$

so that $\Delta_{K_n} \doteq n(t^2 - 2t + 1) + t$. Note that K_0 is the unknot and that this formula gives $\Delta_{K_0} \doteq t$. That is in accord with the result of the previous example, as t is a unit in $\mathbb{Z}[t^{-1}, t]$. Of course, K_1 is the trefoil knot 3_1, and so that has Alexander polynomial $t^2 - t + 1$. Similarly, K_2 is the knot 5_2, and this has polynomial $2t^2 - 3t + 2$.

EXAMPLE 6.9. Let p, q and r be odd integers and let $P(p, q, r)$ be the pretzel knot shown in Figure 6.4. Once again the crossings are in the sense shown for positive integers and in the opposite sense for negative integers.

A Seifert surface is shown, together with generators f_1 and f_2. Then the Seifert matrix is given by

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

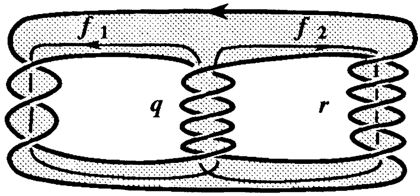


Figure 6.4

and so

$$\Delta_{P(p,q,r)}(t) \doteq \det(tA - A^\tau) = \frac{1}{4}\Big((pq+qr+rp)(t^2 - 2t + 1) + t^2 + 2t + 1\Big).$$

Note that if p, q and r are such that $(pq + qr + rp) = -1$ (for example, $(p, q, r) = (-3, 5, 7)$), then $\Delta_{P(p,q,r)}(t) \doteq t$, which is the Alexander polynomial for the unknot. The knot $P(-3, 5, 7)$ is known as Seifert's knot with unit Alexander polynomial; it can be shown to be a non-trivial knot by, for example, calculating its Jones polynomial.

As a special example, consider $P(3, 3, -3)$ (which is also listed as 9_{46}). The Seifert matrix A is $\begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix}$ and $(tA - A^\tau) = \begin{pmatrix} 3t-3 & 2t-1 \\ t-2 & 0 \end{pmatrix}$. The first elementary ideal of the Alexander module is then the ideal generated by the determinant $-2t^2 + 5t - 2$ (that is, the Alexander polynomial). The second elementary ideal is that generated by the 1×1 minors, so that is the ideal of $\mathbb{Z}[t^{-1}, t]$ generated by $(t - 2)$ and $(2t - 1)$. It is not the whole ring, as the evaluation at $t = -1$ gives a surjection $\mathbb{Z}[t^{-1}, t] \to \mathbb{Z}$ that maps the ideal in question to $3\mathbb{Z}$. This should be contrasted with the situation for the knot 6_1. This has a diagram the same as that of Figure 6.3 with $n = 3$ and the top two crossings of the diagram changed. For this, $A = \begin{pmatrix} -1 & 1 \\ 0 & 2 \end{pmatrix}$ and $(tA - A^\tau) = \begin{pmatrix} 1-t & t \\ -1 & 2t-2 \end{pmatrix}$. Here the Alexander polynomial is again $-2t^2 + 5t - 2$, but now the second elementary ideal is the whole of $\mathbb{Z}[t^{-1}, t]$. Thus these two knots are distinguished by the second, but not by the first, Alexander ideal.

Thus, the Alexander polynomial does not distinguish some pairs of knots. Nevertheless it is *quite* good at distinguishing knots; there follows soon a list of the Alexander polynomials of the prime knots up to eight crossings which this invariant certainly distinguishes from one another. First, though, there follow some easy properties of the Alexander polynomial.

Theorem 6.10.
(i) *For any oriented link* L, $\Delta_L(t) \doteq \Delta_L(t^{-1})$.
(ii) *For any (oriented) knot* K, $\Delta_K(1) = \pm 1$.

Analogues of these results hold for the r^{th} *Alexander polynomials.*

PROOF. (i) Suppose that A is an $n \times n$ Seifert matrix for L. Then

$$\Delta_L(t) \doteq \det(tA - A^\tau) = \det(tA^\tau - A) = (-t)^n \det(t^{-1}A - A^\tau) \doteq \Delta_L(t^{-1}).$$

(ii) Let A be the Seifert matrix for K coming from a standard base of $2g$ oriented curves $\{f_i\}$ on a genus g Seifert surface F as shown in Figure 6.1. Now, $\Delta_K(1) = \pm\det(A - A^\tau)$, but

$$(A - A^\tau)_{ij} = \operatorname{lk}(f_i^-, f_j) - \operatorname{lk}(f_i^+, f_j),$$

and this is the algebraic number of intersections of f_i and f_j on the surface F. Hence $(A - A^\tau)$ consists of g blocks of the form $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ down the diagonal and zeros elsewhere. The determinant of that is 1. □

Note that for a link L of more than one component, $\Delta_L(1) = 0$ by essentially the same proof (the blocks on the diagonal of $(A - A^\tau)$ are now followed by some zeros).

Corollary 6.11. *For any knot K,*

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

where the a_i are integers and a_0 is odd.

PROOF. By Theorem 6.10(i), $\Delta_K(t)$ can be written in the form $\Delta_K(t) = b_0 + b_1 t + b_2 t^2 + \cdots + b_N t^N$, where $b_{N-r} = \pm b_r$ with the same choice of sign for all r. If N were odd, $\Delta_K(1)$ would be even, which contradicts (ii) of the theorem. Hence N is even. If $b_{N-r} = -b_r$ for all r, then $b_{N/2} = 0$ and so $\Delta_K(1) = 0$, again a contradiction. Thus $b_{N-r} = b_r$ for all r and $b_{N/2}$ is odd, and so, within the indeterminacy of multiplication by units, $\Delta_K(t)$ is of the required form. □

In the following table, the coefficients $a_0, a_1, a_2, \ldots,$ that occur in the expression $\Delta_K(t) \doteq a_0 + a_1(t^{-1} + t) + a_2(t^{-2} + t^2) + \cdots$ are recorded. The signs are chosen so that $\Delta_K(1) = +1$, this being Conway's normalisation. For example,

$$\Delta_{8_7}(t) = -5 + 5(t^{-1} + t) - 3(t^{-2} + t^2) + (t^{-3} + t^3).$$

Proposition 6.12. *Let L be an oriented link. Then $\overline{L}$ and $\mathrm{r}L$, the reflection and the reverse of L, have the same Alexander polynomial as L up to multiplication by units.*
If K_1 and K_2 are oriented knots, $\Delta_{(K_1+K_2)}(t) \doteq \Delta_{K_1}(t)\Delta_{K_2}(t)$.

PROOF. If A is a Seifert matrix for L , $-A$ is a Seifert matrix for $\overline{L}$ and A^τ is a Seifert matrix for $\mathrm{r}L$.

If A_1 and A_2 are Seifert matrices for K_1 and K_2, then $\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$ is a Seifert matrix for $K_1 + K_2$. □

TABLE 6.1. Alexander Polynomial Table

Knot	a_0	a_1	a_2	a_3
3_1	−1	1		
4_1	3	−1		
5_1	1	−1	1	
5_2	−3	2		
6_1	5	−2		
6_2	−3	3	−1	
6_3	5	−3	1	
7_1	−1	1	−1	1
7_2	−5	3		
7_3	3	−3	2	
7_4	−7	4		
7_5	5	−4	2	
7_6	−7	5	−1	
7_7	9	−5	1	
8_1	7	−3		
8_2	3	−3	3	−1
8_3	9	−4		
8_4	−5	5	−2	
8_5	5	−4	3	−1
8_6	−7	6	−2	
8_7	−5	5	−3	1
8_8	9	−6	2	
8_9	7	−5	3	−1
8_{10}	−7	6	−3	1
8_{11}	−9	7	−2	
8_{12}	13	−7	1	
8_{13}	11	−7	2	
8_{14}	−11	8	−2	
8_{15}	11	−8	3	
8_{16}	−9	8	−4	1
8_{17}	11	−8	4	−1
8_{18}	13	−10	5	−1
8_{19}	1	0	−1	1
8_{20}	3	−2	1	
8_{21}	−5	4	−1	

Proposition 6.13. *If a knot K has genus g, then $2g \geq$ breadth $\Delta_K(t)$.*

PROOF. Let F be a genus g Seifert surface for K. Then $tA - A^\tau$ is a $2g \times 2g$ matrix, and so the degree in t of the polynomial $\det(tA - A^\tau)$ is at most $2g$. □

The last result can be considered as an application of the Alexander polynomial. Although it is only in the form of an inequality, it gives geometric information about individual knots. The surface constructed by removing the interiors of disjoint discs from a genus g surface is said to still have genus g. Proposition 6.13 generalises at once to show that if a link L with c components bounds a connected orientable surface of genus g, then

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

Now, it is a theorem discovered by R. H. Crowell [22] (see also [17]) that if L has an *alternating* diagram that gives, by means of Seifert's method of Theorem 2.2, a connected Seifert surface of genus $\hat{g}$, then breadth $\Delta_L(t) = 2\hat{g} + c - 1$. Thus the genus is always minimal for a Seifert surface coming in this way from any alternating diagram.

There are oriented links of two or more components that have their Alexander polynomials equal to *zero*. The next proposition describes some of them, but there are even more.

Proposition 6.14. *Suppose an oriented link L bounds a disconnected oriented surface in S^3; then $\Delta_L(t)$ is the zero polynomial.*

PROOF. Suppose Σ is a disconnected oriented surface with boundary L. Form a connected surface F by connecting the components of Σ together with thin "pipes". Take a set of oriented curves $\{f_i\}$ that give a base for $H_1(F)$, choosing f_1 to be a curve encircling once around one of the pipes and ensuring that f_1 is disjoint from the other f_i. This f_1 bounds a disc D in S^3 with $D \cap F = \partial D$. Then $\text{lk}(f_1, f_i^{\pm}) = 0$ for all i. Hence the corresponding Seifert matrix A has its first row and first column consisting entirely of zeros. Of course then $\det(tA - A^\tau) = 0$. □

The idea of a satellite knot was mentioned in Chapter 1. There is a simple formula that gives the Alexander polynomial of a satellite knot in terms of those of its companion and its pattern. This will now be explained.

Theorem 6.15. *In S^3, let T be a standard, unknotted, solid torus that contains a knot K. Let $e : T \to S^3$ be an embedding of T onto a neighbourhood of a knot C, so that e maps a longitude of T (coming from the inclusion of T in S^3) onto a longitude of C. Then*

$$\Delta_{eK}(t) \doteq \Delta_K(t)\Delta_C(t^n),$$

where K represents n times a generator of $H_1(T)$.

PROOF. Construct Seifert surfaces for the pattern knot K and the satellite eK in the following way: The unknotted solid torus T projects onto an annulus in the plane. Apply the Seifert method (Theorem 2.2) to the projection of K, with some orientation, into this annulus. Seifert circuits in the annulus, connected by twisted strips at the crossings, are obtained. Cap off, with discs just above the annulus, any circuits that bound in the annulus; then use annuli to cap off adjacent pairs of curves that encircle the annulus in opposite directions. Add a vertical annulus to each remaining curve so that the result is an oriented surface F contained in T, with ∂F being the union of K and n longitudes of T oriented in the same direction. A Seifert surface $F \cup nD$ for K then consists of the union of F and n parallel copies of a spanning disc of T. Similarly, a Seifert surface $eF \cup nG$ for eK consists of the union of eF and n parallel copies of a genus g Seifert surface G of the companion knot C (this G being regarded as in the closure of $S^3 - eT$).

Note that if f is an oriented simple closed curve in $T - K$, then $\operatorname{lk}(f, K) = f \sqcap F$, where "$\sqcap$" denotes the algebraic sum of the transverse intersection points. Of course, $f \sqcap F = ef \sqcap eF = ef \sqcap (eF \cup nG) = \operatorname{lk}(ef, eK)$. Thus linking numbers of curves in T are preserved by the embedding e. Note, as well, that if f' is a simple closed curve in the interior of G (or is near to such a curve), then $\operatorname{lk}(ef, f') = 0$. This is because ef is homologous in eT to a sum of longitudes of C, and they bound copies of G that can be taken to be disjoint from f'.

A Seifert matrix B for the satellite knot eK can be obtained as follows: Use as base for $H_1(eF \cup nG)$ the image under e of curves in F that give a base for $H_1(F \cup nD)$, together with n parallel copies, each in one of the n copies of G, of curves that provide a base for $H_1(G)$. Using the above remarks, the resulting Seifert matrix has the form $\begin{pmatrix} M & 0 \\ 0 & X \end{pmatrix}$, where M is a Seifert matrix for K and X is the following $n \times n$ block matrix, in which A is a Seifert matrix for C:

$$X = \begin{pmatrix} A & A & A & \dots & A \\ A^\tau & A & A & \dots & A \\ A^\tau & A^\tau & A & \dots & A \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A^\tau & A^\tau & A^\tau & \dots & A \end{pmatrix}.$$

It is consideration of linking numbers of curves in the various parallel copies of G that gives rise to these off-diagonal copies of A and A^τ.

Consider now the linear combination $\sum_{i=1}^{n} t^{n-i}(\text{row } i)$ of the rows of blocks of the block matrix

$$tX - X^\tau = \begin{pmatrix} tA - A^\tau & tA - A & tA - A & \dots & tA - A \\ tA^\tau - A^\tau & tA - A^\tau & tA - A & \dots & tA - A \\ tA^\tau - A^\tau & tA^\tau - A^\tau & tA - A^\tau & \dots & tA - A \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ tA^\tau - A^\tau & tA^\tau - A^\tau & tA^\tau - A^\tau & \dots & tA - A^\tau \end{pmatrix}.$$

That linear combination produces a row of blocks in which every entry is $t^n A - A^\tau$. Thus, replacing the first row of $tX - X^\tau$ by this row and subtracting the first column from all the other columns, it is seen that

$$\det(tX - X^\tau) = t^{-2g(n-1)}\det\begin{pmatrix} t^n A - A^\tau & 0 & 0 & \dots & 0 \\ \star & t(A - A^\tau) & \star & \dots & \star \\ \star & 0 & t(A - A^\tau) & \dots & \star \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \star & 0 & 0 & \dots & t(A - A^\tau) \end{pmatrix}.$$

Now, by Theorem 6.10, $\det(A - A^\tau) = 1$, so that $\det t(A - A^\tau) = t^{2g}$. Thus

$$\det(tB - B^\tau) = \det(tM - M^\tau)\det(t^n A - A^\tau),$$

and this is the required formula. □

In Chapter 8, a Conway normalisation of the Alexander polynomial will be defined, and then the above result will become $\Delta_{eK}(t) = \Delta_K(t)\Delta_C(t^n)$.

Corollary 6.16.
(i) *If $\Delta_{C_1}(t) = \Delta_{C_2}(t)$, then satellites of C_1 and C_2 with the same pattern have the same Alexander polynomial.*
(ii) *Reversing the direction of C has no effect on $\Delta_{eK}(t)$ (though it can change the knot eK).*
(iii) *A Whitehead double of any knot has Alexander polynomial equal to* 1.

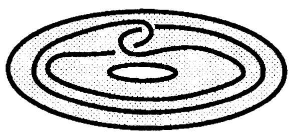


Figure 6.5

The final statement needs a little clarification. A *Whitehead double* of C is a satellite formed by using for $K \subset T$ the curve shown in Figure 6.5 or its reflection. Note that K is unknotted in S^3 and represents zero in $H_1(T)$, so that $n = 0$ in the formula of Theorem 6.15.

There is no formula for the Jones polynomial of a satellite knot analogous to that just proved for the Alexander polynomial. Indeed, the fact that interesting phenomena are encountered when searching for such an analogue underlies the discussion of Chapter 13.

One further satisfying view of the Alexander polynomial of a *knot* gives an interpretation of it as a characteristic polynomial in the following way: Suppose that throughout the preceding theory the field of rational numbers, $\mathbb{Q}$, is used instead of the ring of integers, $\mathbb{Z}$. Not very much would is changed. In particular, if A is a

Seifert matrix, the matrix $(tA - A^\tau)$ presents $H_1(X_\infty; \mathbb{Q})$ as a $\mathbb{Q}[t^{-1}, t]$-module. Information about the elementary ideals of this (marginally) new module can be extracted from $(tA - A^\tau)$ as before, though in general the information obtained is slightly weaker than when using integer coefficients. However, a generator of the first elementary ideal is still $\det(tA - A^\tau)$. Thus the Alexander polynomial of the knot is, up to multiplication by a unit (now an element of the form $qt^{\pm n}$ for any $q \in \mathbb{Q}$), equal to the determinant of any other square matrix representing this new module.

Theorem 6.17. *Let K be a knot in S^3 and let $t : X_\infty \to X_\infty$ be the (covering) translation of X_∞ (the infinite cyclic cover of the exterior of K). Then $H_1(X_\infty; \mathbb{Q})$ is a finite-dimensional vector space over the field $\mathbb{Q}$. The characteristic polynomial of the linear map $t_\star : H_1(X_\infty; \mathbb{Q}) \to H_1(X_\infty; \mathbb{Q})$ is, up to multiplication by a unit, equal to the Alexander polynomial of K.*

PROOF. The ring $\mathbb{Q}[t^{-1}, t]$ is a principal ideal domain. A proof of this, using the Euclidean algorithm, is much the same as the proof that shows the ring of ordinary polynomials over a field to be a principal ideal domain. Over $\mathbb{Q}[t^{-1}, t]$ the module $H_1(X_\infty; \mathbb{Q})$ is finitely presented by the matrix $(tA - A^\tau)$. However, over a principal ideal domain, any finitely presented module is just a direct sum of cyclic modules (see, for example, [38]). This is the same as saying that the module is presented by a square diagonal matrix. Thus $H_1(X_\infty; \mathbb{Q})$ is presented by a matrix $\text{diag}(p_1, p_2, \ldots, p_N)$, where $p_i \in \mathbb{Q}[t^{-1}, t]$, and $H_1(X_\infty; \mathbb{Q})$ is isomorphic as a module to $\bigoplus_{i=1}^N (\mathbb{Q}[t^{-1}, t]/p_i)$. None of the p_i is zero, for then the Alexander polynomial, the determinant of the matrix, would be zero. However, for a *knot* K, $\Delta_K(1) = \pm 1$.

Consider, then, a typical summand of the form $\mathbb{Q}[t^{-1}, t]/p$ where, multiplying by a unit, it may be assumed that $p = a_0 + a_1 t + a_2 t^2 + \cdots + a_r t^r$ with $a_r = 1$. Over the field $\mathbb{Q}$, the vector space $\mathbb{Q}[t^{-1}, t]/p$ has a *finite* base $\{1, t, t^2, \ldots, t^{r-1}\}$, for the relation "$p = 0$" expresses other powers of t linearly in terms of these. Of course, the action of $t_\star$ is just multiplication by t. With respect to this base, then, $t_\star$ is represented by the matrix

$$M = \begin{pmatrix} 0 & 0 & 0 & & & -a_0 \\ 1 & 0 & 0 & & & -a_1 \\ 0 & 1 & 0 & & & -a_2 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & \ldots & 1 & 0 & -a_{r-2} \\ 0 & 0 & 0 & \ldots & 1 & -a_{r-1} \end{pmatrix}.$$

As a polynomial in x, the characteristic polynomial of this is the determinant of $(M - xI)$. Multiplying the i^{th} row of this matrix by x^{i-1} and, for $i \geq 2$, adding it to the top row, this determinant is seen to be the determinant of

$$\begin{pmatrix} 0 & 0 & 0 & & & -\sum_{i=0}^{r} a_i x^i \\ 1 & -x & 0 & & & -a_1 \\ 0 & 1 & -x & & & -a_2 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & -x & -a_{r-2} \\ 0 & 0 & 0 & \dots & 1 & -x - a_{r-1} \end{pmatrix},$$

which is $(-1)^r \sum_{i=0}^{r} a_i x^i$. This is $\pm p$. Now, up to a unit, the Alexander polynomial is the determinant of the presentation matrix $\mathrm{diag}(p_1, p_2, \dots, p_N)$ for $H_1(X_\infty; \mathbb{Q})$. This is just $\prod_{i=1}^{N} p_i$, and the above consideration applied to the summands of $\bigoplus_{i=1}^{N}(\mathbb{Q}[t^{-1}, t]/p_i)$ shows that (with x in place of t) this is the characteristic polynomial of $t_\star$. □

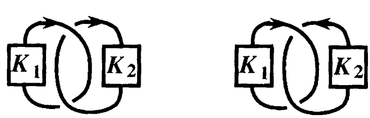
For more on the Alexander polynomial viewed as part of algebraic topology, see the survey by C. McA. Gordon [35].

Exercises

1. Find a Seifert surface F for the knot 7_3, select a convenient base for $H_1(F; \mathbb{Z})$ and find the Seifert matrix with respect to this base. Calculate the Alexander polynomial of 7_3 and check that your answer agrees with that given in the table of Alexander polynomials.

2. Calculate the Alexander polynomial of the two oriented links shown below.

3. Determine the way that the Alexander polynomial of each of the oriented links shown below is related to the Alexander polynomials of knots K_1 and K_2.



4. Show that for a knot K, the Alexander polynomial satisfies $\Delta_K(t) \doteq 1$ if and only if $H_1(X_\infty; \mathbb{Z}) = 0$.

5. What polynomials can arise as Alexander polynomials of genus 1 knots?

6. Figure 12.7 (b) shows (neglecting the zeros) a very symmetric diagram of a three-component link called the Borromean rings. Different choices of directions for the components produce eight possible orientations for the link. Calculate the Alexander polynomial for each of the oriented links so formed.

7. Suppose that B is any $2n \times 2n$ matrix of integers with the property that $B - B^\tau$ consists of n blocks of the form $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ running down the diagonal and zeros elsewhere. Prove that there exists a knot for which B is a Seifert matrix.

8. Calculate the Alexander polynomial of the knot K shown below. What is the genus of K? Is K a prime knot?

9. Show that for any knot K with Alexander polynomial $\Delta_K(t)$, there is, for any positive integer n, another knot with Alexander polynomial $\Delta_K(t^n)$.

10. Show that any knot C has a (non-trivial) satellite knot of genus 1 with the same Alexander polynomial as the trefoil knot 3_1.

11. A fibred knot K is a knot with the property that its exterior X is a bundle over S^1 with fibre an orientable surface F. This means that X is homeomorphic to $F \times [0, 1]$ quotiented by the identification $(x, 0) \equiv (hx, 1)$ for some homeomorphism $h : F \to F$. What is the Alexander polynomial of such a knot K? Prove that $g(K)$ is the genus of the surface F. If a genus 1 knot is fibred, what can be said about its Alexander polynomial?

7

Covering Spaces

In order to bring to a satisfactory conclusion the theory of the last chapter, it is necessary to show that the space X_∞, together with the given action on it by the infinite cyclic group, is uniquely defined by the oriented link L under consideration. Here it will be seen that X_∞ is a certain covering space of the exterior of L, and the theory of coverings will show it to be well defined. That is the present motivation, but it should be understood that the theory of covering spaces is an important part of many areas of mathematics (particularly Riemann surfaces and geometric structures on manifolds). It is intimately related to the study of the (appropriately named) fundamental group of a fairly general type of topological space. Thus the following discussion will be in the language of general topological spaces.

In the whole of this chapter, B will be a path-connected, locally path-connected topological space. By definition, the locally path-connected condition means that each point has a base of path-connected neighbourhoods (that is, there are "arbitrarily small" such neighbourhoods for each point).

Definition 7.1. A continuous map $p : E \to B$ is a covering map if

(i) E is path-connected and non-empty and

(ii) for each $b \in B$, there exists an open neighbourhood V of B such that $p^{-1}V$ is a disjoint union of open sets in E, each of which is mapped homeomorphically by p onto V.

The map p is called the projection of the covering space E to the base space B.

A covering map $p : E \to B$ is, in other terminology, a locally trivial fibre map with discrete fibre. As an exercise, observe that the restriction of the covering map p to any proper subset of E fails to give a covering of B.

Examples 7.2.

(i) $p : \mathbb{R} \longrightarrow S^1 = \{z \in \mathbb{C} : |z| = 1\}$ given by $p(t) = \exp(2\pi i t)$.

(ii) $p : S^1 \longrightarrow S^1$ given by $p(z) = z^n$.

(iii) $p : S^n \longrightarrow \mathbb{R}P^n = S^n/(x \sim -x)$, where p is the quotient map.

(iv) $\overline{p} : S^3 \longrightarrow L_{p,q}$, where for p and q coprime integers, $L_{p,q}$ is the lens space defined as the quotient of S^3 by a certain action of the cyclic group of order p. Regard S^3 as $\{(z_1, z_2) \in \mathbb{C}^2 : |z_1|^2 + |z_2|^2 = 1\}$. If g generates the group, the action is defined by $g(z_1, z_2) = (z_1 \exp(2\pi i/p),\ z_2 \exp(-2\pi i q/p))$. The projection $\overline{p}$ is the quotient map.

Easy Properties 7.3.

(i) The covering map $p : E \to B$ maps open sets to open sets. It is *locally* a homeomorphism. In particular, E is locally path-connected.

(ii) The covering map $p : E \to B$ is surjective.

(iii) The open set V of the definition can be taken to be path-connected.

(iv) B has the quotient topology induced by $p : E \to B$.

(v) If b_1 and b_2 belong to B, then there is a bijection between $p^{-1}b_1$ and $p^{-1}b_2$ (this follows from the next lemma).

Lemma 7.4. *A covering map $p : E \to B$ has the path lifting property. That is, given a point $e_0 \in E$ and a continuous map $f : [0, 1] \to B$ such that $f(0) = p(e_0)$, there exists a unique continuous map $\widehat{f} : [0, 1] \to E$ such that $\widehat{f}(0) = e_0$ and $p\widehat{f} = f$.*

PROOF. The space B is the union of open sets $\{V\}$, as in the definition of a covering. Thus, by the compactness of $[0, 1]$ there is a dissection $0 = t_0 < t_1 < t_2 < \cdots < t_n = 1$ so that $f[t_{i-1}, t_i] \subset V_i$ for some such open set V_i. Assume that $\widehat{f} \mid [0, t_{i-1}]$ has been defined with $\widehat{f}(t_{i-1}) \in W_{i,j}$ where $W_{i,j}$ is one of the open subsets of $p^{-1}V_i$ for which $p : W_{i,j} \to V_i$ is a homeomorphism. Define $\widehat{f} \mid [t_{i-1}, t_i]$ to be equal to $(p \mid W_{i,j})^{-1} f$. For the uniqueness, suppose $\widehat{\phi}$ is a second lift of f, with $\widehat{\phi}(0) = e_0$. Let $\tau = \sup\{t : \widehat{\phi} \mid [0, t] = \widehat{f} \mid [0, t]\}$; by continuity, $\widehat{\phi}(\tau) = \widehat{f}(\tau)$. Then, if $\tau < 1$, the above argument shows that $\widehat{\phi}(\tau + \epsilon) = \widehat{f}(\tau + \epsilon)$ for all sufficiently small ϵ, contradicting the definition of τ. □

Lemma 7.5. *A covering map $p : E \to B$ has homotopy-lifting property for paths. That is, given a continuous map $\widehat{f} : [0, 1] \times \{0\} \to E$ and a continuous map $f : [0, 1] \times [0, 1] \to B$ such that $f(t, 0) = p\widehat{f}(t, 0)$, there exists a unique continuous extension of $\widehat{f}$ to $\widehat{f} : [0, 1] \times [0, 1] \to E$ such that $p\widehat{f} = f$.*

PROOF. The proof of this is entirely analogous to the proof of the previous lemma; here a dissection of the square $[0, 1] \times [0, 1]$ into a mesh of small squares, each mapping into some V_i, is used. □

Elementary homotopy theory assigns to every topological space X, equipped with a selected base point x_0, a group $\Pi_1(X, x_0)$ called its *fundamental group*. Recall that an element of the fundamental group is represented by a loop in X based at x_0 (that is, a continuous function $\alpha : [0, 1] \to X$ with $\alpha(0) = \alpha(1) = x_0$), an actual element being a homotopy class, keeping ends fixed at x_0, of such loops. The product of loops α and β, written $\alpha \cdot \beta$, is formed by following around the loop α and then β; the inverse of α is the loop $\overline{\alpha}$, where $\overline{\alpha}(t) = \alpha(1 - t)$. These operations induce the group structure on the homotopy classes. A continuous function f from one based space to another induces a homomorphism $f_\star$ between their fundamental groups with the usual functorial properties. In particular, homeomorphic based spaces have isomorphic fundamental groups. A path in X from x_0 to x_1 induces, by means of path-composition, an isomorphism from $\Pi_1(X, x_0)$ to $\Pi_1(X, x_1)$, the isomorphisms induced by different paths being related by inner automorphisms. Thus usually one restricts consideration to path-connected spaces, and then choice of base point is irrelevant up to group isomorphism; the base point is then often omitted from the notation. However, base points can never be neglected completely; any attempt to do so usually produces the first homology group. In general the fundamental group of a space is *not* abelian. If one "makes it abelian" by inserting relations that declare that all elements commute, then the result is indeed the first homology group. This is a one-dimensional version of the Hurewicz isomorphism theorem: for a connected cell complex X, the quotient of $\Pi_1(X, x_0)$ by its commutator subgroup (the subgroup generated by all elements of the form $aba^{-1}b^{-1}$) is isomorphic to $H_1(X; \mathbb{Z})$.

The Homotopy Exact Sequence 7.6. An immediate consequence of Lemma 7.5 is the following homotopy exact sequence for a covering map:

$$\{1\} \longrightarrow \Pi_1(E, e_0) \xrightarrow{p_\star} \Pi_1(B, b_0) \longrightarrow \Pi_0(F) \longrightarrow \{1\}.$$

Here $p(e_0) = b_0$ and $F = p^{-1}b_0$. Note that $\Pi_0(F)$ is just the *set* of path components of F (which are just the individual points of F) with a "zero", the component $\{e_0\}$. The map $\Pi_1(B, b_0) \to \Pi_0(F)$ is defined as follows. A loop α in B based at b_0 lifts to a path $\widehat{\alpha}$ starting at e_0. The required map sends the element $[\alpha]$ represented by α to $\widehat{\alpha}(1)$.

In this theory of lifting a path (or homotopy of paths) in the base space to a path in a covering space, one thinks of E as "above" B so that "lifting" has some intuitive feel about it. The next result answers speculation about whether a map from any space into B might be lifted. The answer, for a reasonable type of space, is that it can be lifted unless fundamental group considerations forbid the enterprise.

Proposition 7.7. *Let $p : E \to B$ be a covering map with base points $e_0 \in E$ and $b_0 \in B$, chosen so that $pe_0 = b_0$. Suppose X is a path-connected, locally path-connected, space with base point x_0, and let $f : (X, x_0) \to (B, b_0)$ be continuous. Then there exists a continuous map $g : (X, x_0) \to (E, e_0)$ such that $pg = f$ if and only if*

$$f_\star \Pi_1(X, x_0) \subset p_\star \Pi_1(E, e_0).$$

When such a g exists, it is unique.

PROOF. If g exists, then $p_\star g_\star = f_\star$, and the result is clear. Conversely, suppose $f_\star \Pi_1(X, x_0) \subset p_\star \Pi_1(E, e_0)$. If $x \in X$, choose a path $\alpha : [0, 1] \to X$ so that

$\alpha(0) = x_0$ and $\alpha(1) = x$. By Lemma 7.4, the path $f\alpha$ lifts to a path $\widehat{f\alpha} : [0, 1] \to E$ with $\widehat{f\alpha}(0) = e_0$. Note that if g exists as advertised, then $g(x) = \widehat{f\alpha}(1)$ by the uniqueness in Lemma 7.4, because $g\alpha$ is a lift of $f\alpha$. Thus if g exists, it is unique. Now *define* g by $g(x) = \widehat{f\alpha}(1)$. To check that is well defined, let β be another path in X from x_0 to x_1. Then $f_\star[\alpha \cdot \overline{\beta}] \in f_\star \Pi_1(X, x_0) \subset p_\star \Pi_1(E, e_0)$, so there exists a loop $\gamma : [0, 1] \to E$ with $\gamma(0) = e_0 = \gamma(1)$ so that $p\gamma$ is homotopic, relative to $\{0, 1\}$, to $f(\alpha \cdot \overline{\beta})$. By Lemma 7.5 that homotopy can be lifted, relative to $\{0, 1\}$, so that (at the end of the homotopy) there is a loop $\widetilde{\gamma} : [0, 1] \to E$ with $\widetilde{\gamma}(0) = e_0 = \widetilde{\gamma}(1)$ such that $p\widetilde{\gamma} = f(\alpha \cdot \overline{\beta})$. Thus the lift of $f(\alpha \cdot \overline{\beta})$ starting at e_0 is $\widetilde{\gamma}$, a *loop* at e_0. Hence $p\widetilde{\gamma}(t) = f\alpha(2t)$ and $p\overline{\widetilde{\gamma}}(t) = f\beta(2t)$ for all $0 \leq t \leq 1/2$. Thus $\widehat{f\alpha}(1) = \widetilde{\gamma}(1/2) = \widehat{f\beta}(1)$, and so g is well defined. The continuity of g follows from the fact that X is locally path-connected, and so on sufficiently small open sets g is $p^{-1}f$. □

Suppose $p : (E, e_0) \to (B, b_0)$ is a covering map with base points as above. The subgroup $p_\star\Pi_1(E, e_0)$ of $\Pi_1(B, b_0)$ is called the *group of the covering*. Note that $p_\star\Pi_1(E, e_0)$ is, as explained in the above proof, the set of homotopy classes, relative to $\{0, 1\}$, of loops $\alpha : [0, 1] \to B$ based at b_0 such that $\widehat{\alpha}(1) = e_0$, that is, such that α lifts to a *loop*. Note too that $p_\star$ is injective (from the homotopy exact sequence) so that $p_\star\Pi_1(E, e_0)$ is isomorphic to $\Pi_1(E, e_0)$.

Proposition 7.8. *Suppose $p : (E, e_0) \to (B, b_0)$ and $p' : (E', e_0') \to (B, b_0)$ are two based coverings of B with the same group. Then these are equivalent in the sense that there exists a homeomorphism $h : (E', e_0') \to (E, e_0)$ such that $ph = p'$.*

PROOF. By Proposition 7.7 , the map p' lifts to a *map* $h : (E', e_0') \to (E, e_0)$ such that $ph = p'$. Similarly, by Proposition 7.7 applied to the map p and covering p', there is a map $h' : (E, e_0) \to (E', e_0')$ such that $p'h' = p$. But then $hh' : (E, e_0) \to (E, e_0)$ is a lift of the map p with respect to the covering p. The identity map is another such lift. Hence, by the uniqueness of Proposition 7.7, hh' is the identity. Similarly, $h'h$ is the identity, and so h and h' are mutually inverse homeomorphisms. □

Now recall from Chapter 6 the map $p : X_\infty \to X$, where X is the exterior of an oriented link L, F is a Seifert surface and X_∞ is the space constructed by gluing together countably many copies of Y, where Y is X-cut-along-F.

Theorem 7.9. *The covering space $p : X_\infty \to X$ of the exterior X of an oriented link L does not depend on the choice of Seifert surface used in its construction. Further, the action of the infinite cyclic group on X_∞ is likewise independent of F.*

PROOF. It is clear from the construction of X_∞ that a loop $\alpha : [0, 1] \to X$ lifts to a *loop* $\widehat{\alpha}$ (that is, $\widehat{\alpha}(0) = \widehat{\alpha}(1)$) in X_∞ provided $\widehat{\alpha}(0)$ and $\widehat{\alpha}(1)$ are in the same copy of Y. This is so if and only if α intersects F zero times algebraically, for every time α crosses F, its lift moves from one copy of Y to an adjacent copy. Thus α lifts to a loop if and only if the linking number of α with L (that is, the sum of the linking numbers with the components of L) is zero. Now, that statement is independent of the choice of Seifert surface for L, so the group of the cover does not depend on F. Using the preceding proposition, the the first result follows at once.

Consider the action by the infinite cyclic group $\langle t \rangle$ on X_∞. If $\gamma : [0, 1] \to X_\infty$ is any path from some point a to ta, then, by the above reasoning, $p\gamma$ is a loop in X having linking number 1 with L. Conversely the lift of any such loop in x is a path from some a to ta. Suppose $p' : X'_\infty \to X$ is a second version of X_∞ constructed from Seifert surface F' and $h' : X_\infty \to X'_\infty$ is the homeomomorphism such that $p'h' = p$. Trivially $p'h'\gamma = p\gamma$, so that $h'\gamma$, being a lift of the loop $p\gamma$ with respect to the covering p', is a path in X'_∞ from a point to its t-translate. Hence $th'(a) = h'(ta)$, and the homeomorphism h' preserves the t-action. □

This then concludes the proof of the fact that the Alexander polynomial of an oriented link L is well defined (up to multiplication by a unit). The covering space X_∞ of X is called the infinite cyclic covering of the link exterior. A loop in X lifts to a *loop* in X_∞ if and only if it has zero linking number with L. In the case when L is a knot, this means, by Theorem 1.5, that the loop represents the zero element in $H_1(X)$. Then $p_\star \Pi_1(X_\infty)$ is the kernel of the natural map $\Pi_1(X) \to H_1(X)$, which, for any X, is the commutator subgroup of $\Pi_1(X)$.

In determining the Alexander polynomial of a knot, any convenient method of constructing X_∞ may be used. It is just necessary to construct a covering of the knot exterior with the property that a loop lifts to a loop if and only if it has zero linking number with the knot. The diagrams of Figure 7.1 show such a method for the exterior of the 4-crossing knot. The exterior of the knot 4_1 in the first diagram can be obtained by the following "surgery" procedure on the exterior of the knot of the second diagram (which is unknotted). Remove a (shaded) solid torus as shown and replace it with a solid torus in such a way that on the boundary of the (shaded) toral hole, the curve shown bounds a disc in the replacing torus. To contemplate that replacement, imagine cutting across a disc spanning the outside of the toral hole. This creates discs on either side of the cut. Twist one of these discs through 2π about an axis through its centre, thereby reinserting two crossings into the knot; then glue the discs together again. The curve on the boundary of the hole has been changed to become a meridian of the toral hole. Then the solid torus fits neatly into the hole, and the first diagram is recreated.

The third diagram is the same as the second, up to isotopy; care has been taken to keep track of the curve on the boundary of the toral hole. Now the infinite cyclic cover can be created by cutting across a disc spanning the unknot in this diagram and taking infinitely many copies glued end to end. The result is a copy of $D^2 \times \mathbb{R}$ from which infinitely many solid tori have been removed, as shown, and which are to be replaced so that the indicated curves become the boundaries of discs. (For this to happen it is important that the shaded solid torus was chosen to have zero linking number with the knot.) The t action on the cover is "translation to the right by one unit". Then $H_1(X_\infty; \mathbb{Z})$ is generated as a module by the class of the

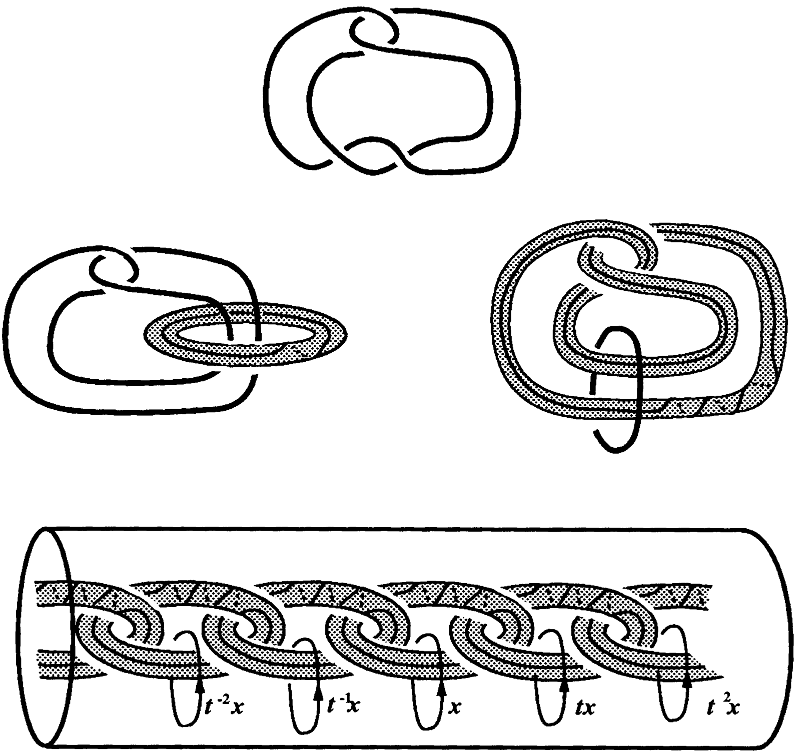


Figure 7.1

curve x shown in the diagram, and there is one relator represented by the curve shown on the boundary of one toral hole. (The relators corresponding to curves on the other toral holes are translates of the first by powers of t.) This relator is $-t^{-1}x + 3x - tx$. Thus the module is represented by the 1×1 matrix $-t^{-1} + 3 - t$, and so (taking its determinant) $-t^{-1} + 3 - t$ is the Alexander polynomial of the knot 4_1.

Note that the essence of the preceding discussion is that the diagram of the knot 4_1 can be changed to a diagram of the unknot by changing one crossing. That crossing is then encircled by the shaded solid torus. If K is a knot with a diagram that can be unknotted with m crossing changes, then the procedure can be repeated using m solid tori, each encircling one of these crossings. The result is a presentation of the Alexander module with m generators and m relators. Thus this module has an $m \times m$ presentation matrix, and so its rth elementary ideal is $\mathbb{Z}[t^{-1}, t]$ for every $r > m$. This proves the following result about unknotting numbers (see Chapter 1).

Theorem 7.10. *If the rth elementary ideal of the Alexander module of a knot K is not the whole of $\mathbb{Z}[t^{-1}, t]$, then K has unknotting number $u(K) \geq r$.*

As an example, consider the pretzel knot $P(3, 3, -3)$ discussed in Example 6.9. There it was shown that the second elementary ideal of the Alexander module is not $\mathbb{Z}[t^{-1}, t]$, and so $u(P(3, 3, -3)) \geq 2$. It is easy to see that two crossing changes do undo the knot, and so $u(P(3, 3, -3)) = 2$. More information on the results of this technique can be found in [103].

Another example of a covering may be useful. Let $X_\infty \to X$ be, as before, the infinite cyclic covering of the exterior of an oriented n-component link L. The cyclic group $\langle t \rangle$ acts on X_∞. Then $X_\infty/\langle t^2 \rangle \to X$ is a 2-fold covering of X called the *cyclic double cover* of X. Denote $X_\infty/\langle t^2 \rangle$ by $\widehat{X_2}$. (This $\widehat{X_2}$ can, if desired, be constructed from two copies of Y, where Y is X cut along a Seifert surface, gluing together parts of the boundary in the obvious way.) A loop in X lifts to a loop in $\widehat{X_2}$ if and only if it has linking number zero modulo 2 with L. The covering is that corresponding to the kernel of the map $\Pi_1(X, x_0) \to H_1(X; \mathbb{Z}) \to \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$, where the second map sends each meridian to $1 \in \mathbb{Z}$. Consider loops on the boundary of the solid torus neighbourhood N_i of any component L_i of L. A longitude lifts to a loop in $\widehat{X_2}$. A meridian does not lift to a loop, but the square of a meridian does lift to a loop. Thus, identifying ∂N_i with $S^1 \times S^1$, with longitude and meridian corresponding to the two factors, the covering restricted to the part of it over ∂N_i is a covering of a torus by a torus. It is equivalent to $(z_1, z_2) \mapsto (z_1, z_2^2)$, where S^1 is the unit complex numbers. (This is also clear from the construction of $\widehat{X_2}$ by gluing together two copies of Y.) That map extends to a map $S^1 \times D^2 \to S^1 \times D^2$ defined by $(z_1, z_2) \mapsto (z_1, z_2^2)$. This is a covering map *except* on $S^1 \times \{0\}$. It is called a covering *branched* over $S^1 \times \{0\}$. Thus n solid tori can be glued to the boundary components of $\widehat{X_2}$ to create X_2, another n solid tori can be glued to the boundary of X to recreate S^3 and the double covering map $\widehat{X_2} \to X$ can be extended, as described above over each solid torus, to achieve a map $X_2 \to S^3$ called the double cover of S^3 branched over L. This is a two-fold cover when restricted to (a map to) the complement of L. Note that this construction is independent of the orientation of L, since $1 = -1$ in $\mathbb{Z}/2\mathbb{Z}$. The construction can be generalised at once to construct an r-fold cyclic cover of S^3 branched over an oriented link.

Two-bridge links provide a simple example. As explained in Chapter 1, a 2-bridge (or rational) link is obtained by gluing together the boundaries of two trivial 2-string tangles. The double cover of a ball branched over a trivial 2-string tangle is a solid torus. Thus the double cover of S^3 branched over the link is two solid tori with their boundaries glued together. That is a lens space or, exceptionally, S^3 or $S^1 \times S^2$. In fact, $L_{p,q}$ is the double cover of S^3 branched over the (p, q) 2-bridge link. It has been shown that only this link has $L_{p,q}$ as its double branched cover [46].

Further facts about covering spaces will be useful in Chapter 11. As has been noted, a covering map $p : (E, e_0) \to (B, b_0)$ induces an injection on fundamental groups, and $p_\star \Pi_1(E, e_0)$ is called the group of the covering. The chief further result is that provided B is "semi-locally simply connected" (locally contractible will do fine), then for any given subgroup G of $\Pi_1(B, b_0)$ there exists a covering space with G as group. Conjugate subgroups produce equivalent (base point free) covers in the sense that $p : E \to B$ and $p' : E' \to B$ are equivalent if there is a homeomorphism $h : E \to E'$ such that $p'h = p$. All this is fairly simple once one can do the theory when G is the trivial one-element subgroup of $\Pi_1(B, b_0)$.

Definition 7.11. A covering $p : \widetilde{B} \longrightarrow B$ in which $\widetilde{B}$ is simply connected is called the universal covering of B.

Note that by Lemma 7.8 a space B has at most one universal covering up to equivalence. The aim now will be to show that a path-connected and locally path-connected space B, with one extra property, does have a simply connected covering space. The definition of the extra property is as follows:

Definition 7.12. A space B is semi-locally simply connected if for each $b \in B$ there exists a neighbourhood V of b with the property that every closed curve in V is null-homotopic in B.

Note that if B has this property, then the set V can be taken to be open and path-connected. The property is then the same as the assertion that inclusion induces the constant map $\Pi_1(V) \to \Pi_1(B)$. Suppose there does exist a covering map $p : E \to B$ for some simply connected space E. If $b \in B$, there is (from the definition of a covering map) an open set $U \subset E$ such that $p|U : U \to V$ is a homeomorphism onto some open neighbourhood V of b. Then $pi_U = i_V(p|U)$, where i_U and i_V are inclusion maps. Of course, $(i_U)_\star : \Pi_1(U) \to \Pi_1(E)$ is constant because $\Pi_1(E)$ is trivial, and as $(p|U)_\star$ is an isomorphism, it follows that $(i_V)_\star$ is the trivial constant map. Thus the semi-locally simply connected condition is certainly needed if B is to have a simply connected cover. Note that if B is a manifold or a finite complex, it certainly has this property (as V can be taken to be contractible).

Theorem 7.13. *Let B be a path-connected, locally path-connected, semi-locally simply connected space. Then there exists a simply connected space $\widetilde{B}$ and covering map $p : \widetilde{B} \longrightarrow B$. Furthermore, the group $\Pi_1(B)$ acts freely as a group of homeomorphisms on (the left of) $\widetilde{B}$, the quotient map $q : \widetilde{B} \longrightarrow \widetilde{B}/\Pi_1(B)$ is a covering map and there is a homeomorphism $h : \widetilde{B}/\Pi_1(B) \longrightarrow B$ such that $hq = p$.*

PROOF. Let $b_0 \in B$ be a base point and let X be the set of all paths $\alpha : [0, 1] \to B$ such that $\alpha(0) = b_0$. Define an equivalence relation on X by letting $\alpha \sim \beta$ if and only if $\alpha(1) = \beta(1)$ and $\alpha \approx \beta$, where "$\approx$" denotes homotopy of paths in B keeping the end points $\{0, 1\}$ fixed. Let $\widetilde{B}$ be the quotient set $X/\sim$ and define $p : \widetilde{B} \to B$ by $p[\alpha] = \alpha(1)$, where $[\alpha]$ is the equivalence class of α. Suppose that $\alpha \in X$ and that V is an open neighbourhood of $\alpha(1)$ in B. Let $\langle \alpha, V \rangle \subset \widetilde{B}$

be defined by

$$\langle \alpha, V \rangle = \{[\alpha \cdot \beta] : \beta : [0, 1] \to V, \ \beta(0) = \alpha(1)\}.$$

Take all possible $\langle \alpha, V \rangle$ to be a base for a topology on $\widetilde{B}$ (so that a subset of $\widetilde{B}$ is defined to be open if and only if it is a union of some of these basic sets). Note that if $[\alpha] \in \langle \alpha_1, V_1 \rangle \cap \langle \alpha_2, V_2 \rangle$, then

$$\langle \alpha, V_1 \cap V_2 \rangle \subset \langle \alpha_1, V_1 \rangle \cap \langle \alpha_2, V_2 \rangle,$$

so that the given sets do form a base of a genuine topology.

Now $p\langle \alpha, V \rangle$ is the path component of V that contains $\alpha(1)$. This is open in B, since B is locally path-connected, so p maps open sets to open sets. Further, if V is open in B, then

$$p^{-1}V = \bigcup_{\alpha} \{\langle \alpha, V \rangle : \alpha(1) \in V\}.$$

By definition this is open, and so p is continuous. The space $\widetilde{B}$ is path-connected, since $[\alpha]$ is joined to the class of the path that is constant at b_0 by $\{[\alpha_s] : s \in [0, 1]\}$, where $\alpha_s(t) = \alpha(st)$.

If V is open in B and $[\gamma] \in \langle \alpha, V \rangle$, then $\langle \gamma, V \rangle = \langle \alpha, V \rangle$. Thus any $\langle \alpha, V \rangle$ and $\langle \beta, V \rangle$ are either disjoint or identical, and so $p^{-1}V$ is the disjoint union of open sets of the form $\langle \alpha, V \rangle$. If $b \in B$, use the given properties of B to select, an open path-connected neighbourhood V of b for which $\Pi_1(V) \to \Pi_1(B)$ is the trivial map. Then p is injective on $\langle \alpha, V \rangle$. This is because if $p[\alpha \cdot \beta] = p[\alpha \cdot \beta']$, where β and β' are paths in V with the same end points, then $\beta \approx \beta'$, so that $\alpha \cdot \beta \approx \alpha \cdot \beta'$ and hence $[\alpha \cdot \beta] = [\alpha \cdot \beta']$. Thus $p : \langle \alpha, V \rangle \to V$ is a homeomorphism and, as $p^{-1}(V)$ is a disjoint union of sets of the form $\langle \alpha, V \rangle$, p is a covering map.

Suppose that $[\gamma] \in \Pi_1(B, b_0)$, where γ is a loop based at b_0. Define a map $[\gamma] : \widetilde{B} \to \widetilde{B}$ by $[\gamma]([\alpha]) = [\gamma \cdot \alpha]$. This gives a well-defined map that sends basic open sets to basic open sets and, as it has $[\overline{\gamma}]$ as an inverse, it is a homeomorphism. Thus the group $\Pi_1(B, b_0)$ acts on $\widetilde{B}$. Note that $[\gamma \cdot \alpha] = [\alpha]$ only if $[\gamma]$ is the identity of $\Pi_1(B, b_0)$ so that the action is a free action. The projection $p : \widetilde{B} \to B$ commutes with this action and so induces a map $h : \widetilde{B}/\Pi_1(B, b_0) \to B$. If q denotes the quotient map $q : \widetilde{B} \to \widetilde{B}/\Pi_1(B, b_0)$, then $hq = p$. As p has these properties, this h is continuous and open, and it is easy to check that h is a bijection. Thus h is a homeomorphism, and the fact that p is a covering implies that q is a covering.

Finally, it is necessary to check that $\widetilde{B}$ is simply connected. By the injectivity of $p_\star$, it suffices to show that for any loop $\gamma : [0, 1] \to \widetilde{B}$, the loop $p\gamma$ is null-homotopic in B by a homotopy that keeps $\{0, 1\}$ fixed. For each t, $\gamma(t) = [\alpha_t]$ for some path α_t in B from b_0 to $p\gamma(t)$. By the continuity of γ and the compactness of $[0, 1]$, there is a dissection of the interval

$$0 = t_0 \le t_1 \le \ldots \le t_n = 1$$

so that $\gamma([t_i, t_{i+1}]) \subset \langle \alpha_{t_i}, V_i \rangle$ for each i, each V_i is open in B and path-connected, and the map $\Pi_1(V_i) \to \Pi_1(B)$ induced by inclusion is the constant map. Now,

$\alpha_{t_{i+1}} \approx \alpha_{t_i} \cdot \beta_i$ for some path β_i in V_i from $\alpha_{t_i}(1)$ to $\alpha_{t_{i+1}}(1)$. Because $\Pi_1(V_i) \to \Pi_1(B)$ is constant, β_i can be chosen to be any path in V_i between these end points. Thus, choose β_i to be a reparametrisation of the restriction of $p\gamma$ to the subinterval $[t_i, t_{i+1}]$. But $\alpha_{t_{i+1}} \approx \alpha_{t_i} \cdot \beta_i$ implies that $\overline{\alpha_{t_i}} \cdot \alpha_{t_{i+1}} \approx \beta_i$, and so $p\gamma \approx \beta_0 \cdot \beta_1 \cdot \dots \cdot \beta_{n-1} \approx \overline{\alpha_{t_0}} \cdot \alpha_{t_n}$, and as $[\alpha_{t_0}] = [\alpha_{t_n}]$, this is homotopic to a constant loop keeping $\{0, 1\}$ fixed. □

A further remark is in order using the notation of the above proof. Suppose that $[\gamma] \in \Pi_1(B, b_0)$. By definition of the group action, $[\gamma]\langle \alpha, V\rangle = \langle \gamma \cdot \alpha, V\rangle$. Suppose that $\Pi_1(V) \to \Pi_1(B)$ is constant. If $\langle \gamma \cdot \alpha, V\rangle = \langle \alpha, V\rangle$, then $\gamma \cdot \alpha \approx \alpha$; this means that $[\gamma]$ is the identity element of $\Pi_1(B, b_0)$. Otherwise $[\gamma]\langle \alpha, V\rangle$ and $\langle \alpha, V\rangle$ are disjoint. Thus the action of $\Pi_1(B, b_0)$ (or of any of its subgroups) on the universal cover $\widetilde{B}$ of B has the following property: Each point of $\widetilde{B}$ has an open neighbourhood that is disjoint from every one of its translates by a non-trivial element of the group. This property will now be explored.

Theorem 7.14. *Suppose that a group G acts as a group of homeomorphisms on a path-connected, locally path-connected, space Y. Suppose that each y belonging to Y has an open neighbourhood U such that $U \cap gU = \emptyset$ for all $g \in G - \{1\}$. Then the quotient map $q : Y \to Y/G$ is a covering map. If Y is simply connected, then $\Pi_1(Y/G)$ is isomorphic to G.*

PROOF. If $y \in Y$, there is an open neighbourhood U of y such that $U \cap gU = \emptyset$ for all $g \in G - \{1\}$. Now $q^{-1}(qU) = \bigcup_{g \in G} gU$. This is open because each gU is open (because g is a homeomorphism). Hence qU is open in the quotient topology on Y/G. Similarly, if U' is any open subset of U, then qU' is open. The map $q : U \to qU$ is an injection because $U \cap gU = \emptyset$ for all $g \neq 1$, and so it is a homeomophism. Of course, $qg^{-1} = q$, so that $q : gU \to qU$ is also a homeomorphism. Thus q is a covering map.

Suppose now that Y is simply connected. Let y_0 be a base point in Y and let g belong to G. Define a function $\phi : G \to \Pi_1(Y/G, q(y_0))$ as follows: Let α be a path in Y from y_0 to gy_0 and let $\phi(g) = [q\alpha]$. If β is another such path, $\alpha \approx \beta$ as Y is simply connected. So $[q\alpha] = [q\beta]$, and ϕ is well defined. Let α_1 be a path from y_0 to $g_1 y_0$ and α_2 be a path from y_0 to $g_2 y_0$. Then $\alpha_1 \cdot g_1\alpha_2$ is a path from y_0 to $g_1 g_2 y_0$. Thus $\phi(g_1 g_2) = [q(\alpha_1 \cdot g_1\alpha_2)] = [q(\alpha_1) \cdot q(\alpha_2)] = \phi(g_1)\phi(g_2)$, and so ϕ is a group homomorphism. The path lifting property of a covering (Lemma 7.4) implies at once that ϕ is surjective, and the homotopy lifting property (Lemma 7.5) implies it is injective. □

This theorem can sometimes be used in an elementary way to determine the fundamental group of a space if that space can easily be expressed as Y/G, where G acts on a simply connected space Y as in the theorem. Thus, referring back to Examples 7.2, it is clear that

$$\Pi_1(S^1) \cong \mathbb{Z}, \quad \Pi_1(\mathbb{R}P^n) \cong \mathbb{Z}/2\mathbb{Z}, \quad \text{and} \quad \Pi_1(L_{p,q}) \cong \mathbb{Z}/p\mathbb{Z}.$$

A good exercise is to construct the famous Klein bottle as the quotient of the plane with respect to a group action, and then to use the theorem to determine, as a subgroup of the isometries of the plane, the (non-abelian) fundamental group of the Klein bottle. This theorem must be accompanied by the usual caution that when considering quotient spaces, it is possible that Y may be Hausdorff but that Y/G may fail to be Hausdorff.

Suppose that a group G acts on a space Y as in the last theorem and that H is a subgroup of G. Then there is a commutative diagram of maps

$$\begin{array}{ccc} Y & \xrightarrow{q_H} & Y/H \\ \downarrow 1 & & \downarrow p \\ Y & \xrightarrow{q_G} & Y/G \end{array}$$

where q_H and q_G are the two quotient maps and p is the map that makes the diagram commute (it exists as $H \subset G$). Of course, as the action of G satisfies the condition in the theorem, so does the action of H, so that q_H is a covering map. If U is an open neighbourhood of $y \in Y$ such that $U \cap gU = \emptyset$ for all $g \in G - \{1\}$, then $p^{-1}(q_G U) = GU/H = \bigcup_{g \in G} q_H(HgU)$. For any right coset Hg of H, the set $q_H(HgU)$ is open in Y/H, it projects by p homeomorphically onto $q_G U$, and distinct cosets give distinct open sets in Y/H. Thus $p : Y/H \to Y/G$ is a covering map. By the previous theorem, if Y is simply connected, then $H \cong \Pi_1(Y/H)$, and the inclusion $H \subset G$ corresponds to the injection $p_\star : \Pi_1(Y/H) \to \Pi_1(Y/G)$.

It might be pleasing if the group action of G on Y were to induce a group action on Y/H with Y/G as the resulting quotient. For that to happen, one requires that $q_H(y_1) = q_H(y_2)$ imply that $q_H(gy_1) = q_H(gy_2)$ for all $g \in G$. Trivially, $y_1 = hy_2$ if and only if $gy_1 = ghg^{-1}gy_2$ for all $g \in G$. Thus the requirement is that H should be a normal subgroup of G. Then the quotient group G/H does act on Y/H with quotient space Y/G. These remarks and the previous two theorems (starting with subgroups of $\Pi_1(B)$ acting on the universal cover $\widetilde{B}$) produce the following theorem:

Theorem 7.15. *Let B be a path-connected, locally path-connected, semi-locally simply connected space. Then for any subgroup G of $\Pi_1(B)$, there exists a covering map $p : E_G \longrightarrow B$, unique up to equivalence, such that $p_\star : \Pi_1(E_G) = G$. If H is a subgroup of G, then E_H covers E_G and the covering maps compose in a natural way. If G is a normal subgroup of $\Pi_1(B)$, then $\Pi_1(B)/G$ acts freely on E_G and the quotient map is equivalent to p.*

When G is a normal subgroup of $\Pi_1(B)$, the covering $E_G \to B$ is called a regular covering. Note that the universal cover $\widetilde{B}$ is so called because it covers any other cover of B. From Theorem 7.15, it is clear that understanding all covering spaces of B is, in some sense, equivalent to understanding all subgroups of $\Pi_1(B)$. That will not always be easy. However, a little practice can be obtained from consideration of covering spaces of the space $S^1 \vee S^1$, two circles with one point [illegible] shown in Figure 7.2(a). The space, of course, covers itself, and the

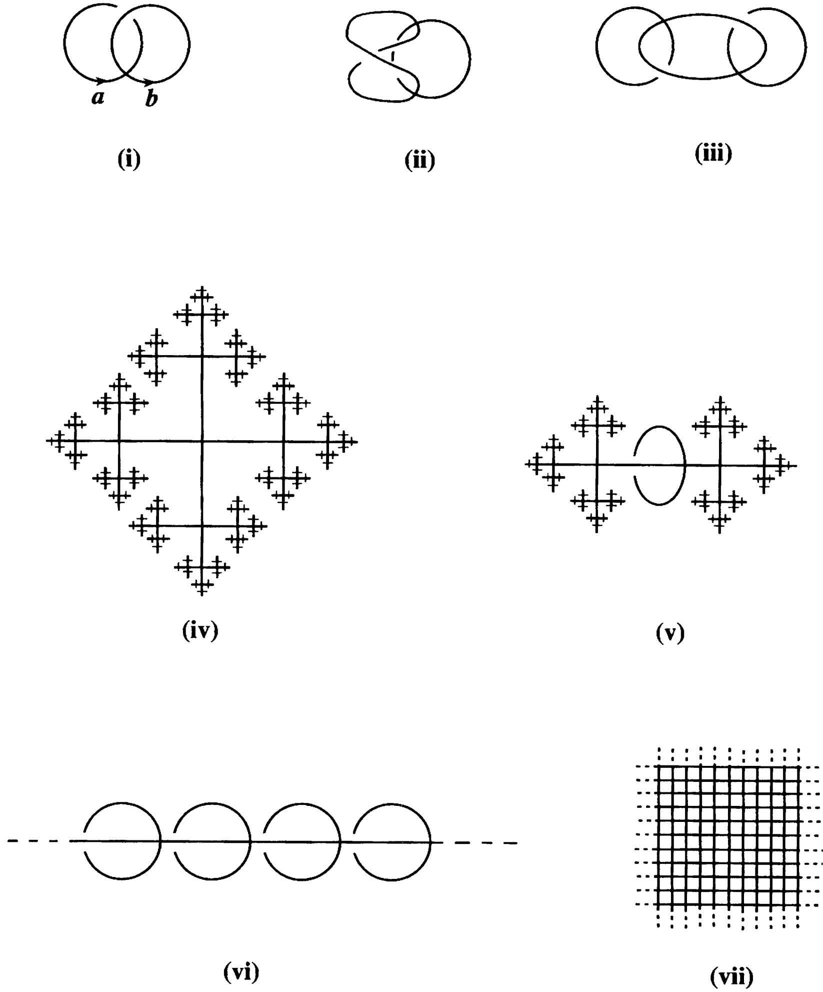


Figure 7.2

group of the cover is the whole of $\Pi_1(S^1 \vee S^1)$, which is the free group on two generators a and b. The next six parts of Figure 7.2 show other covering spaces of $S^1 \vee S^1$ where the covering maps are defined in a sensible and obvious way. As an exercise, determine in terms of a and b the groups of these various covers, and determine whether each covering is regular.

A final useful example concerns manifolds. Suppose B is now a connected non-orientable manifold. The subgroup of $\Pi_1(B)$ consisting of all elements represented by loops that preserve orientation is a normal subgroup of index 2. The covering space corresponding to this, for which the covering map is a two-to-one map, is an orientable manifold called the *orientable double cover of* B

Exercises

1. Write out the details of a proof of the exactness of the homotopy sequence (7.6) associated with a covering map.
2. Figure 7.1 illustrates a method for finding the Alexander polynomial of the knot 4_1. Use this method to find the Alexander polynomial of 6_3.
3. Let B be a θ-curve, that is, a graph of two vertices and three edges, each edge having the two vertices as its end points. Describe (i) the universal cover of B, (ii) a cover of B with infinite cyclic fundamental group and (iii) a finite cover of B.
4. Work through the exercises suggested in association with Figure 7.2.
5. What is the orientable double cover of (i) the Möbius band, (ii) the real projective plane, (iii) the Klein bottle and (iv) the connected sum of n real projective planes?
6. The group $\mathbb{Z} \oplus \mathbb{Z}$ acts on $\mathbb{R}^2$ by $(m, n)(x, y) = (x + m, y + n)$. The quotient space is a torus, and the quotient map $q : \mathbb{R}^2 \to \mathbb{R}^2/\mathbb{Z} \oplus \mathbb{Z}$ is the universal covering map of the torus. By considering the projection of the straight line in $\mathbb{R}^2$ from the origin to the point (p, q), show that if p and q are coprime, then the element $(p, q) \in \mathbb{Z} \oplus \mathbb{Z} \equiv H_1(S^1 \times S^1)$ is represented by a simple closed curve. By cutting the torus along any given non-separating simple closed curve and noting that the result is a annulus, prove the converse is also true.
7. Find a specific action of a group G on the plane $\mathbb{R}^2$ so that the quotient space $\mathbb{R}^2/G$ is a Klein bottle and the action satisfies the condition of Theorem 7.14. Prove that the fundamental group of the Klein bottle is non-abelian.
8. A diagram of a (null-homotopic) simple closed curve C in a solid torus T is shown in Figure 6.5. By considering linking numbers between different lifts of C to the universal cover of T, show that C is not the boundary of any disc embedded in T. Let $\overline{C}$ be the curve in T represented by this diagram reflected in the plane of the paper (that is, with the two crossings changed). Show, again by considering lifts to the universal cover, that there is no orientation preserving (piecewise linear) homeomorphism of T to itself sending C to $\overline{C}$.
9. Suppose that $p : \widetilde{B} \to B$ is a universal covering map and $X \subset \widetilde{B}$. Show that $p|X$ is an injection if and only if $gX \cap X = \emptyset$ for all $g \in \Pi_1(B)$ with g not equal to the identity. Suppose that $\mathbb{R}^3$ is the universal covering space of a closed connected 3-manifold M. Show that any 2-sphere piecewise linearly embedded in M separates M into two components, the closure of one of which is a 3-ball.
10. The fundamental group of a graph (a possibly infinite 1-dimensional complex) is a *free* group. Prove that any subgroup of a free group is a free group.
11. Prove that if knots K_1 and K_2 are related by mutation, then the double cover of S^3 branched over K_1 and the double cover of S^3 branched over K_2 are homeomorphic.

8

The Conway Polynomial, Signatures and Slice Knots

The Conway polynomial [20] for an oriented link is really just the Alexander polynomial without the ambiguity concerning multiplication by units of $\mathbb{Z}[t^{-1}, t]$. Although that might seem a small improvement, it enables two such polynomials to be added together, which would be meaningless if the signs were in doubt, and this in turn permits a "skein formula" for the Alexander polynomials of links to be produced. The method for this given below uses Seifert matrices as before, but it abandons any interpretation by means of the homology of the infinite cyclic cover. (Use of the L-matrix of Reidemeister, as in [107] or [108], can also produce this theory.)

Definition 8.1. Suppose that F is a Seifert surface for an oriented link L in S^3. Suppose there is a solid cylinder, parametrised as $[0, 1] \times D^2$, in S^3 such that $([0, 1] \times D^2) \cap F = \{0, 1\} \times D^2$, the solid cylinder being on the same side of F near $\{0, 1\} \times D^2$. Let $F' = (F - \{0, 1\} \times D^2) \cup [0, 1] \times \partial D^2$. Then F' is said to be obtained from F by means of (embedded) surgery along the arc $[0, 1] \times 0$.

Theorem 8.2. *Suppose that F_1 and F_2 are Seifert surfaces for an oriented link L in S^3. Then there is a sequence of Seifert surfaces $\Sigma_1, \Sigma_2, \dots \Sigma_N$, with $\Sigma_1 = F_1$ and $\Sigma_N = F_2$, such that for each i, either Σ_i is obtained from Σ_{i-1} or Σ_{i-1} is obtained from Σ_i by surgery along an arc embedded in S^3, or they are related by an isotopy of S^3.*

PROOF. After a small homeomorphism (which is isotopic to the identity) of S^3, it may be assumed that F_1 and F_2 intersect transversely in finitely many simple closed curves, including their common boundary L. Suppose that the closure M of some component of $S^3 - (F_1 \cup F_2)$ is a 3-manifold with the property that wherever M abuts either F_i, it always does so from the same side of F_i. Let $\partial M = \partial_1 M \cup \partial_2 M$, where $\partial_i M = \partial M \cap F_i$. Any triangulation of S^3 with F_1 and F_2 as subcomplexes includes a triangulation T of M. Let A be a (collar) neighbourhood of $\partial_1 M$ in M together with a neighbourhood in M of all the 1-simplexes of T. Let B be the closure of $M - A$. To be more precise, A is the simplicial neighbourhood, in the second derived subdivision $T^{(2)}$ of the union of

$\partial_1 M$ with the 1-simplexes of T. Then B is the simplicial neighbourhood in $T^{(2)}$ of the union of all cones with vertex the barycentre of some 3-simplex σ of T and base the barycentres of the 2-simplexes in $\partial\sigma - \partial_1 M$. Change F_1 to F_1' by removing $\partial_1 M$ and inserting the closure of $\partial A - \partial_1 M$. Change F_2 to F_2' similarly by removing $B \cap F_2$ and inserting the closure of $\partial B - (B \cap F_2)$. These changes can be achieved by moving F_1 by an isotopy across the collar and then across the the neighbourhood of the graph of 1-simplexes by isotopies and by surgeries along embedded arcs. Similarly (only without the collar), F_2 can be changed to F_2'. Now, $F_1' \cap F_2' = F_1 \cap F_2 \cup \overline{(\partial A - \partial_1 M)}$, and a small displacement of F_1' removes $\overline{\partial A - \partial_1 M}$ from this intersection and so reduces the number of components of $F_1' \cap F_2'$ to less than the number of components of $F_1 \cap F_2$. If $L \subset M$, readjust F_1' by an isotopy that slides $\partial F_1'$ back down the collar until $\partial F_1' = \partial F_2' = L$.

In this inductive way the number of components of intersection of the two Seifert surfaces can be steadily reduced until $F_1 \cap F_2 = L$. Then one more application of the above procedure finishes the proof. However, it is important to show that at any stage of this induction, the manifold M that abuts each F_i from one side only does indeed exist. How to find M? Recall the infinite cyclic cover X_∞ of the exterior X of L that is constructed by gluing together infinitely many copies of Y, where Y is X-cut-along-F_1. As proved in Theorem 7.9, this is the same as the cover constructed in a similar way by cutting along F_2. Thus X_∞ contains infinitely many copies of F_2 (less a small neighbourhood of L) which are the lifts to the cover of the second Seifert surface. The infinite cyclic group $\langle t \rangle$ acts on X_∞; the homeomorphism t moves one lift of F_i to the next lift. Let $\widehat{F}_1 \subset X_\infty$ be a fixed lift of F_1 and let $\widehat{F}_2 \subset X_\infty$ be a fixed lift of F_2. Suppose that $(F_1 \cap F_2) - L \neq \emptyset$. Let n be the maximal integer such that $\widehat{F}_2 \cap t^n \widehat{F}_1 \neq \emptyset$. The surface $\widehat{F}_2$ separates X_∞ into two components C_L and C_R, with $t^r \widehat{F}_2 \subset C_L$ if and only if $r < 0$. Let Y_n be the copy of Y between $t^n \widehat{F}_1$ and $t^{n+1} \widehat{F}_1$, and let $\widehat{M}$ be the closure of some component of $C_L \cap Y_n$. The boundary of $\widehat{M}$ is contained in $t^n \widehat{F}_1 \cup \widehat{F}_2 \cup \partial X_\infty$, and clearly $\widehat{M}$ lies on only one side of $t^n \widehat{F}_1$ and one side of $\widehat{F}_2$. The projection map $p : X_\infty \to X$ is injective when restricted to $\widehat{M}$, as $\widehat{M} \subset Y_n - t^{n+1} \widehat{F}_1$. Let M be $p\widehat{M}$. Now observe that M is just the closure of some component of $X - (F_1 \cup F_2)$. If this were not so, some lift $t^r \widehat{F}_2$ of F_2 would intersect $\widehat{M}$ for some $r \neq 0$. But if $r > 0$, $t^r \widehat{F}_2$ is disjoint from the closure of C_L and so disjoint from $\widehat{M}$. If $r < 0$, $t^r \widehat{F}_2 \cap t^n \widehat{F}_1 = \emptyset$ by the maximality of n. □

Definition 8.3. Let A be a square matrix over $\mathbb{Z}$. An elementary enlargement of A is a matrix B of the form

$$B = \begin{pmatrix} A & \xi & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} A & 0 & 0 \\ \eta^\tau & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

for some column ξ or row η^τ. The matrix A is called an elementary reduction of B. Square matrices A and B over $\mathbb{Z}$ are called S-equivalent if they are related by a sequence of elementary enlargements, elementary reductions and unimodular congruences (this last being a relation of the form $B = P^{\tau}AP$, where $\det P = \pm 1$).

Theorem 8.4. *Let A and B be Seifert matrices for an oriented link L. Then A and B are S-equivalent.*

PROOF. Suppose that A is an $n \times n$ matrix corresponding to a Seifert surface F, with respect to some base of $H_1(F;\mathbb{Z})$. Changing the base used for $H_1(F;\mathbb{Z})$ changes A to a matrix of the form $P^{\tau}AP$, where P is the unimodular base-change matrix. Thus it suffices to check what happens when the Seifert surface is changed, and to do that it suffices, by Theorem 8.2, to check (with respect to any base) the effect of surgery along an arc. Suppose F is changed to F' by surgery along an arc. A base for $H_1(F';\mathbb{Z})$ can be chosen to be the homology classes of curves $\{f_i\}$ that constitute a base for $H_1(F;\mathbb{Z})$ together with the classes of a curve f_{n+1} that goes once over the solid cylinder defining the surgery and of a curve f_{n+2} around the middle of the cylinder (that is, $f_{n+2} = 1/2 \times \partial D^2$ in the notation of Definition 8.1). Then, because f_{n+2} bounds a disc $(1/2 \times D^2)$ that is disjoint from $\bigcup\{f_i : i \le n\}$, $\operatorname{lk}(f_{n+2}^{\pm}, f_i) = 0$ for all $i \neq n+1$. Further, as f_{n+1} meets this disc at one point in its boundary, choosing orientations carefully gives either $\operatorname{lk}(f_{n+1}^{+}, f_{n+2}) = 0$ and $\operatorname{lk}(f_{n+1}^{-}, f_{n+2}) = 1$, or $\operatorname{lk}(f_{n+1}^{+}, f_{n+2}) = 1$ and $\operatorname{lk}(f_{n+1}^{-}, f_{n+2}) = 0$. In the first case the new Seifert matrix is of the form

$$\begin{pmatrix} A & \xi & 0 \\ ? & ? & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{which is congruent to} \quad \begin{pmatrix} A & \xi & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

The second case leads to a Seifert matrix of the form

$$\begin{pmatrix} A & 0 & 0 \\ \eta^{\tau} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

□

It follows from this theorem that any invariant well-defined on S-equivalence classes of square matrices of integers gives at once an invariant of oriented links. For example, let A be a Seifert matrix for L, and *define* $\Delta_L(t) \in \mathbb{Z}[t^{-1/2}, t^{1/2}]$ to be $\det(t^{1/2}A - t^{-1/2}A^{\tau})$. Here $t^{1/2}$ is just an indeterminate for the ring of Laurent polynomials $\mathbb{Z}[t^{-1/2}, t^{1/2}]$, but it should be thought of as a formal square root of t, so that $\mathbb{Z}[t^{-1}, t] \subset \mathbb{Z}[t^{-1/2}, t^{1/2}]$. Note that if A is an $r \times r$ matrix, then $\Delta_L(t) = t^{-r/2}\det(tA - A^{\tau})$, so that up to a unit of $\mathbb{Z}[t^{-\frac{1}{2}}, t^{\frac{1}{2}}]$, it follows that $\Delta_L(t)$ is just the Alexander polynomial of L. However, it will now be shown that this normalised $\Delta_L(t)$ has no ambiguity of sign or units. Thus call this the Conway normalisation of the Alexander polynomial.

Theorem 8.5. *The Conway-normalised Alexander polynomial is a well-defined invariant of the oriented link L.*

PROOF. It is only necessary to check the invariance of the Conway-normalised polynomial when A changes by S-equivalence. Firstly, note that

$$\det(t^{1/2}P^\tau AP - t^{-1/2}P^\tau A^\tau P) = (\det P)^2 \det(t^{1/2}A - t^{-1/2}A^\tau),$$

so that the normalised $\Delta_L(t)$ is invariant under unimodular congruence. If now

$$B = \begin{pmatrix} A & \xi & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

then

$$(t^{1/2}B - t^{-1/2}B^\tau) = \begin{pmatrix} t^{1/2}A - t^{-1/2}A^\tau & t^{1/2}\xi & 0 \\ -t^{-1/2}\xi^\tau & 0 & t^{1/2} \\ 0 & -t^{-1/2} & 0 \end{pmatrix},$$

which has the same determinant as $(t^{1/2}A - t^{-1/2}A^\tau)$. Similarly, the other type of elementary enlargement of A has no effect on this determinant. □

Note that for a *knot* K the Conway-normalised Alexander polynomial belongs to $\mathbb{Z}[t^{-1}, t]$, it is symmetric between t and t^{-1} and $\Delta_K(1) = +1$ by the proof of Theorem 6.10. The polynomials quoted in the table in Chapter 6 are indeed Conway-normalised. The sign needed for the normalisation cannot be determined in this simple way for an oriented link L of two or more components because $\Delta_L(1) = 0$.

Theorem 8.6. *For oriented links L, the Conway-normalised Alexander polynomial $\Delta_L(t) \in \mathbb{Z}[t^{-\frac{1}{2}}, t^{\frac{1}{2}}]$ is characterised by*
(i) $\Delta_{\text{unknot}}(t) = 1$,
(ii) *whenever three oriented links L_+, L_- and L_0 are the same except in the neighbourhood of a point where they are as shown in Figure 3.2, then*

$$\Delta_{L_+} - \Delta_{L_-} = (t^{-1/2} - t^{1/2})\Delta_{L_0}.$$

PROOF. Construct a Seifert surface F_0 for L_0 that meets the neighbourhood of the point in question as shown in Figure 8.1. The Seifert circuit method described in Chapter 2 will do this. Now form Seifert surfaces F_+ for L_+ and F_- for L_- by adding short twisted strips to F_0 as also shown in Figure 8.1. Let $H_1(F_0; \mathbb{Z})$ be

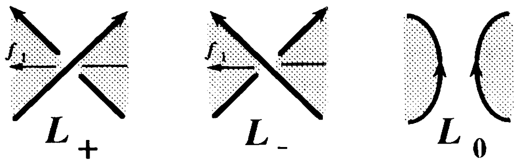


Figure 8.1

generated by the classes of oriented closed curves $\{f_2, f_3, \dots, f_n\}$, and for generators of $H_1(F_\pm; \mathbb{Z})$, take the classes of the same curves together with the class of an extra curve f_1 that goes once along the twisted strip. If A_0 is the resulting Seifert matrix for L_0, the Seifert matrix for L_- is of the form $\begin{pmatrix} N & \xi^\tau \\ \eta & A_0 \end{pmatrix}$ for some integer N and columns ξ and η, whereas that for L_+ is $\begin{pmatrix} N-1 & \xi^\tau \\ \eta & A_0 \end{pmatrix}$. Consideration of $\det(t^{1/2}A - t^{-1/2}A^\tau)$ when A is each of these three Seifert matrices immediately produces the required formula. □

The formulae of this theorem is the promised analogue of the similar formulae (Proposition 3.7) for the Jones polynomial. Just as for the Jones polynomial, repeated use of these formulae allow $\Delta_L(t)$ to be calculated for any oriented link L. It is easy to see that the result is always a polynomial (not a Laurent polynomial) in $(t^{-1/2} - t^{1/2})$, so make the substitution $(t^{-1/2} - t^{1/2}) = z$, and define the Conway polynomial, or potential, for L to be $\nabla_L(z) \in \mathbb{Z}[z]$, where $\nabla_L(t^{-1/2} - t^{1/2}) = \Delta_L(t)$, the Conway-normalised Alexander polynomial.

A paraphrase of the last theorem is that, using the theory of S-equivalence of Seifert matrices, the Conway polynomial invariant of oriented links is well defined. It is characterised by $\nabla_{\text{unknot}}(z) = 1$ and (with reference to Figure 3.2) the skein formula

$$\nabla_{L_+}(z) - \nabla_{L_-}(z) = z\nabla_{L_0}(z).$$

In theory at least, this suffices for calculation of $\nabla_L(z)$. Some easily deduced properties follow; $\nabla_L(z)$ is written as $\nabla_L(z) = \sum_{i\geq 0} a_i(L)z^i$, where $a_i(L) \in \mathbb{Z}$.

Proposition 8.7. *For an oriented link L with $\#L$ components, the Conway polynomial has the following properties.*

(i) *If L is a split link, then $\nabla_L(z) = 0$.*

(ii) *$a_i(L) = 0$ for $i \equiv \#L$ modulo 2 and also for $i < \#L - 1$.*

(iii) *If L is a knot, so $\#L = 1$, then $a_0(L) = 1$.*

(iv) *If $\#L = 2$, then $a_1(L) = \text{lk}(L)$, where $\text{lk}(L)$ is the linking number of the two components of L.*

(v) *If L_+, L_- and L_0 are related in the manner of Figure 3.2 and $\#L_+ = \#L_- = 1$, then $a_2(L_+) - a_2(L_-) = \text{lk}(L_0)$.*

PROOF. (i) This follows from the stronger Proposition 6.14. However, it also follows at once by applying the skein formula to links L_+, L_- and L_0 shown in Figure 8.2. As L_+ and L_- are here the same link, $\nabla_{L_0}(z) = 0$.

(ii) This follows by induction on the number of crossings in a diagram and the number of crossing changes needed to make it a diagram of the trivial link of unknots.

(iii) When $z = 0$, the skein formula becomes $\nabla_{L_+}(0) = \nabla_{L_-}(0)$, so any crossings can be changed without altering $\nabla_L(0)$. Of course, $a_0(\text{unknot}) = 1$.

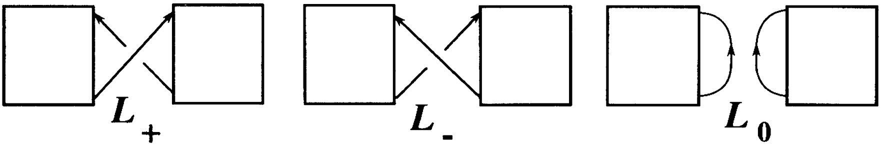


Figure 8.2

(iv) Suppose the skein formula is considering a crossing between the two components of L. Using (iii), consideration of the coefficient of z shows that $a_1(L_+) - a_1(L_-) = 1$. But $\text{lk}(L_+) - \text{lk}(L_-) = 1$, so the result follows by using (i) and considering a collection of crossing changes that yield a split link.

(v) This follows at once from (iv). □

A good exercise is to use the skein formula to show that the Conway polynomial is equal to 1 for the generalised Kinoshita–Terasaka knot shown in Figure 8.3. The symbols in Figure 8.3 denote the numbers of crossings in the tassels, and d is required to be even.

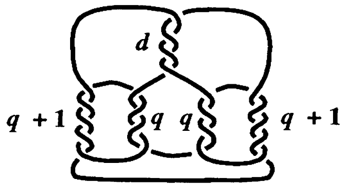


Figure 8.3

Although this completes the theory that establishes the Conway polynomial, it is convenient here to establish also the existence of the ω-signature of an oriented link. This will be done in a direct matrix-oriented manner. This ω-signature was first introduced by A. G. Tristram [123], generalising work of Murasugi [102].

Definition 8.8. Let L be an oriented link in S^3 and let ω be a unit modulus complex number, $\omega \neq 1$. The ω-signature $\sigma_\omega(L)$ of L is defined to be the signature of the Hermitian matrix

$$(1 - \omega)A + (1 - \overline{\omega})A^\tau,$$

where A is a Seifert matrix for L .

Theorem 8.9. *The ω-signature $\sigma_\omega(L)$ is well defined as an invariant of L.*

Proof. The signature of a Hermitian matrix is not changed by congruence (that fact is Sylvester's famous law of inertia), so it is only necessary to see whether the definition changes under an elementary enlargement of a Seifert matrix A.

Suppose

$$B = \begin{pmatrix} A & \xi & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix};$$

then

$$(1-\omega)B + (1-\overline{\omega})B^{\tau} = \begin{pmatrix} (1-\omega)A + (1-\overline{\omega})A^{\tau} & (1-\omega)\xi & 0 \\ (1-\overline{\omega})\xi^{\tau} & 0 & (1-\omega) \\ 0 & (1-\overline{\omega}) & 0 \end{pmatrix}.$$

As $(1-\omega) \neq 0$, the terms in ξ and ξ^{τ} can be removed by congruence (subtracting multiples of the last row and column from predecessors), so that the signature of $(1-\omega)A + (1-\overline{\omega})A^{\tau}$ and the signature of $(1-\omega)B + (1-\overline{\omega})B^{\tau}$ differ by the signature of $\begin{pmatrix} 0 & (1-\omega) \\ (1-\overline{\omega}) & 0 \end{pmatrix}$. Of course, this last signature is zero, as the matrix clearly has one positive eigenvalue and one negative one. Consideration of the other type of elementary enlargement is exactly the same. □

Note that $(1-\omega)A + (1-\overline{\omega})A^{\tau} = -(1-\overline{\omega})(\omega A - A^{\tau})$, so that the Hermitian matrix is non-singular except when ω is a zero of the Alexander polynomial of L. In fact, it can be shown that for a fixed link L, the invariant $\sigma_{\omega}(L)$, when viewed as a function of ω, is continuous except at zeros of the Alexander polynomial. As signatures are integers, this means that $\sigma_{\omega}(L)$ takes finitely many values as ω varies on S^1, with possible jumps at roots of $\Delta_L(t) = 0$.

Sometimes $\sigma_{-1}(L)$ is called *the* signature of L. Table 8.1 records the value of this signature for the knots up to eight crossings as depicted in Chapter 1.

Theorem 8.10. *If L is an oriented link in S^3 and $\overline{L}$ is its reflection, then for any unit complex number $\omega \neq 1$,*

$$\sigma_{\omega}(L) = -\sigma_{\omega}(\overline{L}).$$

PROOF. If A is a Seifert matrix for L, then $-A$ is a Seifert matrix for $\overline{L}$. □

TABLE 8.1. Signatures of Knots

$\mathbf{3_1}$	2	$\mathbf{7_1}$	6	$\mathbf{8_1}$	0	$\mathbf{8_8}$	0	$\mathbf{8_{15}}$	4
$\mathbf{4_1}$	0	$\mathbf{7_2}$	2	$\mathbf{8_2}$	4	$\mathbf{8_9}$	0	$\mathbf{8_{16}}$	2
$\mathbf{5_1}$	4	$\mathbf{7_3}$	−4	$\mathbf{8_3}$	0	$\mathbf{8_{10}}$	−2	$\mathbf{8_{17}}$	0
$\mathbf{5_2}$	2	$\mathbf{7_4}$	−2	$\mathbf{8_4}$	2	$\mathbf{8_{11}}$	2	$\mathbf{8_{18}}$	0
$\mathbf{6_1}$	0	$\mathbf{7_5}$	4	$\mathbf{8_5}$	−4	$\mathbf{8_{12}}$	0	$\mathbf{8_{19}}$	−6
$\mathbf{6_2}$	2	$\mathbf{7_6}$	2	$\mathbf{8_6}$	2	$\mathbf{8_{13}}$	0	$\mathbf{8_{20}}$	0
$\mathbf{6_3}$	0	$\mathbf{7_7}$	0	$\mathbf{8_7}$	2	$\mathbf{8_{14}}$	2	$\mathbf{8_{21}}$	2

A corollary is, of course, that if $L = \overline{L}$ then $\sigma_\omega(L) = 0$. A direct calculation shows that the signature of the trefoil knot is 2 (or -2 if the reflected diagram is used). That is a pre–Jones polynomial proof of the fact that the trefoil and its reflection are distinct knots.

The remainder of this chapter takes a brief look at 4-dimensional topology. The Alexander polynomial and the signatures of knots give information concerning whether a knot K in S^3 bounds some disc embedded in B^4, the 4-ball bounded by S^3.

Definition 8.11. A knot $K \subset S^3$ is a slice knot if there is a flat disc D contained in B^4 such that $K = \partial D = D \cap S^3$. Such a disc is called a slicing disc for K.

Here "flat" means that D has a neighbourhood N that is a copy of $D \times I^2$ meeting S^3 in $\partial D \times I^2$ (of course, $I^2 = I \times I$, and this is just another disc). To avoid triviality such a restriction is needed, for B^4 can be regarded as the cone on S^3, and this contains the cone on *any* knot in S^3. Such a subcone is not flat unless the knot is trivial. It is known that a locally flat condition for D implies flatness. Similarly, if everything is interpreted in terms of differential topology and the disc D is a smooth submanifold of B^4, then it has a trivial normal bundle and so is flat.

Slice knots seem to be fairly rare. In Table 1.1, the knots 6_1, 8_8, 8_9 and 8_{20} are slice. The sum of any knot K with the reverse of its reflection is also slice. This can be seen by creating (B^4, D) from $(S^3 \times I, K \times I)$ by removing a neighbourhood of $\{x\} \times I$, where $x \in K$. An explicit example of a slice knot is needed. Consider, as an analogue, the contour-map description of a mountain with two peaks and one pass (or col) somewhere between the peaks. At low levels there is just one simple closed curve as the contour line. This becomes a curve with one self-intersection point at the level of the pass. Above that, the contour consists of two simple closed curves, which finally become single points at the peaks. Figure 8.4 shows a disc evolving in $S^3 \times [0, \infty)$. The disc meets low levels in a copy of the knot 8_{20}, then meets a critical level in a curve with one self-intersection and meets levels just above that in two curves. The important thing is that these two curves are *unknotted* and *unlinked*, and hence they can be capped off with two discs in a standard way. As will be shown below, any attempt to imitate this with the knot 3_1 will fail. It is known that any slicing disc is obtained in this way, though it may have minima (levels where a curve is "born" unknotted and unlinked from everything else) as well as many passes and maxima. The "slice knots are ribbon knots"

Figure 8.4

conjecture opines that minima are unnecessary. To obtain necessary conditions for sliceness from the theory of the Seifert form, some preliminary lemmas are needed.

Lemma 8.12. *Suppose that for some knot K in S^3, there is a flat surface F in B^4 with $F \cap S^3 = \partial F \cap S^3 = K$. Then the inclusion map induces an isomorphism $H_1(S^3 - K) \longrightarrow H_1(B^4 - F) \cong \mathbb{Z}$.*

PROOF. Let N, a copy of $F \times I^2$, be a neighbourhood of F meeting S^3 in $\partial F \times I^2$. The Mayer-Vietoris theorem gives an exact sequence

$$0 = H_2(B^4) \longrightarrow H_1(F \times \partial I^2) \longrightarrow H_1(N) \oplus H_1(\overline{B^4 - N}) \longrightarrow H_1(B^4) = 0.$$

Exactness implies that the middle map of this must be an isomorphism. Of course,

$$H_1(F \times \partial I^2) = H_1(F) \oplus H_1(\partial I^2),$$

and the $H_1(F)$ component is mapped isomorphically to $H_1(N)$ (and each is the direct sum of copies of $\mathbb{Z}$); $H_1(\partial I^2)$ is mapped to zero in $H_1(N)$. As $H_1(\partial I^2) = \mathbb{Z}$, it follows that $H_1(B^4 - N)$ is also a copy of $\mathbb{Z}$. The map $H_1(\partial I^2) \to H_1(B^4 - N)$ must send generator to generator, as otherwise a matrix representing the map in the above sequence will not have unit determinant. However, a generator of this copy of $H_1(\partial I^2)$ is a meridian of the knot K. Thus the inclusion map from the knot exterior to $B^4 - N$ induces an isomorphism on the first homology, and that is, up to adjustment by a small homotopy equivalence, the required statement. □

Lemma 8.13. *Suppose that $f_1 : F_1 \to B^4$ and $f_2 : F_2 \to B^4$ are maps, of orientable surfaces into the 4-ball, which have disjoint images. Suppose that on ∂F_i the map f_i is a homeomorphism onto a knot K_i in $S^3 = \partial B^4$. Then $\text{lk}(K_1, K_2) = 0$.*

PROOF. After moving the maps into general position, it may be assumed that each f_i has only double points as singularities. That means that near the image of such a singularity in B^4, the image of F_i looks like two standard planes in $\mathbb{R}^4$ meeting in a point P. That is, near P it is the cone from P on a standard Hopf link (a non-trivial two-crossing link) in a copy of S^3. Replace the cone on that link with a Seifert surface of the link. This changes F_i by removing two discs and inserting an annulus, but there is no longer a point of self-intersection. There may also be points at which the image of f_i is locally knotted, points P near which the image is the cone on a knot in a copy of S^3; replace that cone with a Seifert surface of the knot, changing F_i but gaining flatness. In this way it may be assumed that each f_i is an embedding onto a flat surface. Then the existence of $f_1(F_1)$ asserts that K_1 represents the zero homology class in $H_1(B^4 - F_2)$, and so, by the last lemma, K_1 represents zero in $H_1(S^3 - K_2)$. □

Lemma 8.14. *Suppose that F is a Seifert surface for a knot K in S^3 that has a slicing disc D. Then $F \cup D$ bounds some two-sided 3-manifold $M^3 \subset B^4$ with $M^3 \cap S^3 = F$.*

PROOF. The idea here is that M^3 should be ϕ^{-1} (one point), where $\phi : B^4 - D \to S^1$ is a carefully chosen map inducing an isomorphism of first homology groups. It will be more convenient to define ϕ on $B^4 - N$, where N is a standard neighbourhood of D as considered above, with $(B^4 - N) \cap S^3$ being a copy of X, the knot exterior. Define, in the following way, $\phi : X \to S^1$ so that $\phi_\star : H_1(X) \to H_1(S^1)$ is an isomorphism and ϕ^{-1} (one point) $= F$. On a product neighbourhood of F in X, define ϕ to be the projection $F \times [-1, 1] \to [-1, 1]$ followed by the map $t \mapsto e^{i\pi t} \in S^1$, and let ϕ map the remainder of X to $-1 \in S^1$. Extend ϕ over the rest of $\partial(B^4 - N)$ so that the inverse image of $1 \in S^1$ is $F \cup (D \times \star)$ for some point $\star \in \partial I^2$, where $N = D \times I^2$ (note that $\partial D \times \star$ is a longitude of K by Lemma 8.13). This map must now be extended over the whole of $B^4 - N$.

Consider the simplexes of some triangulation of $B^4 - N$. Let T be a tree in the 1-skeleton containing all the vertices of this triangulation that contains a similar maximal tree of $\partial(B^4 - N)$. Extend ϕ over all of T in an arbitrary way. Then on a 1-simplex σ not in T define ϕ so that if c is a 1-cycle consisting of σ summed with a 1-chain in T (joining up the ends of σ), $[\phi c] \in H_1(S^1)$ is the image of $[c]$ under the isomorphism

$$H_1(\overline{B^4 - N}) \overset{\cong}{\longleftarrow} H_1(X) \overset{\phi_\star}{\longrightarrow} H_1(S^1).$$

Trivially, the boundary of a 2-simplex τ of $B^4 - N$ represents zero in $H_1(B^4 - N)$, so $[\phi(\partial\tau)] = 0 \in H_1(S^1)$. Hence ϕ is null-homotopic on $\partial\tau$ and so extends over τ. Finally, ϕ extends over the 3-simplexes and 4-simplexes, as any map from the boundary of an n-simplex to S^1 is null-homotopic when $n \geq 3$.

Now, regard $\phi : B^4 - N \to S^1$ as a simplicial map to some triangulation of S^1 in which 1 is *not* a vertex. Then $\phi^{-1}(1)$ is a 3-manifold M^3, with a neighbourhood $M^3 \times I$, in $B^4 - N$. To see this just consider how ϕ^{-1}(a non-vertex) meets the neighbourhood of any simplex in $B^4 - N$. Of course, ϕ was constructed so that $\partial M^3 = F \cup (D \times \star)$. □

The method used to extend ϕ in this last proof is a very elementary example of the use of "obstruction theory". The proof can be interpreted by saying that $H^1(B^4 - N; \mathbb{Z})$ corresponds naturally to the homotopy classes of maps from $B^4 - N$ to S^1 and ϕ corresponds to a generator of $H^1(B^4 - N; \mathbb{Z})$. If working with smooth manifolds, the final manoeuvre of the proof should be replaced by the the procedure of changing ϕ by a homotopy to make it *transverse* to $1 \in S^1$ and then considering $\phi^{-1}(1)$ as before.

One more lemma is now needed. It concerns the way in which the homology of the boundary of a 3-manifold is related to that of the manifold itself. There seems to be no escape from cohomology theory here, and the proof given below is perhaps a little terse.

Lemma 8.15. *Let M be a compact orientable 3-manifold such that ∂M is a connected surface of genus g. Suppose that $i : \partial M \to M$ is the inclusion map. Then the kernel of $i_\star : H_1(\partial M; \mathbb{Q}) \to H_1(M; \mathbb{Q})$ is a vector subspace of dimension g.*

PROOF. The following commutative diagram has rows that are parts of the homology and cohomology exact sequences of the pair $(M, \partial M)$. Of the vertical arrows, the first and third are Lefschetz duality isomorphisms, and the central one is a Poincaré duality isomorphism.

$$\begin{array}{ccccc} H_2(M, \partial M; \mathbb{Q}) & \stackrel{d}{\longrightarrow} & H_1(\partial M; \mathbb{Q}) & \stackrel{i_\star}{\longrightarrow} & H_1(M; \mathbb{Q}) \\ \downarrow & & \downarrow & & \downarrow \\ H^1(M; \mathbb{Q}) & \stackrel{i^\star}{\longrightarrow} & H^1(\partial M; \mathbb{Q}) & \stackrel{\partial}{\longrightarrow} & H^2(M, \partial M; \mathbb{Q}) \end{array}$$

Now, $H^1(\partial M; \mathbb{Q})$ is the vector space dual to $H_1(\partial M; \mathbb{Q})$, $H^1(M; \mathbb{Q})$ is the space dual to $H_1(M; \mathbb{Q})$ and $i^\star$ and $i_\star$ are dual linear maps. (This follows from the universal coefficient theorem for homology and cohomology and the fact that there is no torsion when coefficients are in the field $\mathbb{Q}$.) Thus, if $r(\ )$ denotes the rank of a linear map, $r(i^\star) = r(i_\star)$. The vertical isomorphisms imply that $i_\star$ and ∂ have the same nullity. Thus $r(i^\star) = 2g - r(i_\star)$. Hence $r(i_\star) = g$, and so g is also the nullity of $i_\star$. □

Corollary 8.16. *There is a base $[f_1], [f_2], \ldots, [f_{2g}]$ over $\mathbb{Z}$ for $H_1(\partial M; \mathbb{Z})$ so that $[f_1], [f_2], \ldots, [f_g]$ map to zero in $H_1(M; \mathbb{Q})$.*

PROOF. One may consider $H_1(\partial M; \mathbb{Z})$ to be $\mathbb{Z}^{2g} \subset \mathbb{Q}^{2g} = H_1(\partial M; \mathbb{Q})$. The g-dimensional subspace U of $\mathbb{Q}^{2g}$, given by Lemma 8.15, has a base consisting of elements in $\mathbb{Z}^{2g}$. Let $\widetilde{U}$ be the $\mathbb{Z}$-span of those elements. As a $\mathbb{Z}$-module $\mathbb{Z}^{2g}/\widetilde{U} = A/\widetilde{U} \oplus B/\widetilde{U}$, where A and B are submodules of $\mathbb{Z}^{2g}$, $A/\widetilde{U}$ is free and $B/\widetilde{U}$ is a torsion module over $\mathbb{Z}$. Thus if $b \in B$ then $nb \in \widetilde{U}$ for some $n \in \mathbb{Z}$; hence $b \in U$. Thus a $\mathbb{Z}$-base for B is a $\mathbb{Q}$-base for U and it extends, using a base of $A/\widetilde{U}$, to a $\mathbb{Z}$-base of $\mathbb{Z}^{2g}$. □

Proposition 8.17. *Suppose that F is a genus g Seifert surface for a slice knot K in S^3. Then a base may be chosen for $H_1(F; \mathbb{Z})$ with respect to which the corresponding Seifert matrix has the form*

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

consisting of a $g \times g$ block of zeros together with $g \times g$ blocks of integers P, Q and R.

PROOF. Let D be a slicing disc for K contained in B^4. By Lemma 8.14 there is contained in B^4 a 3-manifold M having an $M \times [-1, 1]$ neighbourhood such that $\partial M = D \cup F$. Corollary 8.16 gives a certain base $[f_1], [f_2], \ldots, [f_{2g}]$ for $H_1(\partial M; \mathbb{Z})$. It may be assumed that each $[f_i]$ is represented by an oriented closed curve f_i in F. Consider the Seifert matrix A with respect to this basis. In the notation of Chapter 6, $A_{ij} = \operatorname{lk}(f_i^-, f_j)$. (If the f_i are not simple curves, they should here be changed by a very small amount in S^3 to become simple so that "linking number" makes sense.) Now the property of the base proved in Corollary

8.16 means that for $i \leq g$, there exists a non-zero integer n_i so that $n_i[f_i]$ is zero in $H_1(M; \mathbb{Z})$. But $n_i[f_i]$ can be represented by a closed curve that will be denoted $n_i f_i$, and as this bounds a 2-chain with integer coefficients, it bounds a surface mapped into M (dangerous reasoning in higher dimensions). When $n_i f_i$ is moved to $(n_i f_i)^-$, the mapped-in surface can likewise be moved across the neighbourhood of M into $M \times -1$. Thus, for $1 \leq i, j \leq g$, the curves $(n_i f_i)^-$ and $n_j f_j$ bound disjoint surfaces mapped into B^4. By Lemma 8.13, $0 = \operatorname{lk}((n_i f_i)^-, (n_j f_j)) = n_i n_j \operatorname{lk}(f_i^-, f_j)$, and so $A_{ij} = 0$ for $1 \leq i, j \leq g$. □

Now that it has been established that slice knots have Seifert matrices as described in Proposition 8.17, it is easy to produce some necessary conditions for a knot to be a slice knot.

Theorem 8.18. *If K is a slice knot, then the Conway-normalised Alexander polynomial of K is of the form $f(t)f(t^{-1})$, where f is a polynomial with integer coefficients.*

PROOF. Using the Seifert matrix of Proposition 8.17, the required Alexander polynomial is the determinant of

$$\begin{pmatrix} 0 & t^{1/2}P - t^{-1/2}Q^\tau \\ t^{1/2}Q - t^{-1/2}P^\tau & t^{1/2}R - t^{-1/2}R^\tau \end{pmatrix},$$

which is $\det(tP - Q^\tau)\det(t^{-1}P - Q^\tau)$. □

Theorem 8.19. *If K is a slice knot, then the signature of K is zero and, if the unit complex number ω is not a zero of the Alexander polynomial, then $\sigma_\omega(K) = 0$.*

PROOF. This follows at once from the fact that the signature is zero for a quadratic form coming from a non-singular symmetric bilinear form that vanishes on a subspace of half the dimension of the space concerned. A similar result holds for Hermitian forms. □

These two theorems give considerable help in establishing that a knot fails to be a slice knot. A glance at Table 8.1 immediately reveals very many non-slice knots. If the signature is zero, one can wonder if the factorisation of Theorem 8.18 occurs. Note that Theorem 8.18 implies that for a slice knot K, the determinant of K, equal by definition to $|\Delta_K(-1)|$ (see Chapter 9), is the square of an integer. As $|\Delta_K(-1)|$ is an odd integer (see Corollary 6.11), this means that $|\Delta_K(-1)| \equiv 1$ modulo 8. The knot 4_1, for example, has zero signature, but its determinant is 5 and so it cannot be a slice knot. However, the two knots shown in Figure 3.3, the Kinoshita–Terasaka and Conway knots, both have trivial Alexander polynomials and signatures, and so the above results give no information. The Kinoshita–Terasaka knot *is* a slice knot, but the slice status of the Conway knot appears to be unknown.

The topic of slice knots has here given a glimpse of knot theory in dimensions higher than 3. In general, it is quite possible to study knots of any space X in another Y. Usually the spaces are taken to be manifolds. Results in this generality are described in [47], at least in the piecewise linear framework. In that context all knots of an r-sphere S^r in an n-sphere S^n are trivial if $n - r > 2$ (see [140]). Knots of S^{n-2} in S^n have a well-developed theory, with an Alexander polynomial very similar to that for S^1 in S^3 (see [44] or [35]). A motivation for a study of slice knots is their relevance to problems of creating smooth surfaces in 4-manifolds. Suppose a surface embedded in a 4-manifold is locally knotted at a point P. In a neighbourhood of P, the surface is the cone on a knot K. If the knot is a slice knot, the cone on the knot can be replaced by the slicing disc, thus removing a point of local knottedness. Considerable progress has been made in the study of slice knots (for example, see [18]) and the theory of smooth 4-manifolds has virtually become a distinct subject on its own following spectacular progress coming from the use of differential geometry and differential equations (surveys are given in [66] and [25]). The removal of differential or piecewise linear restrictions has a remarkable effect on slice knot theory; the resulting topological slice theory is described in [30].

An extension of the slicing idea is the concept of the 4-ball genus $g^\star(K)$ of a knot K. This is the minimal genus of a surface F with the property that F includes in B^4 as a flat surface and $F \cap S^3 = \partial F \cap S^3 = K$. A slight generalisation of Theorem 8.19 shows that $|\sigma_\omega(K)| \leq 2g^\star(K)$; see [101] and [123]. It is easy to see that $g^\star(K)$ is a lower bound for the unknotting number $u(K)$. Recent work of P. B. Kronheimer and T. S. Mrowka [72], using gauge theory for smooth manifolds, shows that for K the (p, q) torus knot, $g^\star(K) = \frac{1}{2}(p-1)(q-1)$. As it is easy to show that this number of crossing changes will undo that knot, $\frac{1}{2}(p-1)(q-1)$ is the unknotting number.

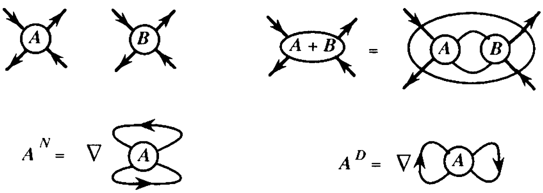
Exercises

1. Let L_n be the $(2n, 2)$ torus link as described in Chapter 1. It has two components, the linking number between them being n. Use the Conway skein formula to calculate, by means of a recurrence formula, the Conway polynomial of this link. If L'_n is L_n with the orientation of one of its components reversed, calculate in a similar way the Conway polynomial of L'_n.

2. Show that the Conway knot (shown in Figure 3.3) has Conway polynomial equal to 1.

3. Show that if knot K_1 is a mutant of knot K_2, then K_1 and K_2 have the same Conway polynomial.

4. Let A and B be diagrams of links of oriented arcs and simple closed curves in balls that meet the boundary at the four oriented points as shown below. The sum of A and B is also a diagram of such a link defined in the way shown. The "numerator" A^N and "denominator" A^D of A are the Conway polynomials of the links formed by joining up the entry and exit points in the way shown.



Prove that $(A + B)^D = A^D B^D$ and $(A + B)^N = A^N B^D + B^N A^D$.

5. Prove that the knots 6_1, 8_8, 8_9 and $3_1 + \overline{3_1}$ are all slice knots.

6. Calculate the signature of the pretzel knot $P(3, 3, -3)$.

7. Two knots K_0 and K_1 are said to be *cobordant* if there is a (piecewise linear) embedding $e : (S^1 \times D^2) \times [0, 1] \to S^3 \times [0, 1]$ so that $e^{-1}(S^3 \times \{i\}) = (S^1 \times D^2) \times \{i\}$ for $i = 0, 1$ and $e(S^1 \times 0) \times \{i\} = K_i$ for $i = 0, 1$. Prove that cobordant knots have the same signatures.

8. Prove that the unknotting number of the knot 8_2 is 2.

9. Show that the knot produced by summing together n copies of the trefoil knot 3_1 has unknotting number n. [Note. It is not known, in general, whether or not $u(K_1 + K_2) = u(K_1) + u(K_2)$.]

10. Show that the 4-ball genus $g^\star(K)$ of a knot K does indeed satisfy the inequality $|\sigma(K)| \leq 2g^\star(K)$.

9

Cyclic Branched Covers and the Goeritz Matrix

Most of this chapter will be concerned with a study of the twofold cyclic cover $X_2 \to S^3$ branched over an n component link L. The link L does not need to be oriented for this to make sense, but it will be sometimes convenient to select an arbitrary orientation in order to consider a Seifert surface. The principle result here is that the order of the first homology group $H_1(X_2)$ is det L—the determinant of the link, where $\det L = |\Delta_L(-1)|$—and that this number is often easy to calculate. As will be explained, the link determinant is, up to sign, the determinant of any Goeritz matrix [34] of the link, a matrix which is easy to write down starting from any diagram of the link.

As explained at the end of Chapter 7, X_2 can be constructed by gluing together n solid tori and two copies Y_0 and Y_1 of Y, where Y is the link exterior cut along a (connected, orientable) Seifert surface F. In the boundary of Y_i are copies $F_{i,+}$ and $F_{i,-}$ of F. The twofold cyclic cover $\widehat{X}_2$ of the link exterior is formed from the disjoint union $Y_0 \sqcup Y_1$ by identifying, in the natural way, $F_{0,+}$ with $F_{1,-}$ and $F_{1,+}$ with $F_{0,-}$. Then X_2 is created by gluing a solid torus to each boundary component of $\widehat{X}_2$, identifying a meridian of a solid torus with a lift of a square of a meridian of each component of L. Of course, $H_1(X_2)$ (with coefficients understood to be $\mathbb{Z}$), being an abelian group, is a $\mathbb{Z}$-module. The next result gives a presentation matrix for $H_1(X_2)$, in the sense of Theorem 6.1, as a $\mathbb{Z}$-module.

Theorem 9.1. *Let X_2 be the cyclic double cover of S^3 branched over a link L and suppose that A is a Seifert matrix for L with respect to some orientation and some Seifert surface. Then $H_1(X_2)$ is presented, as an abelian group, by the matrix $(A + A^\tau)$.*

PROOF. In the above notation, $\widehat{X}_2 = Y_0 \cup Y_1$, where $Y_0 \cap Y_1$ is two disjoint copies of F. A presentation of $H_1(\widehat{X}_2)$ can be obtained from the following exact Mayer–Vietoris sequence:

$$\to H_1(Y_0 \cap Y_1) \xrightarrow{\alpha_\star} H_1(Y_0) \oplus H_1(Y_1) \xrightarrow{\beta_\star} H_1(\widehat{X}_2) \to$$

$$\to H_0(Y_0 \cap Y_1) \xrightarrow{\alpha_\star} H_0(Y_0) \oplus H_0(Y_1)\,.$$

The situation is here very similar to that of Theorem 6.5, and the same sign conventions will be used. There is now a homeomorphism $t : \widehat{X}_2 \to \widehat{X}_2$ with $t^2 = 1$ which interchanges Y_0 and Y_1. As in Theorem 6.5, one can take a base $\{[f_i]\}$ for $H_1(F)$, with corresponding Seifert matrix A and dual base $\{[e_i]\}$ for $H_1(Y)$. Transferring to $\widehat{X}_2$, this gives a base $\{[f_i]\} \cup \{[tf_i]\}$ for $H_1(Y_0 \cap Y_1)$ (since $Y_0 \cap Y_1$ is two copies of F), a base $\{[e_i]\}$ for $H_1(Y_0)$ and a base $\{[te_i]\}$ for $H_1(Y_1)$. Then, with respect to these bases, $\alpha_\star$ is represented by the matrix

$$\begin{pmatrix} -A & A^\tau \\ A^\tau & -A \end{pmatrix}.$$

Similarly, using bases represented by single points, the map $H_0(Y_0 \cap Y_1) \to H_0(Y_0) \oplus H_0(Y_1)$ is represented by $\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$. Thus the kernel of this last map is a copy of $\mathbb{Z}$, and (recalling the definition of the maps in the Mayer–Vietoris sequence) any loop in $\widehat{X}_2$ that cuts each of the two components of $Y_0 \cap Y_1$ at one point maps to a generator of this copy of $\mathbb{Z}$.

Suppose that L has n components and that c_i is a closed curve in $\widehat{X}_2$ which projects to the square of the meridian of L_i, the ith component of L. Then $H_1(X_2)$ is obtained from $H_1(\widehat{X}_2)$ by equating each $[c_i]$ to zero. Consider the genus g surface F with "standard" curves $\{f_i\}$ as shown in Figure 6.1. Suppose that the "outer" boundary in the diagram is L_1. The relation $[c_1] = 0$ simply removes from $H_1(\widehat{X}_2)$ the copy of $\mathbb{Z}$ mentioned above. To achieve $H_1(X_2)$, it is then necessary to add in the relations $[c_i] = [c_1]$ for $i \geq 2$. Now, for $i \geq 2$, the curve e_{2g+i-1} in Y encircles the band of F that has L_i as part of its boundary; when regarded as a curve in the exterior of L, $[e_{2g+i-1}]$ represents the difference between the first and the ith meridians of L. Thus the element $[e_{2g+i-1}] \oplus [te_{2g+i-1}] \in H_1(Y_0) \oplus H_1(Y_1)$ is mapped by $\beta_\star$ to the difference between $[c_i]$ and $[c_1]$ in $H_1(\widehat{X}_2)$. This means that $H_1(X_2)$ is presented by the matrix

$$\begin{pmatrix} -A & A^\tau & B \\ A^\tau & -A & B \end{pmatrix},$$

where B is the $(2g + n - 1) \times (n - 1)$ matrix $\begin{pmatrix} 0 \\ I \end{pmatrix}$, I here being the $(n - 1) \times (n - 1)$ identity matrix. The permitted rules for changing a presentation matrix are described in Theorem 6.1. The operation of subtracting the first row of the blocks from the second, adding the first column of blocks to the second, and then adding each of the last $n - 1$ columns to the column preceding it by $n - 1$ places gives as an equivalent presentation matrix

$$\begin{pmatrix} -A & -A + A^\tau + (0 \oplus B) & B \\ A + A^\tau & 0 & 0 \end{pmatrix},$$

where $(0 \oplus B)$ is B preceded by $2g$ zero columns. Now $(-A + A^\tau + (0 \oplus B))$ consists (see Theorem 6.10) of g blocks of the form $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ followed by the $(n - 1) \times (n - 1)$ identity along its diagonal and zeros elsewhere. This matrix is clearly invertible over $\mathbb{Z}$. Thus the first row of blocks may be discarded, and $H_1(X_2)$ is presented by $(A + A^\tau)$. Of course, the Seifert matrix for L with respect to a different basis of $H_1(F)$ is of the form $P^\tau AP$ for some invertible matrix P, and $(P^\tau AP + P^\tau A^\tau P)$ presents the same group as $(A + A^\tau)$. □

Corollary 9.2. *Let X_2 be the double cover of S^3 branched over a link L. The order of the group $H_1(X_2)$ is the modulus of the determinant of $(A + A^\tau)$, that is*

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

PROOF. Any finitely generated abelian group can be expressed as a direct sum of cyclic groups. Thus it has as a presentation matrix a diagonal matrix, the entries on the diagonal being the orders of the summands, with the convention that an infinite group has order zero. By Theorem 6.1, the determinant of a square presentation matrix is unique up to multiplication by a unit (that is, by ± 1), so the result follows at once. The statement about the Alexander polynomial then follows from Theorem 6.5. □

Note the *caveat* that here a zero corresponds to an infinite order group. However for a *knot* it has already been shown (Corollary 6.11) that the determinant of $(A + A^\tau)$ is an odd integer. Thus the double cover of S^3 branched over a knot always has finite first homology of odd order.

Whenever the exterior X of a link L has been cut by a spanning surface F, it has been required that F be orientable. What happens when F is a non-orientable spanning surface? Suppose, then, that F is a *non-orientable* connected surface that has the link L as boundary, and let W be X-cut-along-F. Recall that X is S^3 less the interior of a regular neighbourhood $N(L)$ of L. If (by removing a small neighbourhood of ∂F in F) F is regarded as being in X, W is formed by removing from X the interior of a regular neighbourhood $N(F)$ of F. Locally $N(F)$ is a product of part of F with the unit interval I. Thus the orientable manifold $N(F)$ is an I-bundle over the non-orientable surface F. The associated ∂I bundle gives a two-to-one covering map from a connected orientable surface $\widetilde{F}$ to F. Thus $\widetilde{F}$ is the orientable double covering space (see Chapter 7) of F, and $N(F)$ is the mapping cylinder of the covering map $p : \widetilde{F} \to F$. If f is a closed loop in F, then $p^{-1}f$ is a single loop (that double covers f) if f is orientation reversing and is the union of two loops if f preserves orientation. In [36], Gordon and Litherland defined a quadratic form

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

by $\mathcal{G}_F([f],[g]) = \operatorname{lk}(p^{-1}f, g)$, where f and g are oriented loops in F. (Thus $\mathcal{G}_F([f],[g])$ is the linking number of g with f pushed off F locally in "both directions".) It is clear that this Gordon–Litherland form gives a well-defined bilinear map and, by considering signs of crossings, that $\mathcal{G}_F$ is symmetric. Of course, this definition still makes sense when F is orientable, and $p^{-1}f$ is always two copies of f, one on either side of F. The form is then sometimes called the Trotter form [124], and it is represented by $A + A^\tau$ where A is any Seifert matrix.

It has already been seen above that $A + A^{\tau}$ is a presentation matrix for $H_1(X_2)$. This will be extended to the non-orientable surfaces in Theorem 9.3.

Returning to the situation where F is a non-orientable connected surface spanning L, the surface $\widetilde{F}$ is a connected subspace of ∂W. The map that interchanges the two end points of each fibre of the above I bundle gives a homeomorphism $t : \widetilde{F} \to \widetilde{F}$ such that $t^2 = 1$ and $W/t = X$. One cannot imitate the orientable situation, taking infinitely many copies of W and gluing them together in a sequence, in any sensible way, for ∂W does not contain two copies of F. However, one can take *two* copies of W, W_0 and W_1, with copies $\widetilde{F}_0$ and $\widetilde{F}_1$ of $\widetilde{F}$ in their boundaries, and for each $x \in \widetilde{F}$ identify the copy of x in $\widetilde{F}_0$ with the copy of tx in $\widetilde{F}_1$. This constructs a cover of X. A loop in X lifts to a *loop* in this cover if and only if it meets F at an even number of points—that is, if and only if it has even linking number with L. As this property characterises the double cyclic cover $\widehat{X}_2$ of X, it is precisely that cover which has been constructed from the two copies of W. Solid tori can then be added, if desired, to obtain the double branched cover X_2.

Theorem 9.3. *Suppose that F is a connected surface spanning a link L; then any matrix representing the form $\mathcal{G}_F : H_1(F) \times H_1(F) \to \mathbb{Z}$ is a presentation matrix for $H_1(X_2)$.*

PROOF. The previous theorem dealt with the case when F is orientable, so suppose that F is a connected non-orientable surface spanning the n-component link L. To calculate $H_1(X_2)$ from this, consider the exact Mayer–Vietoris sequence

$$\longrightarrow H_1(W_0 \cap W_1) \xrightarrow{\alpha_\star} H_1(W_0) \oplus H_1(W_1) \xrightarrow{\beta_\star} H_1(\widehat{X}_2) \to .$$

Because $W_0 \cap W_1$ is a copy of the *connected* surface $\widetilde{F}$, the map $\beta_\star$ is a surjection. In the abstract, regard F as the surface (together with generating curves) shown in Figure 6.1, with the addition of one twisted band, or of two interlocking bands one of which is twisted, together with the extra generating curves as shown in Figure 9.1. Any non-orientable surface can be regarded as being of one of these two types.

Consider the first type of surface with a generating curve g as shown. Exactly as in the orientable case, $H_1(F)$ is freely generated by $\{[f_i] : i = 1, 2, \ldots, 2g + n - 1\} \cup \{[g]\}$, and there is a dual base $\{[e_j] : j = 1, 2, \ldots, 2g + n\}$ freely

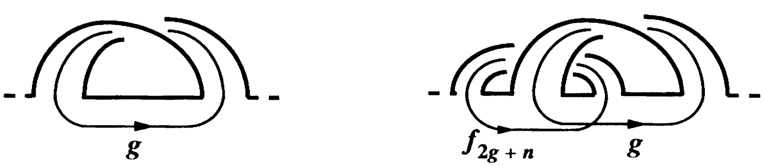


Figure 9.1

generating $H_1(W)$. For each i let $\widetilde{f}_i$ and $t\widetilde{f}_i$ be the two lifts of f_i to $\widetilde{F}$, and let $\widetilde{g}$ be $p^{-1}g$. The classes of these curves are a free base for $H_1(\widetilde{F})$. If $\iota : \partial W \to W$ is the inclusion, there are two $(2g+n) \times (2g+n-1)$ matrices R and S and a $(2g+n) \times 1$ matrix λ such that

$$\iota_\star[\widetilde{f}_i] = \sum_j R_{ji}[e_j], \quad \iota_\star[t\widetilde{f}_i] = \sum_j S_{ji}[e_j] \text{ and } \iota_\star[\widetilde{g}] = \sum_j \lambda_j[e_j].$$

Hence the map $\alpha_\star$ in the above Mayer–Vietoris sequence is represented by

$$\begin{pmatrix} R & S & \lambda \\ -S & -R & -\lambda \end{pmatrix},$$

which is thus a presentation matrix for $H_1(\widehat{X}_2)$.

It remains to consider the effect of gluing solid tori to $\widehat{X}_2$. Consider the curves x and y on the boundary of $N(F)$, as shown in Figure 9.2. Suppose that ξ and η are column matrices such that $\iota_\star[x] = \sum_j \xi_j[e_j]$ and $\iota_\star[y] = \sum_j \eta_j[e_j]$. Inspection of Figure 9.2 shows that $\eta - \xi = \lambda$, and $\eta + \xi$ is the column with 1 in the final place and zeros elsewhere (being the coordinates of $[e_{2g+n}]$).

Figure 9.2

The effect of gluing the first solid torus to $\widehat{X}_2$ is to equate to zero the element $\iota_\star[x] \oplus \iota_\star[y] \in H_1(W_0) \oplus H_1(W_1)$. Gluing on any of the other solid tori has an effect analogous to that observed in Theorem 9.1; it equates to zero the elements of the form $[e_{2g+i-1}] \oplus [e_{2g+i-1}] \in H_1(W_0) \oplus H_1(W_1)$ for $2 \le i \le n$. Hence $H_1(X_2)$ has a presentation matrix of the form

$$\begin{pmatrix} R & S & \lambda & \xi & B \\ -S & -R & -\lambda & \eta & B \end{pmatrix},$$

where B is the $(2g+n) \times (n-1)$ matrix with $B_{2g+j,j} = 1$ for $j = 1, 2, \ldots, n-1$ and $B_{i,j} = 0$ otherwise. Subtracting the second row of blocks from the first produces

$$\begin{pmatrix} R+S & R+S & 2\lambda & -\lambda & 0 \\ -S & -R & -\lambda & \eta & B \end{pmatrix}.$$

Subtracting the first column of blocks from the second and adding twice the fourth column to the third gives

$$\begin{pmatrix} R+S & 0 & 0 & -\lambda & 0 \\ S & S-R & \xi+\eta & \eta & B \end{pmatrix}.$$

As in the proof of Theorem 9.2, $(2g+n)\times(2g+n)$ matrix $(S-R \quad \xi+\eta)$ consists of g blocks up to sign of the form $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ down the diagonal, a 1 in the final place on the diagonal and zeros elsewhere. Hence, if $(0 \oplus B \oplus 0)$ is B preceded by $2g$ zero columns and followed by one zero column, $(S-R \quad \xi+\eta)+(0\oplus B\oplus 0)$ is invertible. Thus another presentation matrix of the same group is $(R+S \quad -\lambda)$ or equivalently $(R+S \quad \lambda)$. However, this matrix represents the quadratic form $\mathcal{G}_F$ with respect to the given base of $H_1(F)$. As with any quadratic form, changing the base changes the matrix to one of the form $P^\tau(R+S \quad \lambda)P$, where P is invertible, and this presents the same group.

Finally, it remains to consider what happens when F is a surface of the second type shown in Figure 9.1. The situation is much the same as before, except that now $H_1(F)$ and $H_1(W)$ each has $2g+n+1$ generators, so that R and S are $(2g+n+1)\times(2g+n)$ matrices. However, $\xi+\eta$ is a column with a 1 in each of the last *two* places and zeros elsewhere. Now, $(S-R)$ has g blocks each up to sign of the form $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ down the diagonal and a 1 in the $(2g+n+1,\ 2g+n)$ place. Thus $(S-R \quad \xi+\eta)+(0\oplus B\oplus 0)$ is again invertible (where the second "$\oplus$ 0" is two columns of zeros). The discussion then proceeds as before. □

A Goeritz matrix for a link is a matrix of integers constructed in the following way: Let D be a connected diagram of a link L and let the regions of the diagram be coloured black and white in chessboard fashion. Given this colouring, an incidence number $\zeta(c) = \pm 1$ can be allocated to any crossing c, as in Figure 9.3. Let $R_0, R_1, \ldots, R_n$ be the white regions of the diagram. Define a "pre-Goeritz matrix" to be the $(n+1)\times(n+1)$ matrix having terms $\{g_{ij}\}$ given, for $i \neq j$, by

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

where the sum is over all crossings at which R_i and R_j come together. Define diagonal terms by

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

The related Goeritz matrix G is this matrix with a row and corresponding column deleted. It may be assumed that the labelling is such that it is the row and column indexed by zero that are deleted. Thus G is the $n\times n$ matrix $\{g_{ij} : 1 \le i, j \le n\}$. Of course G depends on the diagram chosen for L, on which regions are called white, and on the labelling of those white regions. The following result is taken from [36].

Figure 9.3

Theorem 9.4. *Any Goeritz matrix for a link L, associated with the white regions of a diagram of L, represents, with respect to some base, the Gordon–Litherland form*

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

where F is the spanning surface for L given by the black regions of the diagram.

PROOF. Let the white regions, $R_0, R_1, \ldots, R_n$, of the diagram inherit an orientation from the sphere S^2 in which they are assumed to lie; thus each ∂R_i has an orientation. Let f_i be the oriented simple closed curve in F that consists of ∂R_i pushed into the union of the black regions. Then $\{[f_i] : 0 \le i \le n\}$ forms a set of generators for $H_1(F)$; any subset of n of the $\{[f_i]\}$ forms a base for $H_1(F)$. Suppose that the white regions R_i and R_j are both incident at a crossing c where $\zeta(c) = +1$. Then in the above notation, the curve or curves $p^{-1}f_j$, namely the push-off of f_j from F locally to both sides of F, meet R_i in a positive point of intersection and meet R_j in a negative point of intersection near to c. See Figure 9.4. The sign is positive if the orientation of the region is in the sense of a right-hand screw with respect to the orientation of $p^{-1}f_j$. The signs are reversed if $\zeta(c) = -1$. Thus for $i \neq j$, $\text{lk}(p^{-1}f_j, f_i) = \sum \zeta(c)$, where the sum is over all crossings at which R_i and R_j come together, and $\text{lk}(p^{-1}f_j, f_j) = -\sum \zeta(c)$, the sum being over all c at which R_j is incident with *other* regions. Note the two points of $p^{-1}f_j \cap R_j$ near a crossing at which R_j is incident with itself cancel each other. Hence the quadratic form $\mathcal{G}_F$ is represented with respect to the base $[f_1], [f_2], \ldots, [f_n]$ by the Goeritz matrix of the diagram with the above labelling of the white regions. □

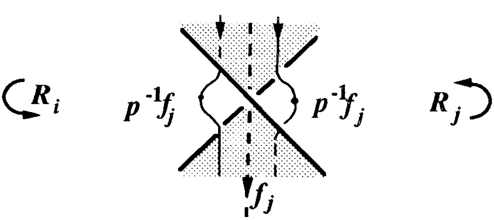


Figure 9.4

Corollary 9.5. *The determinant of L, $|\Delta_L(-1)|$, is equal to $|\det G|$, where G is any Goertiz matrix for L.*

The proof of this is immediate from the last three theorems. It follows that $|\det G|$ is an invariant of L, and, as a Goeritz matrix is often easy to write down, it can be a useful invariant.

As an example, consider the diagram with $n+2$ crossings of a twisted double of the unknot shown in Figure 6.3. The diagram there shown has its regions coloured in chessboard fashion with three white regions. Suppose the outer region is R_0, that R_1 is the region abutting only two crossings, and that R_2 is the other region.

The pre-Goeritz matrix is

$$\begin{pmatrix} n+1 & -1 & -n \\ -1 & 2 & -1 \\ -n & -1 & n+1 \end{pmatrix},$$

so that $\begin{pmatrix} 2 & -1 \\ -1 & n+1 \end{pmatrix}$ is a Goeritz matrix and the determinant of the knot is $|2n + 1|$. Note that this simple invariant is enough to distinguish all these knots from each other when $n \geq 0$.

As a second favourite example, take the pretzel knot or link $P(p, q, r)$ shown by a coloured diagram in Figure 6.4. The pre-Goeritz matrix is

$$\begin{pmatrix} r+p & -p & -r \\ -p & p+q & -q \\ -r & -q & q+r \end{pmatrix},$$

the Goeritz matrix is $\begin{pmatrix} p+q & -q \\ -q & q+r \end{pmatrix}$ and the determinant is $|pq + qr + rp|$. Note that this last determinant can be equal to 1 (for example when $(p, q, r) = (-3, 5, 7)$). Then the double cover of S^3 branched over the link has trivial first homology group; standard results in homology theory then imply that it has all the same homology groups as S^3.

More information about the Goeritz matrix can be found in [36]. In particular, the signature of the link can be calculated from the signature of G together with a simple "correction term". A variant of the proof given here for Theorem 8.2 shows that the Goeritz matrix of a knot is well defined up to moves that change G to $P^\tau G P$ for an invertible matrix of integers P, or to $\begin{pmatrix} G & 0 \\ 0 & \pm 1 \end{pmatrix}$ or the reverse move. Of course, from this the invariance of $|\det G|$ follows at once. One can likewise easily check this invariance directly from the Reidemeister moves. These last remarks must be qualified a little if links are considered rather than knots (see [36]).

A result that connects the idea of the determinant with link polynomials is the following. In the language of Chapters 15 and 16, it states that $(\det L)^2$ is the value of the Kauffman polynomial of L when $(1, 2)$ is substituted for the pair of variables of that polynomial.

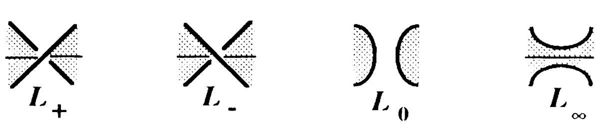


Figure 9.5

Theorem 9.6. *Suppose that L_+, L_-, L_0 and L_∞ are four links that have identical diagrams except near a point where they are as shown in Figure 9.5. Then*

$$(\det L_+)^2 + (\det L_-)^2 = 2((\det L_0)^2 + (\det L_\infty)^2).$$

PROOF. The diagram shows the four links together with connected shaded spanning surfaces F_i for $i = +, -, 0, \infty$. These can always be constructed by using Seifert's method (see Chapter 2) for F_0 and adding bands to get the other three surfaces. The four surfaces are taken to be identical outside the areas shown. Take closed curves in F_0 representing a base of $H_1(F_0)$ and, for bases of $H_1(F_i)$ for $i = +, -, \infty$, take the classes of the extra curves shown in the diagrams (the ends of them are joined up outside the diagrams) together with the set of curves already chosen for F_0. Matrices M_i for the Gordon–Litherland forms $\mathcal{G}_{F_i}$ with respect to these bases are of the following form:

$$M_\infty = \begin{pmatrix} n & \rho \\ \rho^\tau & M_0 \end{pmatrix}, \qquad M_\pm = \begin{pmatrix} n \mp 1 & \rho \\ \rho^\tau & M_0 \end{pmatrix}.$$

Thus

$$\det M_\pm = \det M_\infty \mp \det M_0.$$

Squaring and adding give the required result. □

Recall that the r-fold cyclic cover of S^3 branched over an n-component link is constructed by adding n solid tori to the space formed by gluing together, in a cyclic fashion, r copies of the link's exterior cut along a (connected) Seifert surface. The following result and its proof are direct generalisations of those of Theorem 9.1. The details will thus not be pursued; note however that everything simplifies a little when the link is a knot (that is, when $n = 1$).

Theorem 9.7. *Let X_r be the cyclic r-fold cover of S^3 branched over an n-component oriented link L, and suppose that A is a Seifert matrix for L coming from a genus g Seifert surface. Then $H_1(X_r)$ is presented, as an abelian group, by the $r \times (r+1)$ matrix of blocks*

$$\begin{pmatrix} -A^\tau & & & & A & B \\ A & -A^\tau & & & & B \\ & A & -A^\tau & & & B \\ & & \ddots & \ddots & & \vdots \\ & & & A & -A^\tau & B \end{pmatrix},$$

where B is the $(2g + n - 1) \times (n - 1)$ matrix $\begin{pmatrix} 0 \\ I \end{pmatrix}$.

Corollary 9.8. *The order of the first homology group of X_r, the cyclic r-fold cover of S^3 branched over L, is given by*

$$|H_1(X_r)| = \left|\prod_{j=1}^{r-1} \Delta_L(e^{2\pi i j/r})\right|.$$

Assuming that a "standard" base has been used for the homology of the Seifert surface, the presentation matrix given in the above theorem can be manipulated in the following way: Add to the first column of blocks all the other columns of blocks except the last one, so that every block entry in the first column becomes $A - A^\tau$. Then by rearranging the first $2g + n - 1$ columns of the matrix and the last $n - 1$ columns, deleting zero columns and changing some signs of columns, obtain, as an alternative presentation matrix,

$$\begin{pmatrix} I & & & & A \\ I & -A^\tau & & & \\ I & A & -A^\tau & & \\ \vdots & & \ddots & \ddots & \\ I & & & A & -A^\tau \end{pmatrix}.$$

This is an $r \times r$ matrix of $(2g + n - 1) \times (2g + n - 1)$ blocks. The proof of the corollary consists of the (not entirely trivial) exercise in linear algebra of evaluating the determinant of this matrix, using the fact that $\Delta_L(t)$ is, up to a unit, det $(tA - A^\tau)$.

Exercises

1. Use the Goeritz matrix to find the determinant of the knot 8_{18}.
2. Find some knots K for which the double cover of S^3 branched of K has zero first homology group (and hence has the same homology groups as S^3).
3. Show that Theorem 9.6, together with the fact that the determinant of the unknot is 1, can be used to calculate the determinant of any link. Illustrate the method with a calculation for the knot 4_1.
4. Let C be a knot. Let K be the (cable) satellite of C that consists of a simple closed curve, on the boundary of the solid torus neighbourhood $N(C)$ of C, which is homologous to two longitudes plus $2n + 1$ meridians. Thus K bounds a knotted Möbius band contained in a neighbourhood of C. Use the Gordon–Litherland form associated with this Möbius band to find the determinant of K.
5. Use the Gordon–Litherland form to determine the determinant of the pretzel knot $P(p, q, r)$ when p is an even integer and q and r are both odd.
6. Show directly that the modulus of the determinant of G, the Goeritz matrix associated to the white regions of a knot diagram, does not change if the diagram is changed by a Reidemeister move. What happens if attention is switched to the black regions?
7. If two knots have diagrams giving rise to the same Goeritz matrix, in what way are the knots related?

10

The Arf Invariant and the Jones Polynomial

The original Arf invariant was an invariant of certain quadratic forms on a vector space over a field of characteristic 2. This can be applied to a quadratic form, closely associated to the Seifert form, on the first homology with $\mathbb{Z}/2\mathbb{Z}$ coefficients of a Seifert surface of an oriented link L. The result is a fairly classical link invariant $\mathcal{A}(L) \in \mathbb{Z}/2\mathbb{Z}$ called the Arf (or Robertello) invariant of L ([111], [114]). It must, however, be stated at once that for this theory to work—that is, for $\mathcal{A}(L)$ to be defined—L must satisfy the condition that the linking number of any component with the remainder of the link should be an even number. Before the discovery of the Jones polynomial, efforts to find a sensible generalisation of the Arf invariant to *all* links met with no success. The Jones polynomial $V(L)$ is, of course, always defined. As will be shown in what follows, evaluating $V(L)$ when $t = i$ (with $t^{1/2} = e^{i\pi/4}$) gives

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

where $\#L$ is the number of components of L. In a sense, this shows why a definition of $\mathcal{A}(L)$ for any link could not be found. Interpreted from the point of view of the Jones polynomial, this result gives one of the very few evaluations of the polynomial in terms of previously known invariants that can be calculated in "polynomial time" (see Chapter 16). This chapter will first explore the Arf invariant for vector spaces over $\mathbb{Z}/2\mathbb{Z}$ and then effect liaison with the Jones polynomial.

In what follows, let V be a finite-dimensional vector space over $\mathbb{Z}/2\mathbb{Z}$, the field of two elements $\{0, 1\}$.

Definition 10.1. A function $\psi : V \to \mathbb{Z}/2\mathbb{Z}$ is a quadratic form if for some bilinear map $\mathcal{F} : V \times V \to \mathbb{Z}/2\mathbb{Z}$,

$$\psi(u + v) + \psi(u) + \psi(v) = \mathcal{F}(u, v)$$

for all $u, v \in V$. The quadratic form is called non-singular if $\mathcal{F}$ is non-singular (that is, for each non-zero $u \in V$, $\mathcal{F}(u, v) \neq 0$ for some $v \in V$).

Note that $\psi(0) = 0$, $\mathcal{F}(u, u) = 0$, $\mathcal{F}$ is symmetric (which is here the same as skew-symmetric) and $\psi(\lambda u) = \lambda\psi(u) = \lambda^2\psi(u)$ for $\lambda \in \{0, 1\}$.

If $\mathcal{F}$ is non-singular, the usual arguments for real skew-symmetric forms imply that there is a base $e_1, f_1, e_2, f_2, \ldots, e_n, f_n$ for V with respect to which $\mathcal{F}$ is represented by a matrix of the form

$$\begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & 0 & \cdots & 0 \\ 0 & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{pmatrix}.$$

This implies that V must have even dimension. Such a base is called *symplectic*. Using a symplectic base, it follows (for instance, by induction on n), that

$$\psi(x_1e_1 + y_1f_1 + \cdots + x_ne_n + y_nf_n) = \sum_1^n x_i^2\psi(e_i) + \sum_1^n y_i^2\psi(f_i) + \sum_1^n x_iy_i.$$

Consider the identity $x_1^2 + x_1y_1 = x_1(x_1 + y_1)$. If $\psi(e_1) = 1$ and $\psi(f_1) = 0$, this identity can be used to construct a new symplectic base (starting with $\{e_1 + f_1,\ f_1\}$) so that with the new base neither the term in x_1^2 nor the term in y_1^2 appears. Similarly, when $\psi(e_1)$, $\psi(e_2)$, $\psi(f_1)$ and $\psi(f_2)$ are all non-zero, a new symplectic base starting with

$$\{\, (e_1 + e_2 + f_1),\ (e_1 + f_1 + f_2),\ (e_1 + e_2 + f_2),\ (e_2 + f_1 + f_2) \,\}$$

can be chosen to remove the squared terms; this corresponds to the identity

$$\begin{aligned} &x_1^2 + x_1y_1 + y_1^2 + x_2^2 + x_2y_2 + y_2^2 \\ &\quad = (x_1 + y_1 + x_2)(x_1 + y_1 + y_2) + (x_1 + x_2 + y_2)(y_1 + x_2 + y_2). \end{aligned}$$

Thus, a symplectic base can be chosen with respect to which $\psi(x_1e_1 + y_1f_1 + \cdots + x_ne_n + y_nf_n)$ is of one of the two following "Types".

Type 1: $x_1y_1 + x_2y_2 + \cdots + x_ny_n$,
Type 2: $x_1y_1 + x_2y_2 + \cdots + x_ny_n + x_n^2 + y_n^2$.

Definition 10.2. The Arf invariant $c(\psi)$ of the non-singular quadratic form $\psi : V \to \mathbb{Z}/2\mathbb{Z}$ is the value, 0 or 1, taken more often by $\psi(u)$ as u varies over the 2^{2n} elements of V.

It is easy to show, by induction on n, that the value 1 is taken $2^{2n-1} - 2^{n-1}$ times by $\psi(u)$ if ψ is of Type 1 and $2^{2n-1} + 2^{n-1}$ times if ψ is of Type 2. Hence $c(\psi)$ is always defined, no choice is involved in its definition and

$$c(\psi) = \begin{cases} 0 & \text{if } \psi \text{ is of Type 1,} \\ 1 & \text{if } \psi \text{ is of Type 2.} \end{cases}$$

Note that if ψ_1 and ψ_2 are quadratic forms on V_1 and V_2, respectively, then the form $\psi_1 \oplus \psi_2$ on $V_1 \oplus V_2$ has $c(\psi_1 \oplus \psi_2) = c(\psi_1) + c(\psi_2)$ modulo 2. This follows by checking the possible Types. Note also that if $e_1, f_1, e_2, f_2, \ldots, e_n, f_n$ is any symplectic base, then

$$c(\psi) = \sum_{i=1}^{n} \psi(e_i)\psi(f_i)\,.$$

The above theory of $\mathbb{Z}/2\mathbb{Z}$ quadratic forms is applied to links in the following way: Let L be an oriented link in S^3 with Seifert surface F, the orientation being needed to define F. Define $q : H_1(F; \mathbb{Z}/2\mathbb{Z}) \to \mathbb{Z}/2\mathbb{Z}$ by $q(x) = \alpha_2(x, x) \in \mathbb{Z}/2\mathbb{Z}$, where α_2 is the Seifert form α (see Chapter 6) reduced modulo 2. Thus if x is (represented by) a simple closed curve on F, $q(x)$ is the number, modulo 2, of twists in an annular neighbourhood of x in F. Then

$$q(x + y) + q(x) + q(y) = \alpha_2(x, y) + \alpha_2(y, x) = \mathcal{F}(x, y),$$

where $\mathcal{F}$ is the intersection form (which just counts the number of intersection points of transverse curves) modulo 2 on $H_1(F; \mathbb{Z}/2\mathbb{Z})$. However, a glance at the base shown in Figure 6.1 reveals that $\mathcal{F}$ is non-singular only when L has one component. A second glance shows that $\mathcal{F}$ induces a non-singular form on the quotient $H_1(F; \mathbb{Z}/2\mathbb{Z})/\iota_\star H_1(\partial F; \mathbb{Z}/2\mathbb{Z})$, where ι is the inclusion map. Suppose that L has components $\{ L_i \}$ and that L has the property that

$$(\star) \qquad \operatorname{lk}(L_i,\ L - L_i) \equiv 0 \text{ modulo } 2.$$

Then $q([L_i]) \equiv \operatorname{lk}(L_i^-, L_i) = \operatorname{lk}(L_i,\ L - L_i) \equiv 0$ modulo 2, as L_i is homologous to $L - L_i$ in the complement of L_i^-. For any $x \in H_1(F; \mathbb{Z}/2\mathbb{Z})$, clearly $\mathcal{F}(x, [L_i]) = 0$, so $q(x + [L_i]) = q(x)$, and hence q induces a well-defined non-singular quadratic form $q : H_1(F; \mathbb{Z}/2\mathbb{Z})/\iota_\star H_1(\partial F; \mathbb{Z}/2\mathbb{Z}) \to \mathbb{Z}/2\mathbb{Z}$.

Definition 10.3. The Arf invariant $\mathcal{A}(L)$ of an oriented link L having the property $(\star)$ is $c(q)$, where $q : H_1(F; \mathbb{Z}/2\mathbb{Z})/\iota_\star H_1(\partial F; \mathbb{Z}/2\mathbb{Z}) \to \mathbb{Z}/2\mathbb{Z}$ is the quadratic form described above.

Proposition 10.4. *The Arf invariant $\mathcal{A}(L)$ for an oriented link L having property $(\star)$ is well defined.*

PROOF. It is necessary to check that $\mathcal{A}(L)$ does not depend on the choice of Seifert surface F. By Theorem 8.2, it is only necessary to check what happens when F is changed to F' by embedded surgery along an arc in S^3. Suppose that $\{e_1, f_1, e_2, f_2, \ldots, e_n, f_n\}$ is a symplectic base for $H_1(F; \mathbb{Z}/2\mathbb{Z})/\iota_\star H_1(\partial F; \mathbb{Z}/2\mathbb{Z})$ represented by simple closed curves (for example the first $2g$ curves, renamed, of Figure 6.1). That base can be augmented by $\{e_{n+1}, f_{n+1}\}$ to give a symplectic base for $H_1(F'; \mathbb{Z}/2\mathbb{Z})/\iota_\star H_1(\partial F'; \mathbb{Z}/2\mathbb{Z})$: Choose e_{n+1} to be represented by a simple closed curve encircling once the solid cylinder defining the embedded surgery, that curve being met at exactly one point by a simple closed curve representing f_{n+1}. Note that an isotopy of the end points of the surgery arc α ensures that the two points of $\partial\alpha$ are not separated by any base curve. Then $q(e_{n+1}) = 0$, and so $\sum_{i=1}^{n} q(e_i)q(f_i) = \sum_{i=1}^{n+1} q(e_i)q(f_i)$. □

Note that $\mathcal{A}$(the unknot) $= 0$ and $\mathcal{A}$(the trefoil) $= 1$, for as shown in Figure 6.3 (when $n = 1$), the trefoil has a symplectic base $\{e_1, f_1\}$ for which $q(e_1) = q(f_1) = 1$. Note, too, that the addition formula for the Arf invariant of the direct sum of two quadratic forms implies that $\mathcal{A}(L + L') = \mathcal{A}(L) + \mathcal{A}(L')$ for any two links L and L' having property ($\star$) (whatever components are chosen for the summing operation).

Lemma 10.5. *Suppose that L and L′ are oriented links having property ($\star$) which are the same except near one point, where they are as shown in Figure 10.1; then* $\mathcal{A}(L) = \mathcal{A}(L')$.

PROOF. The two segments shown on one of the two sides of Figure 10.1 must belong to the same component of the link. Suppose, without loss of generality, it is the two segments on the left side. Then using the Seifert circuit method of Theorem 2.2, a Seifert surface can be constructed for the left link that meets the neighbourhood of the point in question in the way indicated by the shading. Adding a band to that produces a Seifert surface for the right link as indicated. Now, as these two surfaces just differ by a band added to the boundary, the $\mathbb{Z}/2\mathbb{Z}$-homology of the second surface is just that of the first surface with an extra $\mathbb{Z}/2\mathbb{Z}$ summand. However, that summand is in the image of the homology of the boundary of the surface; this image is disregarded (by means of the quotienting) in construction of the quadratic form that gives the Arf invariant. □

Figure 10.1

Note that elementary consideration of linking numbers shows the following: If the two segments of the link L shown on one side of Figure 10.1 belong to *distinct* components, and if L has the property ($\star$), then L' also has the property ($\star$).

With the definition of the Arf invariant and its elementary properties now established, its relevance to the Jones polynomial can now be considered. The result linking the two topics is as follows:

Theorem 10.6. *The Jones polynomial of any oriented link L in S^3, evaluated at $t = i$ (with $t^{1/2} = e^{i\pi/4}$), is given by*

$$V(L)_{(t=i)} = \begin{cases} (-\sqrt{2})^{\#L-1}(-1)^{\mathcal{A}(L)} & \textit{if } L \textit{ has property } (\star), \\ 0 & \textit{otherwise,} \end{cases}$$

where #L is the number of components of L and $\mathcal{A}(L)$ is its Arf invariant

PROOF. Define $A(L)$ to be the integer given by

$$A(L) = \begin{cases} (-1)^{\mathcal{A}(L)} & \text{if } L \text{ has property } (\star), \\ 0 & \text{otherwise.} \end{cases}$$

Now suppose that L_+, L_- and L_0 are three oriented links that are exactly the same except near a point where they are as shown in Figure 3.2 (the usual relationship). The proof considers two cases as follows:

Suppose first that the two segments of L_+ near the point in question are parts of the same component of L_+. (Then either both L_+ and L_- have property $(\star)$ or neither of them does.) If L_0 has property $(\star)$ so, by the above remark, do L_+ and L_-, and by Lemma 10.5, $\mathcal{A}(L_0) = \mathcal{A}(L_+) = \mathcal{A}(L_-)$. Thus certainly

$$A(L_+) + A(L_-) - 2A(L_0) = 0,$$

an equation that also, trivially, holds if none of L_+, L_- or L_0 has property $(\star)$. There remains the possibility that L_+ and L_- have property $(\star)$ but that L_0 does not. Consider the two links shown in Figure 10.2. It is easy to check that the first link, X say, has property $(\star)$, and so its Arf invariant exists and by Lemma 10.5, $\mathcal{A}(L_+) = \mathcal{A}(X)$. The second link is just L_- in disguise. It can also be thought of as X first summed with a trefoil knot and then having two components banded together. Thus, again using Lemma 10.5, $\mathcal{A}(X) + 1 = \mathcal{A}(L_-)$ modulo 2. Hence again it is true that $A(L_+) + A(L_-) - 2A(L_0) = 0$.

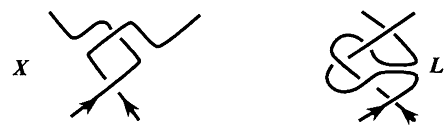


Figure 10.2

Secondly, suppose that the two segments of L_+ near the point in question are parts of different components of L_+. If L_0 does not have property $(\star)$ then neither do L_+ and L_-, and so trivially

$$A(L_+) + A(L_-) - A(L_0) = 0.$$

Otherwise L_0 and *one* of L_+ and L_- has property $(\star)$, and this formula is again true (using Lemma 10.5).

If $\widehat{A}(L)$ denotes $(-\sqrt{2})^{\#L-1}A(L)$, then the two preceding displayed formulae both become

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

and of course if L is the unknot, $\widehat{A}(L) = 1$. However, as discussed in Chapter 3, the Jones polynomial $V(L) \in \mathbb{Z}[t^{-1/2}, t^{1/2}]$ is characterised by being 1 on the unknot and by satisfying

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

Substituting $t^{\frac{1}{2}} = e^{i\pi/4}$ reduces this to exactly the above formula for $\widehat{A}$. □

If, in the notation used in the above proof, L_+ is a knot, then so is L_-, and L_0 is a link of two components. Of course, L_0 has the property $(\star)$ if and only if $\operatorname{lk}(L_0)$, the linking number of the two component of L_0, is even. The second paragraph of the above proof shows that $\mathcal{A}(L_+) - \mathcal{A}(L_-) \equiv \operatorname{lk}(L_0)$ modulo 2 .

Theorem 10.7. *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 $\nabla_K(z)$. The Arf invariant of K is related to the Alexander polynomial by*

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

If K is a slice knot, then $\mathcal{A}(K) = 0$.

PROOF. The formula $\mathcal{A}(L_+) - \mathcal{A}(L_-) \equiv \operatorname{lk}(L_0)$ modulo 2, valid when L_+ has one component, allows calculation of $\mathcal{A}(K)$ from $\mathcal{A}(\text{unknot}) = 0$. However, this gives the same answer as the calculation, modulo 2, of $a_2(K)$ using Proposition 8.7 (v). With the Conway normalisation, $\Delta_K(-1) = \nabla_K(-2i)$. However, $\nabla_K(z) = 1 + a_2(K)z^2 + a_4(K)z^4 + \cdots$, and so $\nabla_K(-2i) \equiv 1 - 4a_2(K)$ modulo 8. Thus, modulo 8,

$$\nabla_K(-2i) \equiv \begin{cases} 1 & \text{if } a_2(K) \equiv 0 \text{ modulo } 2, \\ -3 & \text{if } a_2(K) \equiv 1 \text{ modulo } 2. \end{cases}$$

This gives the required result. As remarked after Theorem 8.19, if K is a slice knot then $\Delta_K(-1) \equiv \pm 1$ modulo 8, and so, from the above discussion, $\mathcal{A}(K) = 0$. □

The result given above (due to J. Levine [75]), relating the determinant of a knot with the Arf invariant, has been stated with the Alexander polynomial determined only up to multiplication by $\pm t^{\pm n}$. However, as shown in the proof, $\Delta_K(-1) \equiv 1$ modulo 4 when the Conway normalisation is employed.

The vanishing of the Arf invariant on slice knots does suggest that the Arf invariant is connected with 4-dimensional topology. In fact, the Arf invariant of a link is intimately related to the Rohlin invariant of a 4-manifold with spin structure. Indeed, A. J. Casson has given a proof of the Rholin theorem based on the Arf invariant of a link. This theorem asserts that the signature of a closed smooth orientable spin 4-manifold is divisible by 16 (see [29] and [66]).

Exercises

1. Make a table of the Arf invariants of the prime knots with at most eight crossings.
2. Determine, directly from a Seifert matrix, the Arf invariant of the pretzel knot $P(p, q, r)$, where p, q and r are odd integers.
3. Prove that cobordant knots have the same Arf invariant (see Exercise 7 of Chapter 8).
4. Use Lemma 10.5 to show that if L is a trivial link of unknotted unlinked components, then the Arf invariant of L is zero. By considering the maxima, minima and saddles

of a slice disc for a slice knot K (as for example in Figure 8.4), show directly from Lemma 10.5 that a slice knot has zero Arf invariant.

5. Suppose L is an oriented link for which the Arf invariant is defined. Suppose that L' is obtained by reversing the orientation of one component of L. Is the Arf invariant of L' defined? If so, how is it related to the Arf invariant of L?

11

The Fundamental Group

It is in its interaction with the theory of the fundamental group that the theory of knots and links becomes almost a part of the general theory of 3-manifolds. It is the exterior of a link (that is, the closure of the complement in S^3 of a small regular neighbourhood of the link) that is studied, by means of its group, as a compact 3-manifold with torus boundary components. In the theory of 3-manifolds this is a very important example, but perhaps not much more than that. Here the view has been taken that to a mathematician it is the proving of results that brings satisfaction, and that this is particularly important in knot theory, wherein a cheerful punter might be satisfied by a good diagram. However, 3-manifold theory is well documented at length elsewhere ([43], [49]), and other more established treatises on knots have dwelt comprehensively on the relationship between links and the fundamental group. Thus what follows in this chapter is but an essay on this topic. It tries to interpret the Alexander polynomial in terms of the fundamental group and to explain what is available in more detail elsewhere.

The fundamental group of a space has, of course, already featured in the discussion of covering spaces in Chapter 7. *The group of a link* L in S^3 is defined to be $\Pi_1(S^3 - L)$, the fundamental group of the complement of L; this is the same as $\Pi_1(X)$, where X is the exterior of L. It is easy to write down a presentation of $\Pi_1(S^3 - L)$ from a given diagram of the link in the following way: Select an orientation of L just for convenience. Now, corresponding to the i^{th} segment

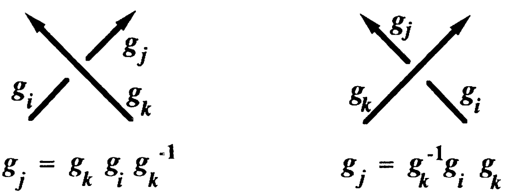


Figure 11.1

of the diagram with the usual breaks at under-passes (that is, an "over-passing" section of the link traversing, maybe, many under-passes or maybe none) take a group generator g_i. Corresponding to each crossing c, take a relator r_c as follows: Suppose at the crossing c the over-pass arc is labelled g_k and the under-pass is labelled g_i as it approaches c and g_j as it leaves c. Then $r_c = g_k g_i g_k^{-1} g_j^{-1}$ if the sign of the crossing is negative and $r_c = g_k^{-1} g_i g_k g_j^{-1}$ if the sign is positive. This is indicated in Figure 11.1. Each relator, when equated to the identity, asserts that the two generators corresponding to the under-passing arc are conjugate by means of the over-passing generator or its inverse (that choice being determined by the sign of the crossing).

The symbol g_i represents the loop that, starting from a base point (the eye of the reader) above the diagram, goes straight to the i^{th} over-passing arc, encircles it in a positive direction (to achieve linking number 1) and returns immediately to the base point. The resulting presentation is called the Wirtinger presentation of the link group. If there are m segments in the diagram and n crossings ($m = n$ unless some link component contains no under-pass), then the group of the link is isomorphic to the group G that has presentation

$$G = \langle g_1, g_2, \ldots, g_m;\ \ r_1, r_2, \ldots, r_n \rangle,$$

this meaning that G is the quotient of the free group on generators $\{g_1, g_2, \ldots, g_m\}$ by the smallest normal subgroup containing $\{r_1, r_2, \ldots, r_n\}$. A proof of this result follows from finding a suitable 2-complex that is a deformation retract of the link complement and using some algorithm for writing down a presentation of the fundamental group of a 2-complex. It is, of course, clear from the geometric interpretation of the g_i that the stated relators are indeed trivial elements of the group; the difficulty is in seeing that no more relators are required. In fact, for $n \geq 1$, at most $(n - 1)$ of the relators are actually needed, for it is easy to see that the product of certain conjugates of any $(n - 1)$ of the relators, in a suitable order, gives the remaining relator. That follows, for a connected diagram, from the fact that the dual graph in $\mathbb{R}^2 \cup \infty$ to the link projection has a four sided region containing each of the original crossings. The boundary of each such four-sided region gives one of the relators. The boundary of the union of $(n - 1)$ of these dual regions is the boundary of the n^{th} region.

It is clear that all the generators of a Wirtinger presentation that correspond to a single link component belong to the same conjugacy class in the group. Further, if the group is abelianised by adding in relations that assert that the g_i all commute with each other, then the group becomes just the direct sum of copies of $\mathbb{Z}$, one for each link component, with all the generators that correspond to a single link component becoming the generator of one of the $\mathbb{Z}$'s. As expected, this is the first homology group of $S^3 - L$; the loops representing generators of $\Pi_1(S^3 - L)$ in the above presentation also represent meridian generators of $H_1(S^3 - L)$.

The group of the unknot is, of course, infinite cyclic. As a simple non-trivial example, consider the trefoil knot 3_1 with the three generators allocated as in Figure 11.2. By the above remark, only two relators are needed, and the group of

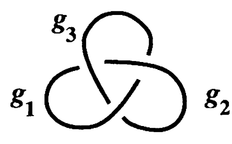


Figure 11.2

the trefoil knot is given by

$$G = \langle g_1, g_2, g_3;\ \ g_3 g_1 g_3^{-1} g_2^{-1},\ g_1 g_2 g_1^{-1} g_3^{-1} \rangle.$$

In this case a group homomorphism can be defined from G to Σ_3, the group of permutations of $\{1, 2, 3\}$, by

$$g_1 \mapsto (1, 2), \quad g_2 \mapsto (2, 3), \quad g_3 \mapsto (3, 1),$$

where as usual $(1, 2)$ is the permutation that interchanges 1 and 2 and fixes 3. That this does give a homomorphism follows from the observation that the two relators do indeed map to the trivial element of Σ_3. It is clear that this homomorphism is surjective; hence G is non-abelian and so certainly it is not cyclic. This proves that the trefoil knot is not the unknot. It is easy to verify that the group of 4_1, the 4-crossing knot, has no surjective homomorphism onto Σ_3. All the generators of a Wirtinger presentation of a *knot* group are conjugate, so any homomorphism will map them into a single conjugacy class. Of course, in a permutation group such a class is determined by the cycle type of a permutation. Homomorphisms can, then, be constructed to Σ_n by assigning permutations in some conjugacy class to the g_i and verifying that the relators map to the identity. This can be done in a systematic way with a computer, and a count can be made of all possible homomorphisms. The count for different knots can then be compared. Thistlethwaite has found such a method to be most effective for distinguishing knots from one another when compiling tables of knots with diagrams of up to fifteen crossings.

Two theorems, basic to the study of 3-manifolds, were proved by C. D. Papakyriakopoulos and published in 1957 ([105]). These are the Loop Theorem and the Sphere Theorem. They are both concerned with changing the assertion that a certain map of a surface into a 3-manifold exists to a statement that an embedding (an injective map) of the surface exists. The proofs are similar (see [43]) and employ the idea of lifting the map up to a succession of covering spaces until the self-intersections of the map can be reduced.

Theorem 11.1 (The Loop Theorem). *Let M be a (possibly non-compact) 3-manifold with boundary ∂M such that the inclusion-induced homomorphism $\Pi_1(\partial M) \to \Pi_1(M)$ is not injective. Then there exists a (piecewise linear) embedding of the disc $e : D^2 \to M$, with $e^{-1}(\partial M) = \partial D^2$ such that the restriction $e : \partial D^2 \to \partial M$ is not homotopic to a constant map*

There follows the application of this to knots; a version of this is sometimes known as Dehn's lemma.

Theorem 11.2. *Let X be the exterior of a knot K in S^3. If K is not the unknot, then the inclusion map induces an injection $\Pi_1(\partial X) \to \Pi_1(X)$.*

PROOF. Suppose $\Pi_1(\partial X) \to \Pi_1(X)$ is not injective. Then, by the loop theorem, there is an embedding $e : D^2 \to X$ sending ∂D^2 into the torus ∂X, to a simple closed curve not homotopically trivial in the torus. Now $e(\partial D^2)$ is certainly the boundary of the disc $e(D^2)$ and so represents a non-trivial element of the kernel of the map $H_1(\partial X) \to H_1(X)$; the longitude of the knot K (with either orientation) is the only simple closed curve representing an element in this kernel (see Definition 1.6). The longitude is parallel to K in a small solid torus neighbourhood of K, so expanding the disc $e(D^2)$ by an annulus gives a disc embedded in S^3 with K as its boundary. Thus $e(D^2)$ when so expanded is a Seifert surface for K. This shows that K is unknotted. □

Corollary 11.3. *A knot K is the unknot if and only if $\Pi_1(S^3 - K)$ is infinite cyclic.*

PROOF. If $\Pi_1(S^3 - K)$ is isomorphic to $\mathbb{Z}$, there can be no injection $\Pi_1(\partial X) \to \Pi_1(X)$ (as $\Pi_1(\partial X)$ is isomorphic to $\mathbb{Z} \oplus \mathbb{Z}$). □

Corollary 11.4. *Let X_1 and X_2 be the exteriors of two non-trivial knots and let M be a 3-manifold formed by identifying their boundaries together using any homeomorphism. Then the inclusion into M of the torus T that comes from the identified boundaries induces an injection $\Pi_1(T) \to \Pi_1(M)$.*

PROOF. This follows at once from the above theorem and from the Van Kampen theorem, which describes how fundamental groups behave when a space is described as a union of subspaces. □

Of course, as there are many invariants for showing that a knot is non-trivial, this corollary provides, if required, a source of orientable 3-manifolds containing tori for which the fundamental group injects. As stated in Definition 4.7, such tori are called incompressible. Thus if the exteriors of two non-trivial knots are glued together by some homeomorphism between their bounding tori, then the result is a 3-manifold, without boundary, containing an incompressible torus.

Theorem 11.5 (The Sphere Theorem). *Suppose that M is an orientable 3-manifold and that there exists a map $S^2 \to M$ that is not homotopic to a constant map (that is, $\Pi_2(M) \neq 0$). Then there exists a (piecewise linear) embedding $S^2 \to M$ that is not homotopic to a constant map.*

The theorem does not assert that the embedding is homotopic to the given map. A slightly stronger version of the theorem can be found in [43]. The application to knots (or to non-split links) are the following two results:

Theorem 11.6. *If K is a knot in S^3 any map $S^2 \to S^3 - K$ is homotopic to a constant map (that is, $\Pi_2(S^3 - K) = 0$).*

PROOF. If the statement is false then, by the sphere theorem, there exists a piecewise linear embedding $e : S^2 \to S^3 - K$ that is not homotopic to a constant in $(S^3 - K)$. Then, by the Schönflies theorem, $e(S^2)$ separates S^3 into two components, the closure of each of which is a ball with boundary $e(S^2)$. The knot K, being connected and disjoint from $e(S^2)$, lies in one of these balls, so e is homotopic to a constant using the other ball. □

Theorem 11.7. *If K is a knot in S^3, any map $S^r \to S^3 - K$ is homotopic to a constant map (that is, $\Pi_r(S^3 - K) = 0$) for all $r \geq 2$.*

PROOF. Let X be the exterior of K and let $\widetilde{X}$ be the universal cover of X. Thus $\widetilde{X}$ is the simply connected cover of X, it is acted upon by $\Pi_1(X)$, and the quotient of $\widetilde{X}$ by this action is X. The operation of lifting maps and homotopies from X to $\widetilde{X}$ shows that, for $r \geq 2$, $\Pi_r(X) = 0$ if and only if $\Pi_r(\widetilde{X}) = 0$ (or equivalently just use the homotopy long exact sequence of the covering). So certainly $\Pi_2(\widetilde{X}) = 0$. Now the third homology of any non-compact connected 3-manifold is zero. A simplicial argument for this uses the fact that any 3-cycle would be a finite sum of oriented 3-simplexes, a neighbourhood of the union of those 3-simplexes is a compact 3-manifold N with non-empty boundary which can be taken to be connected; any such N deformation retracts to a 2-dimensional complex (by collapsing 3-simplexes from the boundary), and so $H_3(N) = 0$. Of course, $\widetilde{X}$ is non-compact because $\Pi_1(X)$ is infinite (as $H_1(X)$ is infinite), and so each simplex has infinitely many different lifts in $\widetilde{X}$. Thus $H_3(\widetilde{X}) = 0$ and $H_r(\widetilde{X}) = 0$ for $r > 3$, as then X has no r-simplex and so its r^{th} chain group is zero. Now, for a simply connected cell complex, the Hurewicz isomorphism theorem asserts that the first non-vanishing homology group and the first non-vanishing homotopy group occur in the same dimension and are isomorphic. Thus $\Pi_r(\widetilde{X}) = 0$ for all r, and so $\widetilde{X}$ is a contractible space. The above remark about lifting ensures that $\Pi_r(X) = 0$ for $r \geq 2$. □

Another way of stating the last theorem is to say that $(S^3 - K)$ is an Eilenberg–MacLane space $\mathbf{K}(G, 1)$, where G is the knot group. An Eilenberg–MacLane space $\mathbf{K}(G, n)$ is a path-connected space that has homotopy group G in dimension n and all other homotopy groups zero. It is a routine task in homotopy theory to establish that two cell complexes that are both $\mathbf{K}(G, n)$'s are homotopy equivalent. Thus the group of a knot K determines the homotopy type of $(S^3 - K)$; any isomorphism between the groups of two knots is induced by some homotopy equivalence between the knot complements. In fact, this result and the given proof of it extend at once to a theorem stating that an irreducible 3-manifold with infinite fundamental group is determined up to homotopy type by that group.

Knots themselves (when not prime) may however not be determined by the homotopy types of their complements. Suppose X_1 and X_2 are the exteriors of oriented knots K_1 and K_2. Consider the knots $K_1 + K_2$ and $K_1 + r\overline{K_2}$, where as usual $r\overline{K_2}$ is the reverse of the reflection of K_2. The exterior of either these two composite knots is formed by identifying an annulus in the boundary of X_1 with an annulus in the boundary of X_2. The two identifications needed are homotopic (they differ by reversing the I factor in the annulus $S^1 \times I$), and so the two spaces obtained are homotopy equivalent. However, in general the two composite knots are distinct; if K_1 and K_2 are each the trefoil knot 3_1, the two composites are distinguished by the Jones polynomial.

Suppose that X_i is the exterior of an oriented knot K_i. On the boundary of X_i are the longitude λ_i and meridian μ_i, simple closed oriented curves that meet at a single point. They are well defined up to homotopy in ∂X_i. Taking $\lambda_i \cap \mu_i$ as a base point, let $[\lambda_i]$ and $[\mu_i]$ be the elements of $\Pi_1(\partial X_i)$ represented by these two curves. If K_1 and K_2 are equivalent oriented knots, there is a homeomorphism $h : X_1 \to X_2$ such that the following diagram commutes.

$$\begin{array}{ccccc} [\lambda_1],\ [\mu_1] & \in & \Pi_1(\partial X_1) & \longrightarrow & \Pi_1(S^3 - K_1) \\ \downarrow h_\star & & \downarrow h_\star & & \downarrow h_\star \\ [\lambda_2],\ [\mu_2] & \in & \Pi_1(\partial X_2) & \longrightarrow & \Pi_1(S^3 - K_2) \end{array}$$

This is immediate. The following converse is however also true. It is a consequence of the, somewhat lengthy, theory of homotopy-equivalent Haken manifolds created by F. Waldhausen *circa* 1966 ([132]). An account is also in [43].

Theorem 11.8. *If there exists an isomorphism from $\Pi_1(S^3 - K_1)$ to $\Pi_1(S^3 - K_2)$ which sends $[\lambda_1]$ to $[\lambda_2]$ and $[\mu_1]$ and $[\mu_2]$, then K_1 and K_2 are equivalent knots.*

Much more recently, the following has been proved by W. Whitten and F. Gonzales-Acuña [134].

Theorem 11.9. *If K_1 and K_2 are prime knots in S^3 and $\Pi_1(S^3 - K_1)$ and $\Pi_1(S^3 - K_2)$ are isomorphic groups, then $(S^3 - K_1)$ and $(S^3 - K_2)$ are homeomorphic spaces.*

Thus, for prime knots, the knot group determines the complement of the knot. It is by no means obvious that this means that the knots are the same. Perhaps the homeomorphism might send a meridian to a non-meridian. That this is not so is the substance of one of the most impressive results in knot theory of the 1980's. It is due to Gordon and J. Luecke [37] and the proof is lengthy and intricate:

Theorem 11.10. *If K_1 and K_2 are unoriented knots in S^3 and there is an orientation preserving homeomorphism between their complements, then K_1 and K_2 are equivalent (as unoriented knots).*

These results proclaim the importance of the group of a knot. It should, however, be observed that nothing as sophisticated as the last two results is needed to show that the knot group determines the Alexander polynomial of the knot. For suppose that a knot K has exterior X and group G. As has already been noted, $H_1(X)$ is the infinite cyclic group G/G' (a generator of which was previously called t), where G' is the commutator subgroup of G. The infinite cyclic cover X_∞ of X has its fundamental group equal to G' because a loop in X lifts to a *loop* in X_∞ if and only if it has zero linking number with K and so represents an element of G'. Now $H_1(X_\infty)$, the abelianisation of $\Pi_1(X_\infty)$, is G'/G'' where G'' is the commutator subgroup of G'. Group-theoretic conjugation gives an action of G on G and this passes to quotients to give an action of G/G' on G'/G''. This is the action of $H_1(X)$ on $H_1(X_\infty)$ that defines the latter as a $\mathbb{Z}[t, t^{-1}]$ module. Roughly, that is because conjugacy in G corresponds to moving a base point around a loop, and this operation lifts to the idea of translating X_∞ along the lift of that loop. The module $H_1(X_\infty)$ can thus be defined entirely in terms of G, and then the definition of the Alexander polynomial (up to a unit) can be given as before.

This means that starting with a presentation of G it should be possible to calculate the Alexander polynomial (it being understood that the abelianisation of G is known to be infinite cyclic). This can be done in the following way, using the *free differential calculus* devised by R. H. Fox. Suppose that G is the group of a knot K, given by any presentation

$$G = \langle x_1, x_2, \ldots, x_n; r_1, r_2, \ldots, r_m \rangle,$$

and let $\alpha : G \to G/G' \cong \langle t \rangle$ be the abelianisation homomorphism. If P is any space with $\Pi_1(P) = G$, and $\widetilde{P}$ is the cover of P corresponding to G', then the above reasoning transferred from X to P shows that $H_1(\widetilde{P}; \mathbb{Z})$, regarded as a module over $\mathbb{Z}[t^{-1}, t]$, is also G'/G'' with action by $\mathbb{Z}[G/G']$, and so it is equivalent to the Alexander module of K. Take for P a complex consisting of one 0-cell V, n oriented 1-cells labelled $x_1, x_2, \ldots, x_n$, having all their end points identified with V to form n loops, and m oriented 2-cells $c_1, c_2, \ldots, c_m$, with each ∂c_i glued to the 1-cells according to the word r_i. All the lifts to $\widetilde{P}$ of all the cells of P give a cell structure on $\widetilde{P}$ which can be used in the following way to investigate the homology of $\widetilde{P}$. Let $\widetilde{V}$ be a chosen lift of the point V, let $\widetilde{x}_i$ be the lift of x_i that starts at $\widetilde{V}$ and let $\widetilde{c}_i$ be the lift of c_i that has as its boundary the lift of r_i that begins at $\widetilde{V}$. The whole of $\widetilde{P}$ is just the union of all translates of these cells under the action of $\langle t \rangle$. Thus the chain complex (with integer coefficients) of $\mathbb{Z}[t^{-1}, t]$ modules for $\widetilde{P}$,

$$C_2(\widetilde{P}) \xrightarrow{d_2} C_1(\widetilde{P}) \xrightarrow{d_1} C_0(\widetilde{P}),$$

has each $C_i(\widetilde{P})$ freely generated as a module by the above specified i-cells in $\widetilde{P}$.

In this chain complex, the boundary map d_2 sends $\widetilde{c}_i$ to the lift of r_i beginning at $\widetilde{V}$ now regarded as an element of the module $C_1(\widetilde{P})$. Any occurrence of x_j in r_i contributes *some* $\langle t \rangle$-translate of $\widetilde{x}_j$ to $d_2(\widetilde{c}_i)$. In fact, if $r_i = w_1 x_j w_2$ this occurrence of x_j contributes to $d_2(\widetilde{c}_i)$ the lift of x_j that begins at the final point of the lift of w_1 which starts at $\widetilde{V}$; thus the contribution is $\alpha(w_1)$ acting on $\widetilde{x}_j$. If $r_i = w_1 x_j^{-1} w_2$, this occurrence of x_j^{-1} contributes $-\alpha(w_1 x_j^{-1})\widetilde{x}_j$. The $\widetilde{x}_j$

term in $d_2(\widetilde{c}_i)$ is thus the sum, over all occurrences of x_j^{-1} and x_j in r_i, of these contributions. It is a simple formality to write down a procedure to determine this sum as

$$d_2(\widetilde{c}_i) = \sum_j \alpha\phi(\frac{\partial r_i}{\partial x_j})\widetilde{x}_j,$$

where the meaning of the terms is as follows: The quotient map (given by the presentation) from F, the free group on generators $x_1, x_2, \ldots x_n$, to G induces a homomorphism of group-rings $\phi : \mathbb{Z}(F) \to \mathbb{Z}(G)$. The map $\frac{\partial}{\partial x_j} : \mathbb{Z}(F) \to \mathbb{Z}(F)$ is the linear extension of the map defined on elements of F by

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij}, \qquad \frac{\partial x_i{}^{-1}}{\partial x_j} = -\delta_{ij} x_i{}^{-1},$$

$$\frac{\partial (uv)}{\partial x_j} = \frac{\partial u}{\partial x_j} + u\frac{\partial v}{\partial x_j}.$$

(In practice, this last formula should be used on a word in which v is the last letter of the word.) Thus the transpose of $\alpha\phi(\frac{\partial r_i}{\partial x_j})$ is a matrix representing d_2, and so that is a presentation matrix for the module $C_1(\widetilde{P})/d_2(C_2(\widetilde{P}))$. Now, as usual, there is a short exact sequence of modules

$$0 \longrightarrow H_1(\widetilde{P}) \longrightarrow C_1(\widetilde{P})/d_2(C_2(\widetilde{P})) \xrightarrow{d_1} d_1C_1(\widetilde{P}) \longrightarrow 0.$$

So it is useful to investigate $d_1C_1(\widetilde{P})$. The boundary map d_1 is determined by $d_1(\widetilde{x}_j) = (\alpha\phi(x_j) - 1)\widehat{V}$. Now, $\alpha\phi(x_j) = t^{a_j}$ for some $a_j \in \mathbb{Z}$, so $d_1C_1(\widetilde{P})$ is $\mathcal{I}\widehat{V}$ where $\mathcal{I}$ is the ideal of $\mathbb{Z}[t, t^{-1}]$ generated by $\{(t^{a_j} - 1) : j = 1, 2, \ldots, n\}$. However, observe that

$$(t^a - 1) + t^a(t^b - 1) = (t^{a+b} - 1),$$

$$(t^a - 1) - t^{a-b}(t^b - 1) = (t^{a-b} - 1).$$

As the t^{a_j} generate $\langle\ t\ \rangle$, it must be that $t = t^{\sum \nu_j a_j}$ for some $\nu_j \in \mathbb{Z}$. Thus, using the above observation, $(t - 1) \in \mathcal{I}$ and, as $(t - 1)$ divides each $(t^{a_j} - 1)$, it follows that $\mathcal{I}$ is just the principal ideal generated by $(t - 1)$. Then $d_1C_1(\widetilde{P})$ is the free rank-one module generated by the element $(t - 1)\widehat{V}$. As $d_1C_1(\widetilde{P})$ is free, the above short exact sequence splits and

$$C_1(\widetilde{P})/d_2(C_2(\widetilde{P})) \cong H_1(\widetilde{P}) \oplus \mathbb{Z}[t, t^{-1}].$$

Now if A is a presentation matrix for a module M over a ring R, a presentation matrix for $M \oplus R$ is A with an extra row of zeros appended. This has the same non-zero minors as has A, but the number of rows deleted to obtain a minor of the new matrix is one more than the number so deleted from A. Thus the r^{th} elementary ideal of M is the $(r+1)^{\text{th}}$ elementary ideal of $M \oplus R$. The transpose of the $m \times n$ matrix $\alpha\phi(\partial r_i/\partial x_j)$ presents $H_1(\widetilde{P}) \oplus \mathbb{Z}[t, t^{-1}]$, so the r^{th} elementary ideal $\mathcal{E}_r$

of $\mathbb{Z}[t, t^{-1}]$ for the module $H_1(\widetilde{P})$ is generated by the $(n-r) \times (n-r)$ minors of this matrix. In particular $\mathcal{E}_1$ is generated by the $(n-1) \times (n-1)$ minors, and this is the ideal that is known (from Chapter 6) to be a principal ideal generated by the Alexander polynomial $\Delta_K(t)$.

As a simple example of the way that this can be used, consider the Wirtinger presentation of the trefoil knot given by

$$G = \langle g_1, g_2, g_3;\ g_3 g_1 g_3^{-1} g_2^{-1},\ g_1 g_2 g_1^{-1} g_3^{-1} \rangle.$$

The formalism of the free differential calculus gives

$$\frac{\partial r_i}{\partial g_j} = \begin{pmatrix} g_3 & -g_3 g_1 g_3^{-1} g_2^{-1} & 1 - g_3 g_1 g_3^{-1} \\ 1 - g_1 g_2 g_1^{-1} & g_1 & -g_1 g_2 g_1^{-1} g_3^{-1} \end{pmatrix}.$$

On abelianising, each generator of the Wirtinger presentation is mapped to t, so

$$\alpha\phi\left(\frac{\partial r_i}{\partial g_j}\right) = \begin{pmatrix} t & -1 & 1-t \\ 1-t & t & -1 \end{pmatrix}.$$

Up to sign, all the (2×2) minors of this are $1 - t + t^2$, and so this is the Alexander polynomial of the trefoil knot. This general method of calculating the Alexander polynomial, when applied to a Wirtinger presentation, is essentially the "L-matrix" method of Reidemeister ([107] or [108]).

The free differential calculus excels in the case of a torus knot. Let T be a standard unknotted torus in S^3 (T can be thought of as the boundary of a neighbourhood of the unknot). A (p, q) torus knot is the knot K contained in T that represents p longitudes and q meridians of the unknot. Such a simple closed curve exists if and only if p and q are coprime (an exercise). The exterior of K consists of two solid tori (one inside T and one outside) glued together along an annulus that goes around T as p longitudes and q meridians. The Van Kampen theorem, which describes the fundamental group of a space obtained by gluing two other spaces together, can then be used. In this case it shows that the group of K has a one relator presentation as

$$\langle\, x_1, x_2;\ x_1^p x_2^{-q} \,\rangle,$$

where x_1 and x_2 are represented by cores of the two solid tori. The relation occurs because the core of the gluing annulus represents both x_1^p and x_2^q. Note that x_1 links K with linking number q and x_2 links K with linking number p. Hence $\alpha x_1 = t^q$ and $\alpha x_2 = t^p$. Thus

$$\frac{\partial r_1}{\partial x_j} = (\,1 + x_1 + x_1^2 + \cdots + x_1^{p-1} \quad x_1^p(-x_2^{-1} - x_2^{-2} - \cdots - x_2^{-q})\,),$$

$$\alpha\phi\left(\frac{\partial r_1}{\partial x_j}\right) = \begin{pmatrix} \dfrac{1 - t^{pq}}{1 - t^q} & \dfrac{-t^{pq}t^{-p}(1 - t^{-pq})}{1 - t^{-p}} \end{pmatrix}.$$

So the Alexander polynomial of K is a generator of the (principal) ideal of $\mathbb{Z}[t, t^{-1}]$ generated by the two elements $1 - t^{pq}/1 - t^q$ and $1 - t^{pq}/1 - t^p$. As p and q

are coprime, the technique used above shows that $(1 - t)$ is in the ideal generated by $(1 - t^p)$ and $(1 - t^q)$. Then it is not hard to see that a highest common factor of $1 - t^{pq}/1 - t^q$ and $1 - t^{pq}/1 - t^p$ is

$$\frac{(1 - t)(1 - t^{pq})}{(1 - t^p)(1 - t^q)},$$

and so this is (up to multiplication by $\pm t^{\pm n}$) the Alexander polynomial of the torus knot.

This discussion of the torus knot depends for its simplicity on the fact that the group of a torus knot has a presentation with just two generators and one relator. Any 2-bridge link also has a two-generator, one-relator presentation, as can be seen from its description as a union of two trivial 2-string tangles. In general, this relator is more complicated than that for the torus knot.

The Alexander polynomial of an oriented link is (up to units) a Laurent polynomial in one variable t. If $\#L$, the number of components of L, is two or more, the theory can be amplified to give a multi-variable Alexander polynomial. Suppose that $l : \{1, 2, \dots, \#L\} \to \{1, 2, \dots, \nu\}$, for some integer $\nu \geq 2$, is a surjection, thought of as labelling (or colouring) of the components $\{L_i : i = 1, 2, \dots, \#L\}$ of L. Let G be, as usual, the group of L. Then G/G' is a free abelian group on $\#L$ meridian generators. Map this on to the free abelian group on ν generators (written multiplicatively as $\langle t_1, t_2, \dots, t_\nu ; t_i t_j t_i^{-1} t_j^{-1} \rangle$) by sending the i^{th} oriented meridian to $t_{l(i)}$, and let α be the composition

$$\alpha : G \to G/G' \to \langle t_1, t_2, \dots, t_\nu ; t_i t_j t_i^{-1} t_j^{-1} \rangle.$$

Then X, the exterior of L, has a covering $\widehat{X}$ corresponding to the kernel of α that is acted upon freely by the group $\langle t_1, t_2, \dots, t_\nu ; t_i t_j t_i^{-1} t_j^{-1} \rangle$. The group ring of this group will be written as $\mathbb{Z}[t_1^{\pm 1}, t_2^{\pm 1}, \dots, t_\nu^{\pm 1}]$, and $H_1(\widehat{X}; \mathbb{Z})$ is a module over this ring. This module is an invariant of the oriented labelled link L. Again, the first elementary ideal of the module is principal, and a generator (well defined up to multiplication by $\pm t_1^{\pm m_1} t_2^{\pm m_2} \cdots t_\nu^{\pm m_\nu}$) is called the multi-variable Alexander polynomial of the oriented labelled link. In [21] a method is given for finding a square matrix that presents the module $H_1(\widehat{X}; \mathbb{Z})$; it is a generalisation of the Seifert surface method of Chapter 6 to the multi-variable situation. Alternatively, one can follow the formalism of the free differential calculus discussed in this chapter. Starting with a Wirtinger presentation of G as $\langle x_1, x_2, \dots, x_n ; r_1, r_2, \dots, r_n \rangle$, form the cell complex P from the presentation as before, and let $\widetilde{P}$ be the cover corresponding to the kernel of α. As before, the transpose of the square matrix $\alpha\phi(\partial r_i / \partial x_j)$ is a presentation matrix for the module $C_1(\widetilde{P})/d_2(C_2(\widetilde{P}))$. Now, however, the module $d_1 C_1(\widetilde{P})$ is not free; it is isomorphic to the ideal of $\mathbb{Z}[t_1^{\pm 1}, t_2^{\pm 1}, \dots, t_\nu^{\pm 1}]$ generated by $\{(t_i - 1) : i = 1, 2, \dots, \nu\}$. Thus the short exact sequence relating these two modules and $H_1(\widetilde{P})$ cannot be split. Nevertheless, it can be shown that a multi-variable Alexander polynomial is obtained by taking the matrix $\alpha\phi(\partial r_i / \partial x_j)$, deleting any row and the j^{th} column, evaluating the determinant of this smaller matrix, and then dividing by $(\alpha\phi(x_j) - 1)$, which will indeed be a factor. More details are in [17]. This has been refined in [39] in order to obtain a canonical normalisation for multi-variable polynomials, which fits into the a generalised skein approach described by Conway [20] (see also [100] and [99]).

A final example of the direct use of the fundamental group of a knot's exterior will now be given. It is a proof found by H. F. Trotter [125] that there are oriented knots that are not equivalent to their reverses. It is included here because the result seems important and because no *invariant* has been found that can ever prove a knot to be a non-reversible knot (though see [70]). Most known methods of achieving such a result are rather *ad hoc*. In this case the technique consists of a detailed investigation of a particular group, understanding it in terms of isometries of the hyperbolic plane. Trotter's result is the following:

Theorem 11.11. *Let p, q and r be odd integers such that $|p|$, $|q|$ and $|r|$ are distinct and greater than one. Then the oriented pretzel knot $P(p, q, r)$ is not equivalent to its reverse.*

SKETCH PROOF. Suppose that $p = 2k + 1$, $q = 2l + 1$ and $r = 2m + 1$. Many of the generators of the Wirtinger presentation of the group of $P(p, q, r)$ can easily be eliminated to give a presentation with meridian generators x, y and z as indicated (for $(p, q, r) = (7, 3, 5)$) in Figure 11.3 and relations

$$\begin{aligned}(xy^{-1})^m x(xy^{-1})^{-m} &= (yz^{-1})^{k+1} z(yz^{-1})^{-k-1},\\ (yz^{-1})^k y(yz^{-1})^{-k} &= (zx^{-1})^{l+1} x(zx^{-1})^{-l-1},\\ (zx^{-1})^l z(zx^{-1})^{-l} &= (xy^{-1})^{m+1} y(xy^{-1})^{-m-1}.\end{aligned}$$

A longitude w represents the element

$$(xy^{-1})^{-m}(yz^{-1})^{k+1}(zx^{-1})^{-l}(xy^{-1})^{m+1}(yz^{-1})^{-k}(zx^{-1})^{l+1}.$$

If $P(p, q, r)$ is reversible, there exists an automorphism α of the group that sends meridians to inverse meridians and w to w^{-1}. (Here a meridian is an element of the group represented by a loop that goes from the base point along some path to the knot, around the knot and back along the same path.) Thus if H is the subgroup generated by the squares of meridians, H is normal and invariant under α. Then α induces an automorphism of G/H, and G/H has a presentaion with generators

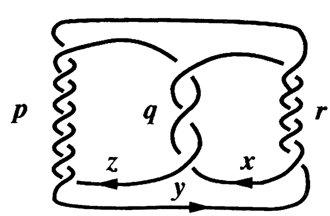


Figure 11.3

x, y and z, with the above three relations and the relations

$$x^2 = y^2 = z^2 = 1.$$

That easily simplifies to become

$$G/H = \langle\, x, y, z\, ;\ x^2 = y^2 = z^2 = 1\, ,\, (xy)^r = (yz)^p = (zx)^q\, \rangle.$$

Now, abelianising G/H produces a cyclic group of order 2 (as p, q and r are odd) with x, y and z all mapping to the generator. Thus the commutator subgroup of G/H is all elements expressible as words of even length in x, y and z. It is thus the subgroup generated by xy, yz and zx (note that $(xy)(yz) = xz$), and each of these three elements commutes with $(xy)^r$. Thus if U is the subgroup of G/H generated by $(xy)^r$, then U is contained in the centre of the commutator subgroup of G/H.

Let W be the quotient $(G/H)/U$ so that W is

$$\langle\, x, y, z\, ;\ x^2 = y^2 = z^2 = (xy)^r = (yz)^p = (zx)^q = 1\, \rangle.$$

This is well known to be a triangle subgroup of isometries of the hyperbolic plane, where x, y and z now represent the reflections in the three sides of a hyperbolic triangle having vertex angles π/p, π/q and π/r. This means that xy, yz and zx represent rotations about the three vertices. (It may be assumed, without effecting the algebra, that p, q and r are now all positive.) It is easy to see that the subgroup generated by those rotations has trivial centre. Hence U *is* the centre of the commutator subgroup of G/H, and so U is mapped to itself by any automorphism of G/H. Hence α induces an automorphism $\overline{\alpha}$ of W. In W, the longitude element w has become $((xy)^{-m}(yz)^{-k}(zx)^{-l})^2$, which represents a translation of the hyperbolic plane along the direction of one of the sides of the triangle. Then $\overline{\alpha}$ must send this element to its inverse, the translation in the opposite direction. However, a little consideration of the triangular tessellation of the hyperbolic plane [125] shows that this would mean that the direction of the side of the triangle would be reversed by some element of W, and this is not possible when p, q and r are all odd. □

Exercises

1. Find a presentation of the group of the knot 8_2 with the minimum possible number of generators and the minimum possible number of relators.

2. Use the loop theorem (or Dehn's Lemma) to show that a non-trivial knot K and a longitude of K never constitute a split link.

3. Use the Loop Theorem and the Schönflies theorem to show that any torus T, piecewise linearly embedded in S^3, bounds, on one side in S^3, a solid torus. [It should be assumed that $S^3 - T$ has two components.]

4. For a given positive integer n find a knot for which the group has no presentation with fewer than n generators.

5. Explore the way in which the free differential calculus, applied to the Wirtinger presentation of a knot group, provides a way of deriving the Alexander polynomial of a knot from the determinant of a matrix directly associated with a diagram of the knot.

6. A knot K has *tunnel number* 1 if an arc α can be embedded (piecewise linearly) in S^3, meeting K precisely in $\partial\alpha$, so that the closure of the complement of a regular neighbourhood of the θ-curve $K \cup \alpha$ is a handlebody of genus 2 (see Chapter 12 for discussion of handlebodies). Show that a 2-bridge knot has tunnel number 1 and so does a torus knot. Prove that the group of a tunnel number 1 knot has a presentation with two generators and one relator. What does that imply about the Alexander ideals of the knot? Prove that the pretzel knot $P(3, 3, -3)$ does not have tunnel number 1.

7. The dihedral group D_{2n} of $2n$ elements is the group of symmetries of a regular n-gon; it has a presentation $\langle x, y\,;\, x^n, y^2, yxyx\rangle$. Suppose that there is given an n-colouring of a diagram of a knot K. This is a function c from the segments of the diagram to $\mathbb{Z}/n\mathbb{Z}$ so that at any crossing, the over-pass is labelled with the average, modulo n, of the labels of the two segments on either side. If g_i is the generator of the Wirtinger presentation of $\Pi_1(S^3 - K)$ corresponding to the ith segment of the diagram, show that $g_i \mapsto yx^{c(i)}$ defines a homomorphism $\Pi_1(S^3 - K) \to D_{2n}$. Show that any surjective homomorphism $\Pi_1(S^3 - K) \to D_{2n}$ must arise in this way. [When n is odd, such a surjection exists if and only if n divides the exponent (the lowest common multiple of the orders of the elements) of the first homology group of the double cover of S^3 branched over K. A necessary condition for the existence of an n-colouring is that n divide the determinant of K.]

8. Show that the genus of the (p, q) torus knot, where p and q are coprime, is $\frac{1}{2}(p - 1)(q - 1)$.

9. Suppose that X is the exterior of a knot K. The 3-manifold $(S^1 \times D^2) \cup_h X$, where $h : \partial(S^1 \times D^2) \to \partial X$ is a homeomorphism with $h(\text{point} \times \partial D^2)$ homologous to the sum of α meridians and β longitudes, is said to be obtained by α/β Dehn surgery on K. Show that if α/β Dehn surgery on a torus knot produces a simply connected manifold, then $(\alpha, \beta) = (\pm 1, 0)$ and the manifold produced is just S^3. [A knot with this property is said to have "Property P"; it is not known if all knots have this property.]

10. Prove that the trefoil knot 3_1 and its reflection are distinct by showing that there is no isomorphism, from the group of one knot to the group of the other, that maps the elements corresponding to meridian and and longitude in one group to those corresponding to meridian and longitude in the other.

12

Obtaining 3-Manifolds by Surgery on S^3

The aim of this chapter is to show, in Theorem 12.14, that every closed connected orientable 3-manifold can be obtained by "surgery" on S^3. The method used is a version of that of [77]. An elementary *r-surgery* on a general n-manifold M is the operation of removing from M an embedded copy of $S^r \times D^{n-r}$ and replacing it with a copy of $D^{r+1} \times S^{n-r-1}$, the replacement being effected by means of the obvious homeomorphism between the boundaries of the removed set and its replacement. Surgery in general is a sequence of elementary surgeries. In the case of surfaces, instances of 1-surgery and 0-surgery have already been employed in earlier chapters, usually when the surface was contained in S^3. The only surgeries needed in this chapter are 1-surgeries on a 3-manifold, and it is easy to see they can be performed "simultaneously". The surgery process will consist of the removal from S^3 of disjoint copies of $S^1 \times D^2$ and their replacement by copies of $D^2 \times S^1$. Of course, the set removed and its replacement are homeomorphic, but the parametrisation of the removed set as disjoint copies of $S^1 \times D^2$, and the canonical method of replacement with respect to that, ensure that the new manifold is usually not S^3. A collection of disjoint solid tori in S^3 is just a regular neighbourhood of a link, and a parametrisation of a neighbourhood of each component by $S^1 \times D^2$ is called a *framing* of the link. Thus it will be shown that 3-manifolds can be interpreted by means of framed links in S^3.

The fact that any 3-manifold M is triangulable, and so can be regarded as a simplicial complex, will be assumed. It is hoped that piecewise linearity, though assumed throughout, will not be obtrusive. When M is closed (that is, compact and with empty boundary) and orientable, a triangulation will lead easily to the fact that M has a Heegaard splitting. This will mean that M is just two "handle-bodies" (see Definition 12.10) with their boundary surfaces identified by some homeomorphism between them. Philosophically, complete knowledge of surface homeomorphisms should tell all about 3-manifolds. Thus a little investigation of surface homeomorphisms is in order.

Firstly, it is desirable to divide homeomorphisms into *isotopy classes*. As already mentioned in Chapter 1, homeomorphisms are isotopic if one can be "slid" to the other. The definition of Chapter 1 is amplified below. If two homeomorphisms between surfaces do not differ significantly, one would not expect much difference between 3-manifolds formed by operations using those homeomorphisms.

Definition 12.1. Piecewise linear homeomorpisms h_0 and h_1 between complexes X and Y are isotopic if they are connected by a path of homeomorphisms $\{h_t : X \to Y,\ t \in [0, 1]\}$ such that the map $H : X \times [0, 1] \to Y \times [0, 1]$ defined by $H(x, t) = (h_t(x), t)$ is a piecewise linear homeomorphism.

If preferred, "smooth" could be substituted for "piecewise linear" in the above definition when X is a smooth manifold. However, it is important that the homeomorphism H should indeed belong in the category of choice. A classical result of Alexander ([113], [47]) states that any piecewise linear homeomorphism of the n-dimensional ball to itself, that is fixed on the boundary, is isotopic to the identity keeping the boundary fixed (by all the h_t). This leads to the result that any piecewise linear orientation-preserving homeomorphism of the n-sphere to itself is isotopic to the identity. (Although the smooth versions of these results are, in general, false, they are true when $n = 2$.) For surfaces it is, in fact, known that homotopic homeomorphisms are isotopic. It is easy to show that for any complex X, the set of all self-homeomorphisms that are isotopic to the identity forms a normal subgroup of the group of all self-homeomorphisms of X. The quotient of the group of all self-homeomorphisms by this normal subgroup is called the *mapping class group* of X. The present motivation for thinking about isotopy comes from the following elementary lemma.

Lemma 12.2. *Suppose that U and V are 3-manifolds with homeomorphic boundaries, and that $h_0 : \partial U \to \partial V$ and $h_1 : \partial U \to \partial V$ are isotopic homeomorphisms. Then $U \cup_{h_0} V$ and $U \cup_{h_1} V$ are homeomorphic.*

PROOF. Choose ([113], [47]) a collar neighbourhood C of ∂U in U; C is a neighbourhood of ∂U homeomorphic to $\partial U \times [0, 1]$, with ∂U identified with $\partial U \times 0$. A homeomorphism $f : U \cup_{h_0} V \to U \cup_{h_1} V$ can be constructed by defining f to be the identity on $(U - C) \cup V$ and on C defining $f(x, t) = (h_1^{-1} h_t x, t)$. □

In what follows, let F be a connected compact oriented surface, possibly with non-empty boundary. Let C be a simple closed curve embedded in F, and let A be an annulus neighbourhood of C. The standard annulus is $S^1 \times [0, 1]$ with some fixed orientation.

Definition 12.3. A twist about C is any homeomorphism isotopic to the homeomorphism $\tau : F \to F$ defined such that $\tau|F - A$ is the identity and, parametrising A as $S^1 \times [0, 1]$ in an orientation-preserving manner, $\tau|A$ is given by $\tau(e^{i\theta}, t) = (e^{i(\theta - 2\pi t)}, t)$.

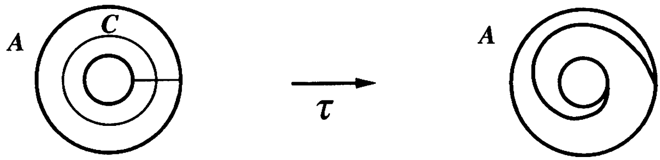


Figure 12.1

Note that the effect of τ on a path crossing C is to sweep that path all the way around the annulus. See Figure 12.1. Strictly, of course, a twist homeomorphism should here be piecewise linear; the fourth power of the piecewise linear homeomorphism shown in Figure 12.2 (which fixes the inner boundary component and moves each vertex on the outer boundary to the next vertex in a clockwise direction) is an appropriate piecewise linear model for a twist rather than the homeomorphism of Figure 12.1.

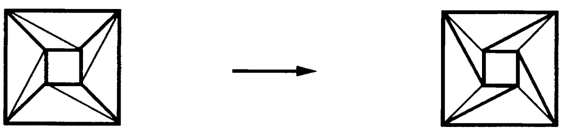


Figure 12.2

Definition 12.4. Oriented simple closed curves p and q contained in the interior of the surface F are called twist-equivalent, written $p \sim_\tau q$, if $hp = q$ for some homeomorphism h of F that is in the group of homeomorphisms generated by all twists of F (which includes homeomorphisms isotopic to the identity).

In this definition h is required to carry the orientation of one curve to that of the other. Of course, in general there may be no homeomorphism of any sort that sends p to q; that is certainly the case if p separates F and q does not.

Lemma 12.5. *Suppose oriented simple closed curves p and q, contained in the interior of the surface F, intersect transversely at precisely one point. Then $p \sim_\tau q$.*

PROOF. The first diagram of Figures 12.3 shows the intersection point of p and q and also a simple closed curve C_1 that runs parallel to, and is slightly displaced from, q. Similarly, C_2 is a slightly displaced copy of p. The second diagram shows $\tau_1 p$, where τ_1 is a twist about C_1. The third diagram shows $\tau_2\tau_1 p$, where τ_2 is a twist about C_2. In this diagram $\tau_2\tau_1 p$ has a doubled-back portion that can easily be moved by a homeomorphism isotopic to the identity (that is, a slide in F) to change $\tau_2\tau_1 p$ to q. □

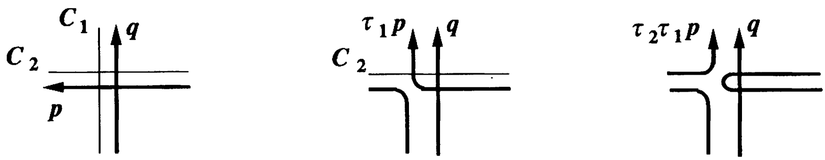


Figure 12.3

Lemma 12.6. *Suppose that oriented simple closed curves p and q contained in the interior of the surface F are disjoint and that neither separates F (that is,* $[p] \neq 0 \neq [q]$ *in* $H_1(F, \partial F)$*). Then* $p \sim_\tau q$.

PROOF. Consideration of the surface obtained by cutting F along $p \cup q$ shows at once that there is a simple closed curve r in F that intersects each of p and q transversely at one point. Then, by Lemma 12.5 , $p \sim_\tau r \sim_\tau q$. □

Proposition 12.7. *Suppose that oriented simple closed curves p and q are contained in the interior of the surface F and that neither separates F. Then* $p \sim_\tau q$.

PROOF. Changing q by means of a homeomorphism of F that is (close to and) isotopic to the identity, it can be assumed that p and q intersect transversely at n points. The proof is by induction on n; Lemmas 12.5 and 12.6 start the induction, so assume that $n \geq 2$ and that the result is true for less that n points of intersection.

Let A and B be consecutive points along p of $p \cap q$. Suppose firstly that p leaves A on one side of q and returns to B from the other side of q. Let r be a simple closed curve in F that starts near A, follows close to p until near B and then returns to its start in a neighbourhood of q. As shown in the first diagram of Figure 12.4, r can be chosen so that $p \cap r$ contains less than n points and $q \cap r$ is one point. Hence $p \sim_\tau r$ by the induction hypothesis, and $r \sim_\tau q$ by Lemma 12.5.

Suppose now that p leaves A on one side of q and returns to B from the *same* side of q. Let r_1 and r_2 be the two simple closed curves shown in the second

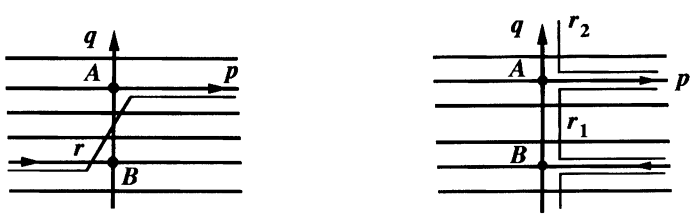


Figure 12.4

diagram of Figure 12.4. Each starts near A, proceeds near p until close to B and then back to its start following near to q. However, r_1 starts on the right of p and r_2 starts on the left. Now in $H_1(F, \partial F)$, $[r_1] - [r_2] = [q]$, and hence at least one of r_1 and r_2 does not separate (as $[q] \neq 0$). Let that curve be defined to be r. Then r is disjoint from q, so $r \sim_\tau q$ by Lemma 12.6 and, as $r \cap p$ has at most $n-2$ points, $p \sim_\tau r$ by the induction hypothesis. □

Corollary 12.8. *Let $p_1, p_2, \ldots, p_n$ be disjoint simple closed curves in the interior of F the union of which does not separate F. Let $q_1, q_2, \ldots, q_n$ be another set of curves with the same properties. Then there is a homeomorphism h of F that is in the group generated by twists, so that $hp_i = q_i$ for each $i = 1, 2, \ldots, n$.*

PROOF. Suppose inductively that such an h can be found so that $hp_i = q_i$ for each $i = 1, 2, \ldots, n-1$. Apply Proposition 12.7 to hp_n and q_n in F cut along $q_1 \cup q_2 \cup \ldots \cup q_{n-1}$. □

The theory of homeomorphisms of surfaces will be left at that point and attention turned back to n-manifolds with particular interest in $n = 3$.

Definition 12.9. Let M be an n-manifold, let $e : \partial D^r \times D^{n-r} \to \partial M$ be an embedding (where, as usual, D^s is the standard s-dimensional disc or ball). Then $M \cup_e (D^r \times D^{n-r})$ is called "M with an r-handle added".

Note that the boundary of this new manifold is ∂M changed by an $(r-1)$-surgery.

Definition 12.10. A handlebody of genus g is an orientable 3-manifold that is a 3-ball with g 1-handles added.

Here, "orientable" can be taken to mean that every simple closed curve in the manifold has a solid torus neighbourhood. It is a straightforward exercise in the elementary technicalities of piecewise linear manifold theory to show that, up to homeomorphism, there is only one genus g handlebody. It is indeed, as already stated, the product of an interval with a g-holed disc. A regular neighbourhood of any finite connected graph embedded in an orientable 3-manifold is a handlebody. This follows by taking the neighbourhood of a maximal tree as the 3-ball and neighbourhoods of the midpoints of the remaining edges as 1-handles.

Definition 12.11. A Heegaard splitting of a (closed, connected, orientable) 3-manifold M is a pair of handlebodies X and Y contained in M such that $X \cup Y = M$ and $X \cap Y = \partial X = \partial Y$.

Note that X and Y have the same genus; namely, the genus of their common boundary surface.

Lemma 12.12. *Any closed connected orientable 3-manifold has a Heegaard splitting.*

PROOF. This is similar to the first part of the proof of Theorem 8.2. Take a triangulation of M as a simplicial complex K. The vertices of the first derived subdivision $K^{(1)}$ of K are the barycentres $\widehat{A}$ of the simplexes A of K. The second derived subdivision $K^{(2)}$ of K is, of course, just $(K^{(1)})^{(1)}$. The 1-skeleton of K (that is, the sub-complex consisting of the 0-simplexes and 1-simplexes of K), being a graph, has, as intimated above, for its simplicial neighbourhood in $K^{(2)}$, a handlebody. The closure of the complement of this is the simplicial neighbourhood in $K^{(2)}$ of another graph. That graph, called the dual 1-skeleton of K, is the sub-complex $\bigcup_A C_A$ of $K^{(1)}$, where the union is over all 3-simplexes A, and C_A is the cone with vertex $\widehat{A}$ on the barycentres of the 2-dimensional faces of A. Thus $K^{(2)}$ is expressed as the union of two handlebodies that intersect in their common boundary, and this is the required Heegaard splitting. □

Theorem 12.13. *Let M be a closed connected orientable 3-manifold. There exists finite sets of disjoint solid tori $T_1', T_2', \ldots, T_N'$ in M and $T_1, T_2, \ldots, T_N$ in S^3 such that $M - \bigcup_1^N \text{Int}(T_i')$ and $S^3 - \bigcup_1^N \text{Int}(T_i)$ are homeomorphic.*

PROOF. By Lemma 12.12, M has a Heegaard splitting, so for handlebodies U and V of some genus g, and some homeomorphism $h : \partial U \to \partial V$, $M = U \cup_h V$. Let $p_1', p_2', \ldots, p_g'$ be disjoint simple closed curves in ∂U, that bound disjoint discs in U and let $q_1, q_2, \ldots q_g$ be disjoint simple closed curves in ∂V (one around each "hole" of the handlebody) as shown in Figure 12.5, so that if ϕ is any homeomorphism $\phi : \partial U \to \partial V$ such that $\phi(p_i') = q_i$ for each i, then $U \cup_\phi V = S^3$.

Figure 12.5

Let $h(p_i') = p_i$ for each i. *If* there were a homeomorphism of V sending each p_i to q_i then $U \cup_h V$ would be S^3. However, by Corollary 12.8, there is a product ψ, of twists and inverses of twists, of ∂V that sends each p_i to q_i. Up to isotopy a twist τ of ∂V is, by definition, supported on an annulus A. By Lemma 12.2 (and using the normality of the subgroup of all homeomorphisms isotopic to the identity) it may be assumed that all the twists concerned are so supported. As in Lemma 12.2, ∂V has a collar neighbourhood in V, a neighbourhood homeomorphic to $\partial V \times [0, 1]$ with ∂V identified with $\partial V \times 0$. Of course, $A \times [0, 1] \subset \partial V \times [0, 1]$, and τ extends to $\tau \times 1$ on $A \times [0, 1/2]$. Then τ extends, by the identity, over the remainder of the closure of $V - (A \times [1/2, 1])$. Thus τ extends over V after the removal of the interior of a solid torus. This means that the product ψ, of twists and inverse twists supported on annuli in ∂V, extends to a homeomorphism from V less the interiors of solid tori to V less the interiors of (in general, different) solid tori. The solid tori that permit successive twists to extend are removed from successively narrower collars of ∂V. Thus, at the cost of removing these solid tori, there is a homeomorphism of V to V sending each p_i to q_i, so gluing on copies of U by means of h to the first copy of V and by ψh to the second copy gives the required result. □

Note that, with the notation of the above proof τ maps the boundary of the meridian disc of the solid torus $A \times [1/2, 1]$ to a curve that is homologous to one longitude plus some number of meridians of the boundary of the solid torus. The solid torus can, then, be parametrised as $S^1 \times D^2$ so that τ maps $\{\star\} \times \partial D^2$ to $S^1 \times \{\star\}$. This translates at once into the following result:

Theorem 12.14. *Any closed connected orientable 3-manifold M can be obtained from* S^3 *by a collection of 1-surgeries, that is, by removing disjoint copies of* $S^1 \times D^2$ *and replacing them with copies of* $D^2 \times S^1$ *in the canonical way. Thus M bounds a 4-manifold that is a 4-ball to which a collection of 2-handles has been added.*

In using this result the disjoint copies of $S^1 \times D^2$ that are to be removed from S^3 are thought of as a neighbourhood of a link in S^3. In order to specify the parametrisation of this neighbourhood by copies of $S^1 \times D^2$, parallels (in the $S^1 \times D^2$ structures) to the link components (the cores of the solid tori) are specified. Each parallel, or *framing curve*, is a simple closed curve on the boundary of a solid torus neighbourhood of a link component that *will* bound a disc when $S^1 \times D^2$ is replaced by $D^2 \times S^1$. Each parallel can be specified by an integer, allocated to the component of the link, that specifies the linking number in S^3 of the component and its parallel (both oriented in the same direction around the solid torus neighbourhood). Alternatively the framed link can be taken to be a link of thin bands (annuli), the two boundary components of each annulus being a component of the link and its parallel. Sometimes the link is drawn with cross-overs in the plane (or some other surface), and it is assumed that the designated parallel always runs beside the link component in the 2-dimensional projection. The framing so encoded by a diagram is sometimes colloquially described as the "blackboard framing".

The representation of a closed connected orientable 3-manifold by means of surgery on a framed link is by no means unique. That certainly seems likely from the proof of Theorem 12.13. There is no unique way of expressing a homeomorphism as a product of twists, for there are relations in the mapping class group of a surface. The following theorem, due to Kirby [65], describes two ways in which the framed links can be changed without changing the 3-manifolds that result from them by means of surgery. It is fairly easy to see that the changes of links by such Kirby moves do not change the 3-manifold. What is not obvious is the fact that iterations of these two types of move relate any two framed links representing the same 3-manifold.

Figure 12.6

Theorem 12.15. *Two framed links in S^3 give, by surgery, the same oriented 3-manifold if and only if they are related by a sequence of moves of two types. In a move of type* 1, *an extra unknotted component, unlinked from all other components, with framing* 1 *or* -1 *is added to or removed from the link. In a move of type* 2, *any two components that are, together with their framing curves, contained in a doubly punctured disc (itself possibly knotted up and linked with other components) in* S^3, *as on the left of Figure 12.6, can be changed to the two curves on the right, the new framing curves again being on the punctured disc.*

For the proof, which uses 4-dimensional Cerf theory, refer to [65]. If one considers the surgery information as a recipe for adding 2-handles on to a 4-ball to create a 4-manifold with the 3-manifold as its boundary, a move of type 2 corresponds to sliding one 2-handle over another. A type 1 move changes the 4-manifold by taking the connected sum with a complex projective plane (oriented in either way), or by removing such a summand. Neither manoeuvre changes the boundary of the 4-manifold. The two moves of Theorem 12.15 can be, and indeed have been, explored at length to give many examples of different framed links representing the same manifold [65]. An interesting exercise is to show that any closed connected orientable 3-manifold can be obtained by surgery on S^3 using a framed link with all its components unknotted (a crossing in a link diagram can be changed by introducing, by a type 1 move, a new component and then employing two type 2

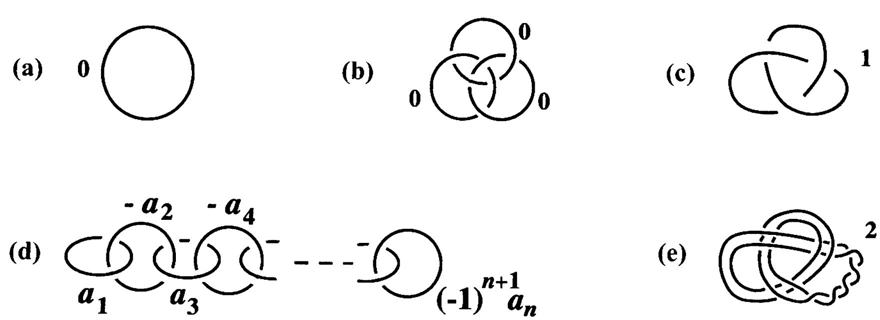


Figure 12.7

moves). A few examples of framed links that yield, by surgery, certain well-known 3-manifolds are shown in Figure 12.7, where the integers indicate the linking numbers of framing curves. Diagram (a) is $S^1 \times S^2$, (b) is $S^1 \times S^1 \times S^1$, (c) is the Poincaré homology 3-sphere with finite fundamental group, (d) is the lens space $L_{p,q}$ where q/p has continued fraction expansion $1/\{a_1 + 1/\{a_2 + 1/\{a_3 + \cdots + 1/a_n\}\}\}$, and (e) is the connected sum of a homology 3-sphere and real projective 3-space. The manifold obtained by surgery on a (p, q) cable knot with framing pq always has $L_{p,q}$ as a connected summand.

At the beginning of this chapter, the mapping class group (self-homeomorphisms up to isotopy) of a space was introduced, and the twist homeomorphisms of a surface were discussed. For a closed orientable surface the isotopy classes of orientation-preserving homeomorphisms form a subgroup of index 2 in the mapping class group (beware that sometimes it is that subgroup that is named the mapping class group). It can be shown that that subgroup is generated by all twists ([77] [24]). Further, a finite collection of twists generate ([78], [80]) this subgroup; a minimal collection of twist generators, found by S. P. Humphries [48], consists of the twists about the set of $(2g + 1)$ curves shown in Figure 12.8. For a torus T these are just the familiar longitude and meridian curves; twists about them induce standard generators of the group of automorphisms of $H_1(T)$ of determinant 1.

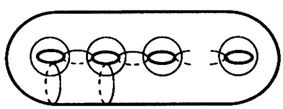


Figure 12.8

A finite *presentation* for the mapping class group of a surface was given by B. Wajnryb [131]. Note that for a surface of genus 2 the five generators for the orientation-preserving group commute with a π rotation about the "horizontal axis" (see Figure 12.9), so this rotation is in the centre of the group. This implies that a 3-manifold with a Heegaard splitting of genus 2 has a self-homeomorphism of period 2. In turn, using results of Thurston, that can be shown to lead to the result that any simply connected closed *genus two* 3-manifold is S^3; that is, the famous Poincaré conjecture is true for genus two 3-manifolds.

Studies of the mapping class group of a closed non-orientable surface can be found in [79], [19] and [8].

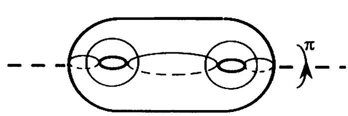


Figure 12.9

Exercises

1. Prove that any piecewise linear homeomorphism $h : B^n \to B^n$, with $h|\partial B^n$ the identity, is isotopic, keeping the boundary fixed, to the identity homeomorphism. [Hint: Consider the $(n+1)$-ball $B^n \times [0, 1]$ as the cone on its boundary, with an interior point for the vertex of the cone.]

2. Let X be a disc with n holes (that is, a 2-sphere from which the interiors of $n + 1$ disjoint discs have been removed). Suppose that $h : X \to X$ is a (piecewise linear) homeomorphism that is the identity on ∂X. Prove that h can be expressed as a product of finitely many twists. [Hint: If α is an arc in X from one boundary component to another, consider the intersection of $h\alpha$ with other such arcs.]

3. Prove that any orientation-preserving homeomorphism of a closed connected surface F to itself is expressible as a product of finitely many twists.

4. Consider the mapping class group of isotopy classes of orientation-preserving homeomorphisms of the torus to itself. Show that this is isomorphic to the group of 2×2 matrices over $\mathbb{Z}$ generated by $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. What is the mapping class group of isotopy classes of all homeomorphisms of the Klein bottle?

5. Suppose that C is a simple closed curve in an orientable surface X and that $h : X \to X$ is an orientation-preserving homeomorphism. How are the twist about C and the twist about hC related? Suppose that p and q are simple closed curves in X that intersect transversely in precisely one point; τ_p and τ_q are the twists about p and q. Show that in the mapping class group of X, $\tau_p \tau_q \tau_p = \tau_q \tau_p \tau_q$.

6. Prove that surgery on the unknot in S^3 with ± 1 framing just produces S^3.

7. Find two distinct non-trivial framed *knots* in S^3 that describe, by means of surgery, the same 3-manifold.

8. Verify that diagrams (a) and (b) of Figure 12.7 are indeed surgery diagrams for $S^1 \times S^2$ and $S^1 \times S^1 \times S^1$. Find surgery diagrams for real projective 3-space $\mathbb{R}P^3$ and (harder) for $S^1 \times F$, where F is a closed connected orientable surface.

9. What is the effect on S^3 of (i) a 0-surgery and (ii) a 2-surgery?

10. Show that any closed connected orientable 3-manifold can be obtained by surgery on a framed link in *any* other such manifold.

11. If M is a 3-manifold with a genus g Heegaard splitting, show that the fundamental group of M has a presentation with g generators and g relators.

12. Suppose an orientable connected surface is described in terms of a handle decomposition with just one 0-handle and some 1-handles. Use the idea of sliding a 1-handle over other 1-handles (a 2-dimensional version of the 4-dimensional handle sliding described in the second type Kirby move) to produce a canonical form for the surface as depicted in Figure 6.1. What happens if there are more 0-handles and some 2-handles?

13

3-Manifold Invariants from the Jones Polynomial

As proved in Chapter 12, any closed connected orientable 3-manifold can be obtained by the process of surgery on a framed link in S^3. Any invariant of framed links can be applied to such a surgery prescription in the hope of finding an invariant of the 3-manifold. That would need to be some entity associated to the 3-manifold and not just to the particular surgery description; it would need to be unchanged by all possible Kirby moves. An elementary example comes from the idea of linking numbers. A framed link (with components temporarily ordered and oriented) has a *linking matrix*. This is the symmetric matrix with entries the linking numbers between the pairs of components of the link. The linking number of a component with itself (a diagonal term of the matrix) is taken to be the integer that gives the framing of that component. This linking matrix can easily be seen to be a presentation matrix (in the sense of Chapter 6) for the first homology of the 3-manifold arising from surgery on the framed link. Thus the modulus of the determinant of the matrix, if it is non-zero, is the order of that homology group and the nullity of the matrix is the first Betti number of the manifold. It is easy to check that these numerical invariants do indeed remain unchanged by Kirby moves on the framed link. This, however, is not too exciting, as homology is long and better understood by other means. One might hope to emulate this procedure by a simple direct application of some link invariant. The Alexander polynomial and the Jones polynomial fail in that respect. This chapter explains how the Jones polynomial can nevertheless be amplified to achieve a 3-manifold invariant. Roughly, the idea is to take a linear sum of the Jones polynomials, evaluated at a complex root of unity, of copies of the link with the components replaced by various parallels of the original components. The resulting invariants are known as Witten's quantum $SU_q(2)$ 3-manifold invariants. The details are somewhat intricate and, as might be expected, will here be eased by the simplifying approach of the Kauffman bracket and the linear skein theory associated with it. The Temperley–Lieb algebras appear as instances of that theory. E. Witten's initiation of this topic can be found in [135].

A proof, using quantum groups, of the existence of these $SU_q(2)$ 3-manifold invariants was given first by N. Y. Reshetikhin and V. G. Turaev [109]; it was amplified by Kirby and P. Melvin [68]. Early proofs using skein theory, or the

Temperley–Lieb algebra, appeared in [82] and [83]; the proof that follows here first appeared in [86].

When thinking about surgery, *framed* links are needed. As remarked in Chapter 12, a framing can be interpreted as annuli twisting along the components of a link, and this can be encoded by a planar diagram of the link. The understanding then is that each annulus is a widening of each component *in the plane of the diagram*. Extra twists of the annulus correspond to extra 'kinks' in the diagram. This means that the writhe of any component of the diagram is the linking number of a boundary curve of the annulus with that link component. This integer is the framing of the link component. A diagram so encoding the framing will be called a diagram for the framed link. Of course, moving a framed link around by isotopy in S^3 will not change at all the result of surgery upon it. This moving around corresponds to the equivalence of regular isotopy on representing diagrams in $\mathbb{R}^2 \cup \infty$. Recall that regular isotopy is generated by the Reidemeister moves of Types II and III.

A Kirby type 2 move on a *diagram* of a framed link can be thought of as dragging a segment of one component of the link up to another and then passing it over to the far side of that component. The framings so encoded by the diagrams are then correct for such a move. A Kirby type 1 move consists of adding to a diagram, or subtracting from it, a curve with precisely one crossing. The theorem of Kirby [65], Theorem 12.15, is then that closed connected oriented 3-manifolds are equivalent if and only if any link diagrams that represent them (with respect to surgery) differ by regular isotopy and a sequence of Kirby moves of the above two types. Thus, to construct a 3-manifold invariant, it is necessary only to associate with each link diagram some algebraic concept that does not change when the diagram changes under regular isotopy or Kirby moves. Of course any link invariant is unchanged under (regular) isotopy. It is in accommodating the type 2 move that difficulty arises; the type 1 move turns out to be almost a piece of administration.

Consider now, for a surface, the following version of the linear skein theory associated to the Kauffman bracket. Let F be an oriented surface with a finite collection (possibly empty) of points specified in its boundary ∂F. A *link diagram* in the surface F consists of finitely many arcs and closed curves in F, with just finitely many transverse crossings with the usual "over and under" information; the end points of the arcs must be precisely the specified points in ∂F. This definition is meant to contain no surprise. Two diagrams are regarded as the same if they differ by a homeomorphism of F that is isotopic to the identity always keeping ∂F fixed. The required linear skein theory of F (inspired by the Kauffman bracket) is defined as follows:

Definition 13.1. Let A be a fixed complex number. The linear skein $\mathcal{S}(F)$ of F is the vector space of formal linear sums, over $\mathbb{C}$, of (unoriented) link diagrams in F qotiented by the relations

(i) $D \cup$ (a trivial closed curve) $= (-A^{-2} - A^2)D$,

(ii)  $= A$  $+ A^{-1}$  .

Here a trivial closed curve in a diagram is one that is null-homotopic and that contains no crossing. The empty set is a permitted diagram if no point is specified in ∂F. The equation in (ii) refers to three diagrams that are identical except where shown. It follows, exactly as in Lemma 3.3, that diagrams that are regularly isotopic in F (that is, related by the Reidemeister Type II and III moves in F) represent the same element of $\mathcal{S}(F)$. Although a linear skein space is in this way associated with any oriented surface, the only surfaces needed in what follows are the plane, the sphere, the annulus and the disc.

The linear skein of the plane, $\mathcal{S}(\mathbb{R}^2)$, is easily seen to be a 1-dimensional vector space with the empty diagram as a (fairly natural) base. $\mathcal{S}(\mathbb{R}^2)$ will thus be identified with $\mathbb{C}$. This is because, by use of (ii), any link diagram in any surface can be expressed uniquely as a linear sum of diagrams with no crossing at all, and, in this case, it follows from (i) that those diagrams are multiples of the empty diagram. Of course, this is the Kauffman bracket approach to the Jones polynomial; the Kauffman bracket of a diagram is the coordinate of the diagram in $\mathcal{S}(\mathbb{R}^2)$ if the zero-crossing diagram of the unknot were the base. The inclusion of $\mathbb{R}^2$ in $\mathbb{R}^2 \cup \infty$ induces an isomorphism of the skein spaces of the plane and the sphere.

The linear skein of the annulus, $S^1 \times I$, similarly has a base consisting of diagrams with no crossing and no null-homotopic closed curve. Each base element is then just a number of parallel curves encircling the annulus. A product of a diagram in an annulus with a diagram in another annulus can be defined by identifying together one boundary component from each annulus. This produces a third annulus containing a diagram that is the union of the two original diagrams. It is easy to see that this operation induces a well-defined bilinear product on $\mathcal{S}(S^1 \times I)$ that turns it into a commutative algebra. Let α denote the base element that consists of one single curve encircling the annulus once with no crossing. Then the base mentioned above is $\{\alpha^0, \alpha^1, \alpha^2, \ldots\}$, where α^0 denotes the empty diagram in the annulus, and α^n is represented by n parallel curves all encircling the annulus. $\mathcal{S}(S^1 \times I)$ is thus $\mathbb{C}[\alpha]$ the polynomial algebra in α with complex coefficients. Next consider the linear skein $\mathcal{S}(D^2, 2n)$ of a disc with $2n$ points in its boundary. Again, this has a base consisting of all diagrams with no crossing and no closed curve. (A combinatorial exercise shows there are $\frac{1}{n+1}\binom{2n}{n}$ such diagrams, this number being the n^{th} Catalan number.) Regarding the disc as a square with n standard points on the left edge and n on the right, a product of diagrams can be defined by juxtaposing squares, identifying the right edge of one (with its n special points) with the left edge of the other. This product of diagrams extends to a well-defined bilinear map that turns $\mathcal{S}(D^2, 2n)$ into an algebra TL_n, the n^{th} *Temperley–Lieb algebra*. As an algebra TL_n is generated by n elements $1, e_1, e_2, \ldots, e_{n-1}$ shown in Figure 13.1, for any of the above base elements is a product of these (an easy exercise). In this and later diagrams, an integer n beside an arc signifies n copies of that arc all parallel in the plane so that, for example, the identity element $\mathbf{1} \in TL_n$ is n parallel arcs going from one side of the square to the other. Note that in practice, some figures will, for convenience, show the square as a rectangle!

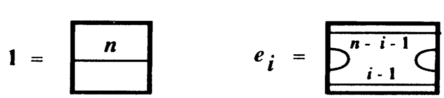


Figure 13.1

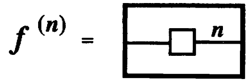


Figure 13.2

Nothing subtle has yet occurred. It is now, however, essential to understand the definition of the Jones–Wenzl idempotent $f^{(n)} \in TL_n$ as defined in [133]. In the following figures $f^{(n)}$ will be shown as a small blank square with n arcs entering and n leaving (see Figure 13.2); indeed, the number of such arcs is used to determine to which value of n, and hence to which Temperley–Lieb algebra, such a blank square refers. Although $f^{(n)}$ will be represented by a linear sum of diagrams, it is sometimes helpful to pretend it is just one diagram! The complex number Δ_n will be that obtained by placing $f^{(n)}$ in the plane, joining the n points on the left of the square by parallel arcs to those on the right (see Figure 13.3) and interpreting the result in $\mathcal{S}(\mathbb{R}^2) \equiv \mathbb{C}$. This type of definition will occur again. More pedantically, Δ_n is the image of $f^{(n)}$ under the linear map $TL_n \to \mathcal{S}(\mathbb{R}^2) \equiv \mathbb{C}$ induced by mapping each diagram in the square (with $2n$ boundary points) to a planar diagram formed by the above standard joining-up process.

Figure 13.3

The element $f^{(n)}$ is defined and characterised in the following lemma:

Lemma 13.2. *Suppose that A^4 is not a k^{th} root of unity for $k \leq n$. Then there is a unique element $f^{(n)} \in TL_n$ such that*
(i) $f^{(n)}e_i = 0 = e_i f^{(n)}$ *for* $1 \leq i \leq n-1$,
(ii) $(f^{(n)} - \mathbf{1})$ *belongs to the algebra generated by* $\{e_1, e_2, \dots, e_{n-1}\}$,
(iii) $f^{(n)}f^{(n)} = f^{(n)}$ *and*
(iv) $\Delta_n = (-1)^n(A^{2(n+1)} - A^{-2(n+1)})/(A^2 - A^{-2})$.

PROOF. Note that if $f^{(n)}$ exists, $\mathbf{1} - f^{(n)}$ is the identity of the algebra generated by $\{e_1, e_2, \dots, e_{n-1}\}$, and so $f^{(n)}$ is then certainly unique. Let $f^{(0)}$ be the empty diagram (so that $\Delta_0 = 1$), let $f^{(1)} = \mathbf{1}$, and inductively assume that $f^{(2)}, f^{(3)}, \dots, f^{(n)}$ have been defined with the above properties (i), (ii), (iii) and (iv). Observe that (i) and (ii) immediately imply (iii) and that this generalises to the identity shown in Figure 13.4 provided that $(i + j) \leq n$.

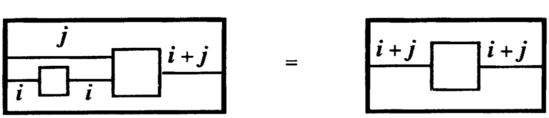


Figure 13.4

Now consider the element x, say, of TL_{n-1} shown at the start of Figure 13.5. The identity of Figure 13.4 implies that $f^{(n-1)}x = x$. But $f^{(n-1)}x$ is, by (i), just some scalar multiple λ of $f^{(n-1)}$ (because x is a linear sum of **1**'s and products of e_i's); the trick of placing squares in the plane and joining points on the left to points on the right, in the standard way, implies that the scalar λ is Δ_n/Δ_{n-1}.

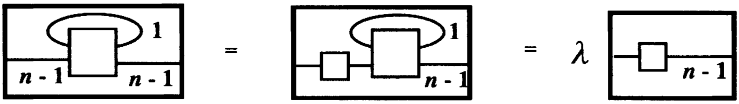


Figure 13.5

Suppose now that $A^{4k} \neq 1$ for $k \leq n+1$, so that $\Delta_k \neq 0$ for $k \leq n$. Define $f^{(n+1)} \in TL_{n+1}$ inductively by the equation of Figure 13.6.

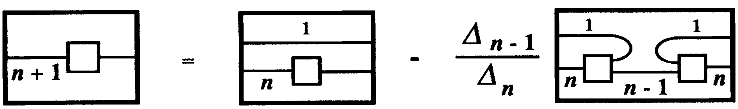


Figure 13.6

Properties (i) and (ii) (and hence (iii)) for $f^{(n+1)}$ follow immediately, except perhaps for the fact that $f^{(n+1)}e_n = 0$. However, Figure 13.7 shows, using the identities of Figure 13.5 and Figure 13.4, why that also is true.

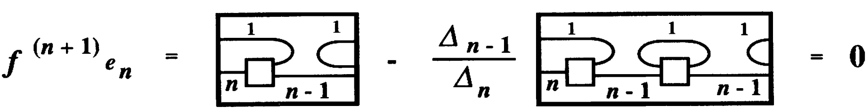


Figure 13.7

It remains to investigate Δ_{n+1}. Consider the operation of placing a square in an annulus and joining k points on one side to k points on the other by parallel arcs encircling the annulus. For each k, this gives a linear map $TL_k \to \mathcal{S}(S^1 \times I)$. The image of $f^{(k)}$ is some polynomial $S_k(\alpha)$ in the generator α of $\mathcal{S}(S^1 \times I)$. $S_0(\alpha) = \alpha^0$ and $S_1(\alpha) = \alpha$. Inserting into the annulus, in this way, the defining relation of Figure 13.6 for $f^{(n+1)}$ gives the formula of Figure 13.8. However, in the last diagram in Figure 13.8 the two small squares representing $f^{(n)}$ can be slid together to become one small square (using $f^{(n)}f^{(n)} = f^{(n)}$), and an application of the formula of Figure 13.5 gives

$$S_{n+1}(\alpha) = \alpha S_n(\alpha) - S_{n-1}(\alpha).$$

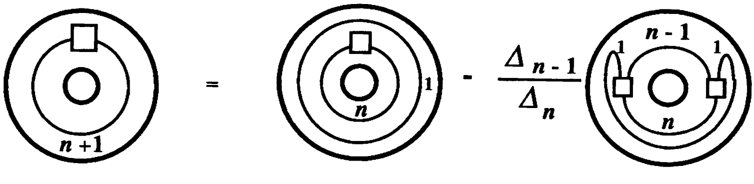


Figure 13.8

This, with the above initial conditions, is the recurrence formula for the n^{th} Chebyshev polynomial (of the second kind, renormalised) in α. If now an annulus is placed in the plane, the ensuing linear map $\mathcal{S}(S^1 \times I) \to \mathcal{S}(\mathbb{R})$ sends α^k to $(-A^{-2} - A^2)^k$, and by definition it maps $S_k(\alpha)$ to Δ_k. Thus

$$\Delta_{n+1} = (-A^{-2} - A^2)\Delta_n - \Delta_{n-1}.$$

An induction argument then easily shows that

$$\Delta_{n+1} = \frac{(-1)^{n+1}(A^{2(n+2)} - A^{-2(n+2)})}{A^2 - A^{-2}}. \qquad \square$$

The proof of Lemma 13.2 could have been slightly shortened by inserting the squares directly into the plane, but consideration of the annulus is important later. Also, attention has been drawn to the Chebyshev polynomial S_n, which, with indeterminate x and integer coefficients, is defined by

$$S_{n+1}(x) = xS_n(x) - S_{n-1}(x); \qquad S_0(x) = 1, \quad S_1(x) = x.$$

It has the important (easy) properties that

$$S_n(x) = (-1)^n S_n(-x) \text{ and } (t - t^{-1})S_n(t + t^{-1}) = t^{n+1} - t^{-(n+1)}.$$

Further, it has been seen that $f^{(n)}$ inserted into $S^1 \times I$ with the boundary points of $f^{(n)}$ connected up by arcs encircling the annulus is $S_n(\alpha) \in \mathcal{S}(S^1 \times I)$. This features in the next most important definition, soon to be extensively employed.

Definition 13.3. For a given integer r, let $\omega \in \mathcal{S}(S^1 \times I)$ be defined by

$$\omega = \sum_{n=0}^{r-2} \Delta_n S_n(\alpha).$$

As a final instance of skein theory, consider the linear skein of an annulus with two points specified on *one* of its boundary components, $\mathcal{S}(S^1 \times I,\ 2 \text{ points})$. Let $a\omega$ and $b\omega$ be the elements of $\mathcal{S}(S^1 \times I,\ 2 \text{ points})$ that consist of ω inserted into the annulus together with an arc, joining the two boundary points of the annulus; the arc goes "above" ω for $a\omega$ or "below" ω for $b\omega$ (see Figure 13.9).

Lemma 13.4. *In $\mathcal{S}(S^1 \times I,\ 2$ points), $a\omega - b\omega$ is a linear sum of two elements, each of which contains a copy of $f^{(r-1)}$ (That is, each of the two elements is*

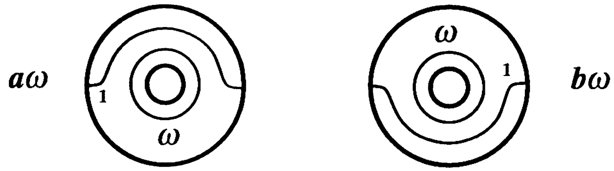


Figure 13.9

the image of $f^{(r-1)}$ *under some map* $TL_{r-1} \to \mathcal{S}(S^1 \times I,\ 2 \text{ points})$ *formed by including a square into an annulus and joining up boundary points in some way.)*

PROOF. Consider the inclusion, shown in Figure 13.10, of the TL_{n+1} recurrence relation of Figure 13.6 into the annulus.

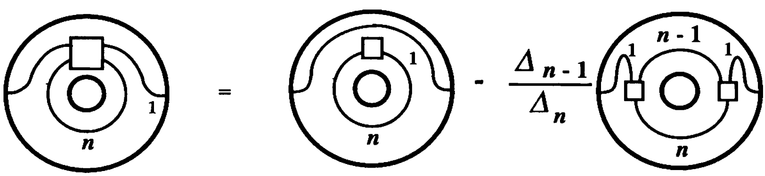


Figure 13.10

The top boundary points on either side of the square are joined to the two points on the annulus boundary, and the other n points on the left of the square are joined to the n on the right by parallel arcs encircling the annulus. As in the proof of Lemma 13.2, the two small squares in the final diagram of Figure 13.10 can be slid together (using $f^{(n)} f^{(n)} = f^{(n)}$) to become one square, and the equality can then be rearranged to become that of Figure 13.11. Sum these equalities from $n = 0$ to $n = r - 2$ (here $\Delta_{-1} = 0$). The right-hand side is $a\omega$. Rotate each annulus of Figure 13.11 through π and sum again. The right-hand side is now $b\omega$. The left-hand sides of the formulae so obtained are almost the same; recalling that $\Delta_{-1} = 0$, the difference of these left-hand sides is the difference of the first term of Figure 13.11, when $n = r - 2$, and its rotation; in each is a copy of $f^{(r-1)}$. □

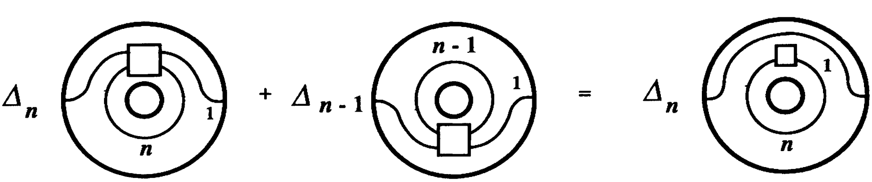


Figure 13.11

If D is a planar diagram of a link of n ordered components, D defines a multilinear map

$$< \ , \ , \ldots , \quad >_D : \mathcal{S}(S^1 \times I) \times \mathcal{S}(S^1 \times I) \times \cdots \times \mathcal{S}(S^1 \times I) \to \mathcal{S}(\mathbb{R}^2).$$

This map is defined, by multilinearity, using the following construction with diagrams: Take n link diagrams in n annuli and immerse the annuli, with their diagrams, in the plane as a regular neighbourhood of the n (ordered) components of D. Over- and under-crossings of D become over- and under-crossings of the immersed annuli and of the diagrams that they contain. In this way the diagrams in the annuli are made to run parallel to the original components of D. Then the n annulus diagrams have produced a diagram in $\mathbb{R}^2$ representing an element of $\mathcal{S}(\mathbb{R}^2) \equiv \mathbb{C}$. It is easy to check that this induces a well-defined map of the required form. As a simple example, let D be the diagram on the left of Figure 13.12; then $< \alpha^2, \alpha, 1 >_D$ is represented in $\mathcal{S}(\mathbb{R}^2)$ by the diagram on the right of Figure 13.12 (remember that in $\mathcal{S}(S^1 \times I)$, α is the generator and 1 is represented by the empty set).

Figure 13.12

Lemma 13.5. *Suppose that A is chosen so that A^4 is a primitive r^{th} root of unity, $r \geq 3$. Suppose that D is a planar diagram of a link of n (ordered) components. Suppose that D' is another such diagram, obtained from D by a Kirby type 2 move, in which a parallel of the first component of D is joined by some band to another component (or, equivalently, a segment of the second component is moved up to and over the first). Then*

$$< \omega, \ , \ldots , \quad >_D = < \omega, \ , \ldots , \quad >_{D'} .$$

PROOF. It must be checked that the elements of $\mathcal{S}(\mathbb{R}^2)$, produced as described above from D and from D', with ω as the "diagram" around the first component and with any given diagrams around the others, are in fact the same element. The difference between these elements is the result of a sequence of moves, each consisting of moving, by regular isotopy, an arc of some component up to that labelled with ω, and changing from an immersed copy of $a\omega$ to one of $b\omega$. By Lemma 13.4, this difference is a linear sum of elements of $\mathcal{S}(\mathbb{R}^2)$, each containing a copy of $f^{(r-1)}$. However, in $\mathcal{S}(\mathbb{R}^2)$ any element containing a copy of $f^{(r-1)}$ is *zero* if A^4 is a primitive r^{th} root of unity. That is because such an element is, for some $x \in TL_{r-1}$, the image in $\mathcal{S}(\mathbb{R}^2)$ of $f^{(r-1)}x$ under the map induced by placing the square in the plane and joining the $r-1$ points on the left to those on the right by parallel arcs. As usual, $f^{(r-1)}x$ is a scalar multiple of $f^{(r-1)}$, but $f^{(r-1)}$ maps to Δ_{r-1} and $\Delta_{r-1} = 0$ because $A^{4r} = 1$. □

Corollary 13.6. *If A^4 is a primitive r^{th} root of unity, $r \geq 3$, and planar diagrams D and D' are related by a sequence of Kirby moves of type 2, then*

$$< \omega, \omega, \ \ldots \ , \omega >_D = < \omega, \omega \ \ldots \ , \omega >_{D'} .$$

In what follows, U_+ and U_- will be planar figure-eight diagrams, each with one crossing, representing the unknot with framings $+1$ and -1 respectively; U will denote the diagram of the 0-framed unknot with no crossing at all. The definition of ω implies at once that $< \omega >_U = \sum_{n=0}^{r-2} \Delta_n{}^2$. When A^4 is a primitive r^{th} root of unity, the substitution

$$\Delta_n = \frac{(-1)^n (A^{2(n+1)} - A^{-2(n+1)})}{A^2 - A^{-2}}$$

and the summation of the ensuing geometric progression produce the formula

$$< \omega >_U = \sum_{n=0}^{r-2} \Delta_n{}^2 = \frac{-2r}{(A^2 - A^{-2})^2} .$$

Note that $< \omega >_U \neq 0$. The next result implies that $< \omega >_{U_+}$ and $< \omega >_{U_-}$ are also non-zero. The proof of this lemma will be given a little later.

Lemma 13.7. *Suppose $r \geq 3$ and A is a primitive $4r^{th}$ root of unity. Then*

$$< \omega >_{U_+} < \omega >_{U} = < \omega >_U = \frac{-2r}{(A^2 - A^{-2})^2} .$$

Now comes the theorem (first proved in another form in [109]) asserting the existence of certain 3-manifold invariants that, up to normalisation, are often called the quantum $SU_q(2)$ invariants. First though, recall the linking matrix of a framed link with ordered oriented components mentioned at the start of this chapter. This matrix changes by congruence under Kirby type 2 moves, so its numbers of positive and of negative eigenvalues do not change under such moves, nor do they change if different orientations or orderings on the link's components are chosen.

Theorem 13.8. *Suppose that a closed oriented 3-manifold M is obtained by surgery on a framed link that is represented by a planar diagram D. Let b_+ be the number of positive eigenvalues and b_- be the number negative eigenvalues of the linking matrix of this link. Suppose $r \geq 3$ and that A is a primitive $4r^{th}$ root of unity. Then*

$$< \omega, \omega, \ldots , \omega >_D < \omega >_{U_+}^{-b_+} < \omega >_{U}^{-b_-}$$

is a well-defined invariant of M.

PROOF. Note that A is a primitive $4r^{th}$ root of unity, and so, by Lemma 13.7, $< \omega >_{U_+}$ and $< \omega >_{U}$ are non-zero. It follows from the Corollary 13.6 and the preceding remarks about the linking matrix that the given expression is invariant under Kirby type 2 moves. The last two factors make it invariant under Kirby type

1 moves, and regular isotopy of D just induces regular isotopies of all the diagrams used in defining the expression. □

The invariant just defined is essentially the $SU_q(2)$ invariant of M at a "level" corresponding to r. Observe however that if ω is replaced throughout by $\mu\omega$, where μ is a constant complex number, then clearly another slightly different invariant is obtained. (The new invariant is the old one multiplied by μ raised to the power of the first Betti number of M which is the nullity of the above linking matrix.) It may often be more convenient to use some such renormalisation. Some small generalisations to this whole approach can be made in several directions. One can take A to be a primitive $2r^{th}$ root of unity when r is odd ([12], or see [86]). One can take A to be an indeterminate symbol rather than a complex number and work with modules over $\mathbb{Z}[A, A^{-1}]$ rather than vector spaces, quotienting when appropriate by a cyclotomic polynomial. One can also rephrase the exposition in terms of the skein theory of framed links in 3-manifolds rather than using link diagrams in surfaces. The invariant generalises at once to become an invariant of framed links in 3-manifolds; just add extra components to the surgery link (see [86]).

A more subtle extension to the theory comes from expressing ω as $\omega_0 + \omega_1$, where

$$\omega_0 = \sum_{\substack{n=0 \\ n \text{ even}}}^{r-2} \Delta_n S_n(\alpha), \qquad \omega_1 = \sum_{\substack{n=0 \\ n \text{ odd}}}^{r-2} \Delta_n S_n(\alpha).$$

If ω is replaced by ω_0 or ω_1 in Figure 13.9, the result analogous to that of Lemma 13.4 is that each of $a\omega_0 - b\omega_1$ and $a\omega_1 - b\omega_0$ is a multiple of an element containing a copy of $f^{(r-1)}$. The theory just described can be altered by decorating some subset of the components of the surgery link with ω_0 the remainder with ω_1. Careful choice of those subsets leads ([11], or see [86]) to invariants of a 3-manifold M with spin structure or with a preferred element of $H^1(M; \mathbb{Z}/2\mathbb{Z})$.

To complete this chapter, a proof of Lemma 13.7 is needed. If the square, with n points specified on each of its two sides, is placed in the plane or in S^2, each element of TL_n (a linear sum of diagrams inside the square) can be regarded as a linear map to $\mathbb{C}$ of the linear skein of diagrams outside the square. This map is induced by taking a diagram inside and a diagram outside the square and regarding the union of the two as an element of $\mathcal{S}(\mathbb{R}^2) = \mathbb{C}$. As has already been noted, if $r \geq 3$ and A^4 is a primitive r^{th} root of unity, $f^{(r-1)}$ defines the zero map of outsides although it is not the zero element of TL_{r-1}. Consider now the element of TL_n shown in Figure 13.13, regarded as a map of outsides, that consists of $f^{(n)}$ encircled by an ω.

Lemma 13.9. *Suppose $r \geq 3$ and A is a primitive $4r^{th}$ root of unity. The element of TL_n shown in Figure 13.13 is the zero map of outsides if $1 \leq n \leq r-2$. When $n = 0$, the element acts as multiplication by $< \omega >_U$.*

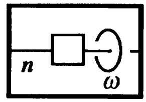


Figure 13.13

PROOF. Consider first the element of TL_n that consists of $f^{(n)}$ encircled by one simple closed curve. This is shown in Figure 13.14 for $n = 4$. Figure 13.14 shows a calculation for that element. Firstly one crossing is removed in the two standard ways, the results being multiplied by A and A^{-1} and added. The two elements obtained are then simplified by removing kinks and multiplying by $-A^{\pm 3}$. Now, in the two resulting elements, removal of any of the crossings depicted in *one* of the standard ways gives zero (as $f^{(n)}e_i = 0$), so only the other standard way need be considered. It follows that $f^{(n)}$ encircled by one simple closed curve is equal, in TL_n, to $(-A^{2(n+1)} - A^{-2(n+1)})f^{(n)}$.

Now the element required in this Lemma is $f^{(n)}$ encircled by an ω, regarded as a map of outsides. Let this be denoted x. A small single unknotted simple closed curve inserted into this changes x, in the usual way, to $(-A^{-2} - A^2)x$. However, that small curve can be slid right over the ω without (by Lemma 13.5) changing the map of outsides, and then removed altogether (by the preceding paragraph) at the cost of multiplying by $(-A^{2(n+1)} - A^{-2(n+1)})$. Thus $(-A^{-2} - A^2)x = (-A^{2(n+1)} - A^{-2(n+1)})x$. Hence either $x = 0$ or $A^{2(n+1)} = A^2$ or $A^{2(n+1)} = A^{-2}$. The two latter possibilities do not occur for $1 \leq n \leq r-2$, as A is a primitive $4r^{th}$ root of unity, so then $x = 0$. When $n = 0$, it is trivial that x acts as multiplication by $< \omega >_U$ because there is nothing but the curve labelled ω to consider. □

PROOF OF LEMMA 13.7. By Corollary 13.6, $< \omega >_{U_+} < \omega >_U$ is equal to a component with one crossing labelled ω simply linked with a component with no self-crossing also labelled ω, (see Figure 13.15). By definition the ω on the first component is $\sum_{n=0}^{r-2} \Delta_n S_n(\alpha)$, and $S_n(\alpha)$ is $f^{(n)}$ inserted into the annulus and joined around the annulus by n parallel arcs. By Lemma 13.9, the linking curve

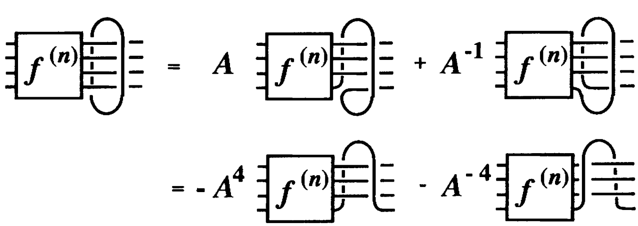


Figure 13.14

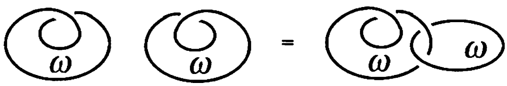


Figure 13.15

labelled ω converts to zero each term of the summation except the first (when $n = 0$). Thus $< \omega >_{U_+} < \omega >_{U_-} = < \omega >_U$. □

That completes the proof of the existence of the $SU_q(2)$ 3-manifold invariants associated with the Jones polynomial. The first proof by V. G. Turaev and H. Wenzl, based on representation theory, of the existence of $SU_q(n)$ invariants (associated with the HOMFLY polynomial of Chapter 15) can be found in [128]. A skein theory $SU_q(n)$ proof—a much harder version of the proof of this chapter—is given by Y. Yokota [139].

Exercises

1. Prove that the elements $1, e_1, e_2, \dots e_{n-1}$ do indeed generate the Temperley–Lieb algebra TL_n.

2. Draw five diagrams that form a base of TL_3 and determine a specific expression for the idempotent $f^{(3)}$ as a linear sum of these base elements.

3. Consider the π-rotations of the square, used to define the Temperley–Lieb algebra TL_n, about axes from north to south, from east to west and perpendicular to the plane of the square. Show that for $n \geq 3$, these rotations induce involutions of TL_n which are not the identity but which fix the element $f^{(n)}$.

4. Let ρ be an element of the permutation group S_n. Let D_ρ be the element of the Temperley–Lieb algebra TL_n that consists of precisely n arcs, one arc joining the point labelled i on the left edge of the square to the point labelled ρi on the right edge (where labellings start at the top). In D_ρ, if $i < j$ and $\rho i > \rho j$, there is one crossing between the arc starting at i and that starting at j and there the first arc is *over* the second. There are no other crossings. Let $|\rho|$ denote the number of crossings in D_ρ. Show that the idempotent $f^{(n)} \in TL_n$ is a scalar multiple of $\sum_{\rho \in S_n} A^{3|\rho|} D_\rho$ and determine that scalar.

5. The operation of placing a square, containing a generating diagram of TL_n, in the plane, joining the n points on the left to the n on the right (introducing no new crossing) and evaluating the result in the skein of the plane induces a linear map $\mathrm{tr} : TL_n \to \mathbb{C}$. (Thus $\mathrm{tr}(f^{(n)}) = \Delta_n$.) Show that $\mathrm{tr}(xy) = \mathrm{tr}(yx)$ and that $(x, y) \mapsto \mathrm{tr}(xy)$ defines a bilinear form on TL_n. If A is not a root of unity show that this form is non-degenerate.

6. Prove that the Chebyshev polynomials have a product formula of the form $S_m(x)S_n(x) = \sum_r S_r(x)$, and determine the range of r for given m and n

7. The collections of elements $\{\alpha^n : n \geq 0\}$ and $\{S_n(\alpha) : n \geq 0\}$ are bases of the space $\mathcal{S}(S^1 \times I)$. Find an expression for α^n as a linear sum of elements in the second base.
[Hint: Prove that $x^{n+1} = \sum_{r=0}^{n} \binom{n}{r} S_{n-2r+1}(x)$.]

8. Suppose a diagram D of a framed knot can be changed by mutation to become a diagram D', where the mutation is effected by rotating, in the usual way, a disc in the plane whose boundary meets D at just four points. Prove that $\langle S_n(\alpha)\rangle_D = \langle S_n(\alpha)\rangle_{D'}$. Is it true that $\langle \alpha^n\rangle_D = \langle \alpha^n\rangle_{D'}$?

9. Prove that the signature of the linking matrix of a framed link is not changed when the link is changed by a Kirby move of the second type.

10. Let A be a primitive $4r^{\text{th}}$ root of unity. Suppose that $\phi \in \mathcal{S}(S^1 \times I)$ is an element of the skein of the annulus with the property that if H is a two-crossing diagram of a non-trivial (Hopf) link, $\langle \phi, \psi\rangle_H = 0$ for all $\psi \in \mathcal{S}(S^1 \times I)$. Show that if D is a diagram of any other two-component link, then $\langle \phi, \psi\rangle_D = 0$ for all $\psi \in \mathcal{S}(S^1 \times I)$. [Hint: Use ω.]

11. Let D be a planar link diagram, $D_1, D_2, \ldots, D_n$ being the sub-diagrams of the individual components. Let A be a primitive $4r^{\text{th}}$ root of unity and suppose that $k \leq r-2$. Let $w(D_1)$ be the writhe of D_1. If $i(2), i(3), \ldots, i(n)$ are non-negative integers, show that

$$\langle S_k(\alpha), \alpha^{i(2)}, \alpha^{i(3)}, \ldots, \alpha^{i(n)}\rangle_D$$
$$= (-1)^{\Lambda+r}\left((-1)^{k+r+1}A^{-r^2}\right)^{w(D_1)} \langle S_{r-2-k}(\alpha), \alpha^{i(2)}, \alpha^{i(3)}, \ldots, \alpha^{i(n)}\rangle_D,$$

where $\Lambda = \sum_{j=2}^{n} i(j)\mathrm{lk}(D_1, D_j)$.

14

Methods for Calculating Quantum Invariants

The quantum $SU_q(2)$ 3-manifold invariants associated with a primitive $4r^{\text{th}}$ root of unity, described in the previous chapter, are fairly new and mysterious. Their use has so far been exceedingly limited in knot theory and in 3-manifold theory. Certainly they do distinguish many pairs of 3-manifolds, even pairs with the same homotopy type, but that has usually been more simply achieved by other means. However, there exist pairs of distinct manifolds with the same invariants for all r (see [85], [55] and [62]). For some manifolds, for some values of r the invariant is known by direct calculation to be zero. Superficially it might seem to be almost impossible to calculate any of these invariants. The calculation, from first principles, of the invariant corresponding to a $4r^{\text{th}}$ root of unity involves taking an $(r-2)$-parallel of a surgery link giving the 3-manifold. If the link's diagram has n crossings, that of the parallel has $n(r-2)^2$ crossings; calculating a Jones polynomial by naive means soon becomes impractical when many crossings are involved. It will be shown here that it is in principle fairly easy to give a formula, as a summation, for the invariants of lens spaces and, more generally, for certain Seifert fibrations. Although in theory any of the invariants can always be calculated, it is sensible to use various simplifying procedures whenever possible. Some of those will be described in this chapter. Tables of specific computer calculations appear in [104] and in [62], where one can search for patterns in the resulting lists of complex numbers.

The basic strategy in making calculations of the quantum $SU_q(2)$ invariants is to make calculations of elements of the skein of S^2, making as much use as possible of the idempotents $f^{(n)}$ of the Temperley–Lieb algebras TL_n. The methods for doing this were first developed by several authors; original accounts can be found in [87], [86], [84], [61], [62] and [137]. The next two important preparatory results relate to the Temperley–Lieb algebras.

Lemma 14.1. *The element of TL_n shown on the left of Figure 14.1, which consists of the idempotent $f^{(n)}$ followed by a complete positive "kink" in all n strands, is $(-1)^n A^{n^2+2n} f^{(n)}$.*

Proof. As shown in Figure 14.1, one strand can be separated a little from the other $n-1$ strands. Now removing the kink in that single strand contributes $-A^3$,

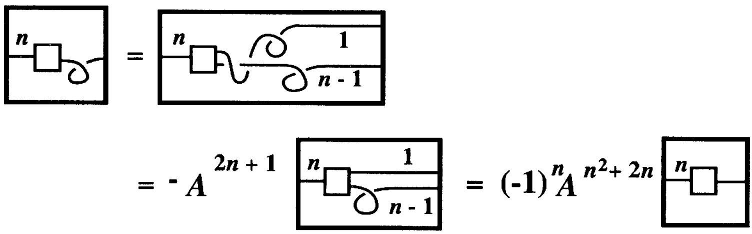


Figure 14.1

and (as in the proof of Lemma 13.9) removing all the other crossings of the single strand with the other $n-1$ strands contributes only a multiplying factor of $A^{2(n-1)}$ (any removal of a crossing in a negative manner gives zero on interacting with $f^{(n)}$). Thus the first diagram of Figure 14.1 is equal to $-A^{2n+1}$ times the third diagram, but that is $(-1)^{n-1}A^{n^2-1}f^{(n)}$ by induction on n. The result follows at once. □

Note that this implies that the removal a negative kink adjacent to an $f^{(n)}$ entails multiplying by a factor of $(-1)^n A^{-(n^2+2n)}$.

Lemma 14.2. *The element of TL_n shown in Figure 14.2, which consists of the idempotent $f^{(n)}$ with all its strands encircled by a parallel strands that join up the ends of an idempotent $f^{(a)}$, is*

$$(-1)^a \frac{A^{2(n+1)(a+1)} - A^{-2(n+1)(a+1)}}{A^{2(n+1)} - A^{-2(n+1)}} f^{(n)}.$$

PROOF. The a parallel strands and the idempotent $f^{(a)}$ can, as explained in Chapter 13, be thought of as $S_a(\alpha)$ contained in an annulus encircling the strands of $f^{(n)}$, where S_a is the a^{th} Chebyshev polynomial. Now, as in the proof of Lemma 13.9, $f^{(n)}$ with a single strand encircling it (to be thought of as α in the annulus) is $(-A^{2(n+1)} - A^{-2(n+1)})f^{(n)}$. Hence the element required here is $S_a(-A^{2(n+1)} - A^{-2(n+1)})f^{(n)}$. This immediately gives the result using the remarks about Chebyshev polynomials after Lemma 13.2. □

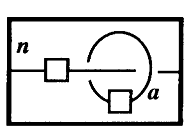


Figure 14.2

The first of these results is sometimes interpreted by saying that the operation of inserting a positive kink induces a linear map $\mathcal{S}(S^1 \times I) \to \mathcal{S}(S^1 \times I)$ and, with respect to this, each $S_n(\alpha)$ is an eigenvector with corresponding eigenvalue $(-1)^n A^{n^2+2n}$. A direct application of this result is the following:

Lemma 14.3. *Suppose A is a primitive* $4r^{\text{th}}$ *root of unity. Then*

$$< \omega >_{U_+} = \frac{G}{2A^{(3+r^2)}(A^2 - A^{-2})},$$

where G is the Gauss sum given by $G = \sum_{n=1}^{4r} A^{n^2}$.

PROOF. Recall that U_+ is the diagram of the unknot with one positive crossing,

$$\omega = \sum_{n=0}^{r-2} \Delta_n S_n(\alpha) \text{ and } \Delta_n = \frac{(-1)^n(A^{2(n+1)} - A^{-2(n+1)})}{A^2 - A^{-2}}.$$

So use of Lemma 14.1 to remove the kink in each $S_n(\alpha)$ shows that $< \omega >_{U_+}$ is

$$\sum_{n=0}^{r-2} \Delta_n^2 (-1)^n A^{n^2+2n} = (A^2 - A^{-2})^{-2} \sum_{n=0}^{r-2} (-1)^n A^{n^2+2n} (A^{2(n+1)} - A^{-2(n+1)})^2.$$

Now elementary manoeuvres of algebraic number theory (see [74] for example) show that the summation in the last term is $\frac{1}{2}\, A^{-(3+r^2)}(A^2 - A^{-2}) \sum_{n=1}^{4r} A^{n^2}$. □

The fact that reflecting diagrams induces in $\mathcal{S}(S^2)$ an interchange of A with A^{-1}, and that Δ_n is unaltered by such interchange, means that

$$< \omega >_{U_-} = \frac{-\overline{G}}{2A^{-(3+r^2)}(A^2 - A^{-2})}.$$

Thus

$$-\overline{G}G/4(A^2 - A^{-2})^2 =< \omega >_{U_+} < \omega >_{U_-}$$

and this has already been shown, in Lemma 13.7, to be $-2r(A^2 - A^{-2})^{-2}$. Thus $\overline{G}G = 8r$. In fact, when $A = e^{i\pi/2r}$, it can be shown that $G = 2\sqrt{2r}\, e^{i\pi/4}$.

It was remarked in Chapter 12 that a lens space has a surgery diagram that consists of a chain of unknotted simple closed curves, each with some framing, each simply linking the curve before it in the chain and the curve after it (except that the curves at the two ends of the chain only link one other curve). Calculation of the invariant of Chapter 13 involves evaluating the element of $\mathcal{S}(S^2)$ that arises from allocating ω to each component of the chain. That can be done by expanding each ω as $\sum_{n=0}^{r-2} \Delta_n S_n(\alpha)$ and using multilinearity, next changing all framings to zero by removing "kinks" using Lemma 14.1, and then removing components from the end of the chain using Lemma 14.2. Factors involving powers of $< \omega >_{U_+}$ and $< \omega >_{U_-}$ can be evaluated using Lemma 14.3. There results a formula that can be given to a computer for determination (see [104] for details). An extensive analysis of such a formula appears in [76]. Work on the formula shows, for example, that for all r the lens spaces $L_{65,8}$ and $L_{65,18}$ have the same invariant, but that the invariant is not a function of the fundamental group. The lens space invariants were also explored in [28] and [51]. The same method works if the unknotted components of the surgery diagram are linked not just in a linear chain but in a tree-like configuration. The 3-manifold then has the structure of the union of Seifert fibre spaces (see [104]).

It has already been intimated that it is expedient to renormalise the $SU_q(2)$ invariant discussed so far by replacing ω with $\mu\omega$, for some carefully chosen $\mu \in \mathbb{C}$. Now choose $\mu \in \mathbb{C}$ so that

$$\mu^{-2} = <\omega>_{U_+} <\omega>_{U_-} = <\omega>_U = \frac{-2r}{(A^2 - A^{-2})^2}$$

(quoting Lemma 13.7 in the last equality). This means that $<\mu\omega>_{U_+} = <\mu\omega>_{U_-}^{-1}$. The renormalisation of the invariant can then be written in terms of the signature of the linking matrix of the surgery link; it is this renormalisation, which will now be defined, that produces some elegant evaluations.

Definition 14.4. Suppose $r \geq 3$ and A is a primitive $4r^{th}$ root of unity. Let M be a closed oriented 3-manifold. Define the invariant $\mathcal{I}_A(M)$ by

$$\mathcal{I}_A(M) = <\mu\omega, \mu\omega, \ldots, \mu\omega>_D <\mu\omega>_{U_-}^{\sigma} \mu,$$

where σ is the signature of the linking matrix of a link diagram D that is a surgery diagram for M.

It follows at once that when A is a primitive $4r^{\text{th}}$ root of unity,

$$\mathcal{I}_A(S^3) = \frac{A^2 - A^{-2}}{\sqrt{-2r}} \qquad \text{and} \qquad \mathcal{I}_A(S^1 \times S^2) = 1.$$

This is because the empty diagram represents S^3, so $\mathcal{I}_A(S^3) = \mu$, this being the term inserted somewhat gratuitously at the end of the above definition. From the definition of μ, $<\mu\omega>_U^{-1} = \mu = (A^2 - A^{-2})/\sqrt{-2r}$. The diagram U, the zero-crossing diagram of the unknot, represents $S^1 \times S^2$, so $\mathcal{I}_A(S^1 \times S^2) = 1$.

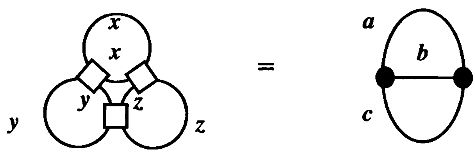


Figure 14.3

To make more general progress in calculating these 3-manifold invariants, it is necessary to develop some expertise in evaluating, in $\mathcal{S}(S^2)$, certain "diagrams" that consist of idempotents of various Temperley–Lieb algebras joined together by arcs in very simple ways. Consider first the diagram shown on the left of Figure 14.3. It consists of x parallel copies of a circle, y of another circle and z of a third with $f^{(x+y)}$, $f^{(y+z)}$ and $f^{(z+x)}$ inserted as shown. Let $\Gamma(x, y, z)$ be the element of $\mathcal{S}(S^2)$ that this diagram represents. This element will now be determined, for it will be important to know when $\Gamma(x, y, z)$ is and is not zero. In what follows, $\Delta_n!$ denotes $\Delta_n \Delta_{n-1} \Delta_{n-2} \ldots \Delta_1$, this being interpreted as 1 if n is -1 or zero.

Lemma 14.5.

$$\Gamma(x, y, z) = \frac{\Delta_{x+y+z}! \Delta_{x-1}! \Delta_{y-1}! \Delta_{z-1}!}{\Delta_{y+z-1}! \Delta_{z+x-1}! \Delta_{x+y-1}!}.$$

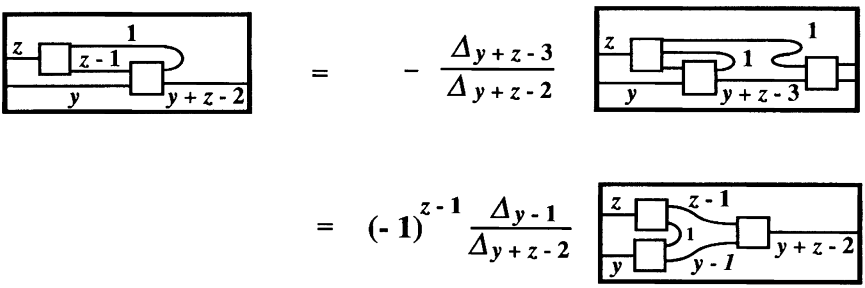


Figure 14.4

PROOF. Consider the equations depicted in Figure 14.4; as usual a symbol beside a line is a count of the number of parallel arcs that it represents. The first equality follows from the defining relation of Figure 13.6 for $f^{(y+z-1)}$ (together with $f^{(z)}e_{z-1} = 0$), and the second line follows by iterating the first line. Next, the defining relation for $f^{(y+z)}$ followed by a double application of Figure 14.4 produces the identity of Figure 14.5.

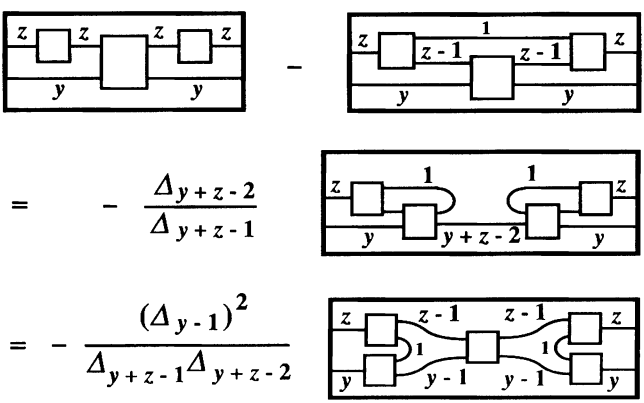


Figure 14.5

Now apply this last identity to Figure 14.3, using the formulae of Figures 13.4 and 13.5. The following recurrence relation results:

$$\Gamma(x, y, z) =$$
$$\Gamma(x, y, z-1)\Delta_{x+z}/\Delta_{x+z-1} - \Gamma(x+1, y-1, z-1)(\Delta_{y-1})^2/(\Delta_{x+z-1}\Delta_{y+z-2}).$$

This is ready for a verification of the given formula by induction on z. That formula is clearly true when $z = 0$, and inserting it into this recurrence relation reduces the proof to a demonstration of the equality

$$\Delta_{x+y+z}\Delta_{z-1} = \Delta_{x+z}\Delta_{y+z-1} - \Delta_{y-1}\Delta_{x}.$$

The truth of this can however easily be checked either directly from the formula for Δ_n or using a double induction on

$$\Delta_{x+y} = \Delta_x\Delta_y - \Delta_{x-1}\Delta_{y-1}.$$ □

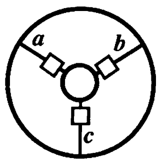


Figure 14.6

Consider the skein space of the disc D with $a + b + c$ specified points in its boundary. The points are partitioned into three sets of a, b and c (consecutive) points. The effect of adding the idempotents $f^{(a)}$, $f^{(b)}$ and $f^{(c)}$ just outside every diagram in such a disc with specified points (and so slightly enlarging the disc), is to map the skein space of the disc into a subspace of itself. That subspace will be be denoted $T_{a,b,c}$. Thus $T_{a,b,c}$ is spanned by all diagrams inserted into the inner disc of Figure 14.6. The dimension of $T_{a,b,c}$ is either one or zero, for the only chance of obtaining a non-zero skein element on inserting a diagram without crossings into Figure 14.6 is when the element obtained is a multiple of that on the left of Figure 14.7. This element, if it exists, will be denoted $\tau_{a,b,c}$. (The insertion of any other zero-crossing diagram into Figure 14.6 always gives zero on interacting with the idempotents.) For $\tau_{a,b,c}$ to exist, it is necessary that there should be non-negative integers x, y and z defined by $a = y + z$, $b = z + x$ and $c = x + y$. This occurs precisely when a, b and c are admissible in the following sense:

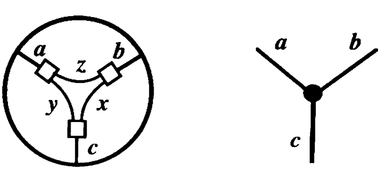


Figure 14.7

Definition 14.6. The triple (a, b, c) of non-negative integers will be called admissible if $a + b + c$ is even, $a \leq b + c$, $b \leq c + a$ and $c \leq a + b$.

When (a, b, c) is admissible, it is easy to see that $\tau_{a,b,c}$ is not the zero element of $T_{a,b,c}$ by considering the **1**-terms of the expansions of the three idempotents as sums of base elements of the various Temperley–Lieb algebras. When (a, b, c) is admissible, define

$$\theta(a, b, c) = \Gamma(x, y, z),$$

where non-negative integers x, y and z are defined in the above way. In the diagrams that follow, a triad consisting of a black dot with three arcs emerging from it labelled a, b and c, will be an abbreviation for the triad diagram $\tau_{a,b,c}$ (see Figure 14.7); it is always then to be assumed that (a, b, c) is admissible. Note that $\theta(a, b, c)$ is the evaluation of the diagram consisting of two black dots joined together by three simple disjoint arcs labelled a, b, and c; see Figure 14.3. It should be observed that each arc emerging from a black dot is automatically decorated with the relevant idempotent. A useful identity that uses this notation is shown in Figure 14.8.

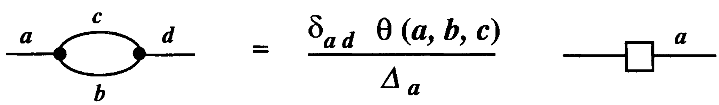


Figure 14.8

The Krönecker delta function occurs because if, say, $a > d$ then the copy of $f^{(a)}$ in the left-hand triad must, in the expansion of the remainder as a sum of diagrams with no crossing, always abut some "e_i", (in each such diagram, some curve leaving the left must return to the left). When $a = d$, the left diagram must be some scalar multiple of $f^{(a)}$, and the multiplier is readily found by joining, in the plane, points on the left of the diagram to points on the right.

Suppose that D is in S^2 and D' is the disc complementary to D with the same specified $a + b + c$ boundary points. Taking unions of diagrams in the two discs induces a bilinear form $\mathcal{S}D \times \mathcal{S}D' \to \mathcal{S}(S^2) = \mathbb{C}$, and using this, $\tau_{a,b,c}$ corresponds to the element $\tau^*_{a,b,c}$ of the dual space to $\mathcal{S}D'$. In this way $T_{a,b,c}$ can be regarded as a space $T^*_{a,b,c}$ of linear maps of the skein outside D; an element of $T_{a,b,c}$ is thus a "map of outsides". Strictly, $T^*_{a,b,c}$ is the quotient of $T_{a,b,c}$ by the kernel of the bilinear form. This is almost unnecessary sophistry for generic A, but is significant when A is a root of unity.

Lemma 14.7. *Let (a, b, c) be admissible and let A be a primitive $4r^{\text{th}}$ root of unity. Then $\tau^*_{a,b,c}$ is non-zero if and only if $a + b + c \leq 2(r - 2)$.*

PROOF. $\mathcal{S}D'$ has a base consisting of all diagrams in D' with no crossing. For all but one of these diagrams there is an arc from a point of one of the three specified subsets (for example, that with a points) to another point of the same subset. As usual (using $f^{(a)}e_i = 0$), $\tau^*_{a,b,c}$ annihilates such an element. There remains to consider the base element of SD' that consists of z arcs from the first boundary subset to the second such subset, x from the second to the third and y from the third to the first. Of course, $T^*_{a,b,c}$ maps this element to $\Gamma(x, y, z)$. It follows from Lemma 14.5 that as $x + y + z$ increases, this is non-zero until $\Delta_{x+y+z}! = 0$ and that this occurs when $x + y + z = r - 1$. □

Definition 14.8. A triple (a, b, c) of non-negative integers will be called r-admissible if it is admissible and $a + b + c \leq 2(r - 2)$.

The substance of the last result is that for A a primitive $4r^{\text{th}}$ root of unity, the space of maps $T^*_{a,b,c}$ is zero unless (a, b, c) is r-admissible, and in that case it has dimension 1.

These ideas are now to be generalised to the disc with an even number, $(a + b + c + d)$, of points specified in its boundary, partitioned consecutively into a , b , c and d points. Let $Q_{a,b,c,d}$ denote the subspace of the skein space of such a disc that comes from placing the idempotents $f^{(a)}$, $f^{(b)}$, $f^{(c)}$ and $f^{(d)}$ just outside every diagram that generates this space (see Figure 14.9).

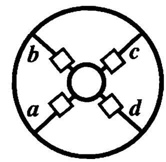


Figure 14.9

Lemma 14.9. *Suppose A is not a root of unity. A base for $Q_{a,b,c,d}$ is the set of elements as in Figure 14.10 (the boundary of the disc is not shown), where j takes all values such that (a, b, j) and (c, d, j) are both admissible.*

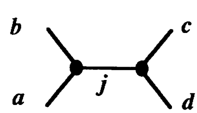


Figure 14.10

PROOF. Note that the proposed base elements each consist of two triads glued together; there is an $f^{(j)}$ on the central line. Certainly $Q_{a,b,c,d}$ is spanned by all elements of the form shown in Figure 14.11, where the lines all represent multiple parallel arcs, for, as usual, any other diagrams interact with the idempotents to give zero. Without loss of generality, it is assumed that $b + d \geq a + c$, and it is clear that the diagonal line represents $\frac{1}{2}\{b + d - a - c\}$ parallel arcs. The number of arcs represented by the other lines can vary. Suppose there are j arcs crossing the vertical dotted line. In the Temperley Lieb algebra TL_j, recall that $\mathbf{1} - f^{(j)}$ is in the ideal generated by the e_i. Thus a diagram with j arcs crossing the dotted line can be replaced with a linear sum of diagrams, one with j arcs containing an $f^{(j)}$ and others that cross the vertical line fewer than j times (coming from the e_i). Thus, by induction on the number of arcs crossing the vertical line, it is seen that the given elements span the space.

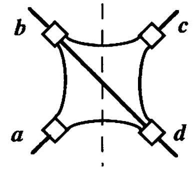


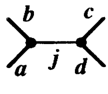
Figure 14.11

Gluing a triad $\tau_{c,d,k}$ onto the right of the disc under consideration produces a linear map from $Q_{a,b,c,d}$ to $T_{a,b,k}$. Of course this operation maps the element of Figure 14.10 to $\delta_{j,k}\ (\theta(c, d, k)/\Delta_k)\tau_{a,b,k}$, using the formula of Figure 14.8. When $j = k$, this is non-zero, and so the proposed base elements are indeed independent. □

Lemma 14.10. *Suppose A is a primitive $4r^{th}$ root of unity. A base for $Q^*_{a,b,c,d}$ (this being $Q_{a,b,c,d}$ regarded as maps of diagrams outside the disc) is the set of elements as in Figure 14.10 where j takes all values such that (a, b, j) and (c, d, j) are both r-admissible.*

PROOF. The proof that the given elements span is the same as in Lemma 14.9 with a small modification. Now, $f^{(n)}$ does not exist for $n \geq r$. However, $f^{(r-1)}$ is the zero map of outsides. Thus working in this dual context, any diagram as in the above proof, with at least $(r - 1)$ arcs crossing the dotted vertical line, can be replaced by a sum of diagrams with fewer such arcs. Further, any triad encountered that is not r-admissible may be discarded, since it represents the zero map. The proof of independence is essentially the same as before (though the map used now goes from and to the dual spaces). □

The bases for $Q_{a,b,c,d}$ and $Q^*_{a,b,c,d}$ given in the last two lemmas have a "horizontal" bias. There is, by symmetry, a base in each case with a "vertical" bias. The change-of-base equation is depicted in Figure 14.12, where the summation is over all i for which the triples (b, c, i) and (a, d, i) are admissible (or, respectively, r-admissible). The terms $\begin{Bmatrix} a & b & i \\ c & d & j \end{Bmatrix}$ of this change-of-base matrix are sometimes called $6j$-symbols. The $6j$-symbols can be evaluated in terms of a diagrammatic presentation by adjoining a triad $\tau_{a,d,k}$ beneath the diagrams of both sides of the equation of Figure 14.12. This produces zero for every term on the right hand side except the k^{th} term. That term becomes (with no summation convention) $\begin{Bmatrix} a & b & k \\ c & d & j \end{Bmatrix} \theta(a, d, k)\Delta_k^{-1}\tau_{b,c,k}$. Now place a copy of $\tau_{b,c,k}$ on the outside of



$$= \sum_i \left\{ \begin{matrix} a & b & i \\ c & d & j \end{matrix} \right\}$$

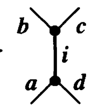


Figure 14.12

both sides of the equation. It transpires (keeping track of what has happened to the left hand side) that the labelled diagram in the shape of the edges of a tetrahedron as in Figure 14.13 is equal to $\left\{ \begin{matrix} a & b & k \\ c & d & j \end{matrix} \right\} \theta(a, d, k)\theta(b, c, k)\Delta_k^{-1}$. A lengthy closed formula is known for this labelled tetrahedral diagram and hence for the $6j$-symbol, but it is not very attractive (except to a computer). It is quoted in [62], a proof is in [12], and the form of the answer is known from quantum field theory [69].

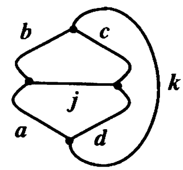
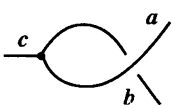


Figure 14.13

One more general formula of use is the identity shown in Figure 14.14. The 1-dimensional nature of $T_{a,b,c}$ asserts that the element on the left of Figure 14.14 is some multiple of $\tau_{a,b,c}$. That the multiplier is $(-1)^{(a+b-c)/2}A^{a+b-c+((a^2+b^2-c^2)/2)}$ is left as a fairly easy exercise.



$$= (-1)^{\frac{a+b-c}{2}} A^{a+b-c+\frac{a^2+b^2-c^2}{2}}$$

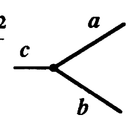


Figure 14.14

Suppose now it is desired to evaluate in $\mathcal{S}(S^2) = \mathbb{C}$, the complex number represented by a link diagram in which each component is replaced by some $S_n(\alpha)$ as encountered in the definition of the 3-manifold invariants. This can be thought of as a diagram with segments representing multiple parallel arcs with some $f^{(n)}$ included (and note that because $f^{(n)}f^{(n)} = f^{(n)}$, as many copies of $f^{(n)}$ as might be desired may be inserted around any component of the link). Near each crossing, two such parallel multiple arcs with labels a and b can be replaced by a linear sum of the above base elements of $Q^+_{a,b,b,a}$, and then the crossing can be removed in each summand using the equality of Figure 14.14. What emerges is a linear sum of weighted trivalent graphs in S^2 with a black dot at each vertex. Again at the expense of taking linear sums, the graphs can be changed using the $6j$-symbol equation to reduce the number of edges of the graph around a region of the graph's complement in S^2. When a region is bounded by just two edges, it can be removed using Figure 14.8. This eventually simplifies completely all the graphs to be considered. Although in principle such a method of calculation will always work, in general it desperately needs computer assistance.

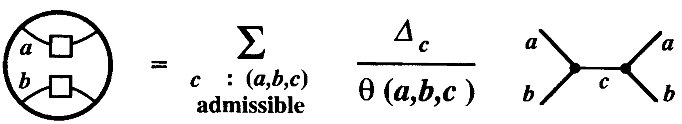


Figure 14.15

Sometimes the $SU_q(2)$ 3-manifold invariant does have a compact formulation with an elegant method of calculation. An example, which will now be described, is the manifold consisting of the product of a closed orientable surface and a circle.

In Figure 14.15 is a picture of a strands with an $f^{(a)}$ inserted beside b strands with an $f^{(b)}$. This can be regarded as an element of $Q_{b,a,a,b}$ or of $Q^*_{b,a,a,b}$. In either case, it must be expressible as a linear sum of the basis elements. The summation is over all c for which (a, b, c) is admissible (or r-admissible), and the coefficients of the sum are determined by adjoining the triad $\tau_{a,b,c}$ and using Figure 14.8.

Lemma 14.11. *In $\mathcal{S}(S^1 \times I)$, $S_a(\alpha)S_b(\alpha) = \sum_c S_c(\alpha)$ where the summation is over all c such that (a, b, c) is admissible. If A is a primitive $4r^{\text{th}}$ root of unity regarding both sides of the equation as maps of outsides (of immersed annuli as in Chapter 13), $S_a(\alpha)S_b(\alpha) = \sum_c S_c(\alpha)$, where now the sum is over all c such that (a, b, c) is r-admissible.*

PROOF. In fact the first part of this lemma is almost immediate. This is because it is a result on Chebyshev polynomials that $S_a(x)S_b(x) = \sum_c S_c(x)$, the sum being over all c such that (a, b, c) is admissible. This follows by induction on b. However, another proof is shown in Figure 14.16, where the result of Figure 14.15 is first applied at the top of the diagram and then the result of Figure 14.8 is applied at the bottom. The advantage of this alternative proof is that it also works in the r-admissible case as well. □

Figure 14.17 shows an element of $\mathcal{S}(S^1 \times I)$ that will temporarily be denoted β. Regarding β as a map of outsides, expanding one of the ω's as $\sum_{a=0}^{r-2} \Delta_a S_a(\alpha)$ and using Figure 14.15, a summation expression for β is obtained. This is also depicted in Figure 14.17, where the a's at the top are to be understood to be joined to those at the bottom by arcs that encircle the annulus. But, by Lemma 13.9, if A is a primitive $4r^{\text{th}}$ root of unity, then the only non-zero contribution to that expression

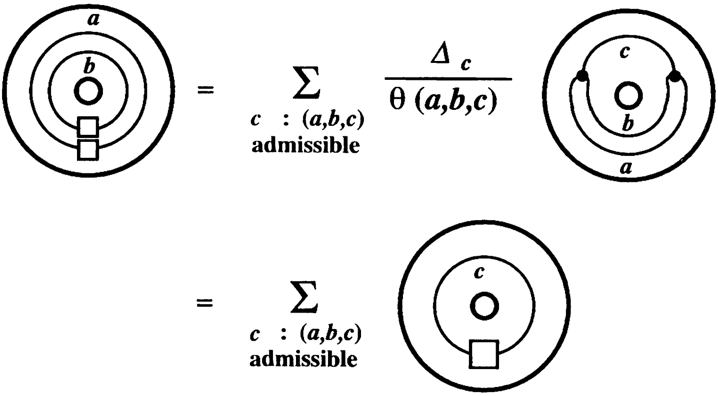


Figure 14.16

is when $c = 0$. Thus as maps of outsides (recall that $\theta(a, a, 0) = \Delta_a$),

$$\beta = \sum_{a=0}^{r-2} < \omega >_U (S_a(\alpha))^2.$$

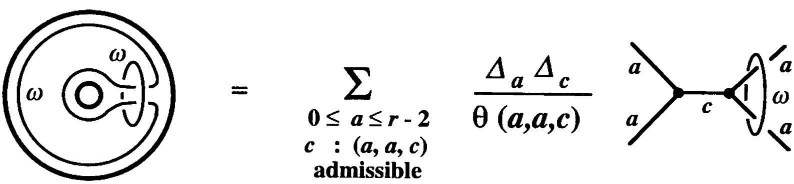


Figure 14.17

Theorem 14.12. *Let F_g be the closed orientable surface of genus g and let A be a primitive $4r^{\text{th}}$ root of unity, $r \geq 3$. Then $\mathcal{I}_A(S^1 \times F_g)$ is an integer. It is $r - 1$ when $g = 1$. Otherwise it is the number of ways of labelling the $3(g - 1)$ edges of the graph of Figure 14.18 with integers a_i, $0 \leq a_i \leq r - 2$, so that the three labels at any node form an r-admissible triple.*

PROOF. The 3-manifold $S^1 \times F_g$ is obtained by surgery on a link that consists of g copies of the Borromean rings summed together on one component, each component having the zero framing. (Proving this is an interesting exercise.) A diagram D for such a link is obtained by taking g annuli, each containing a link as on the left of Figure 14.17, threading an unknotted closed curve through these annuli and then taking the resultant diagram of $2g + 1$ components. Then $< \omega, \omega, \dots, \omega >_D = < \omega, \beta^g >_H$, where H is just the two-crossing diagram of the simple Hopf link of two curves. Thus, as the signature of the linking matrix of

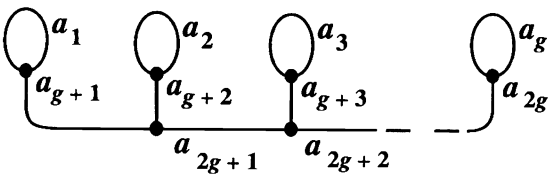


Figure 14.18

this link is zero (and using $< \mu\omega >_U = \mu^{-1}$),

$$\mathcal{I}_A(S^1 \times F_g) = \mu^{2g+2} < \omega, \beta^g >_H .$$

Now $\beta = \sum_{a=0}^{r-2} \mu^{-2}(S_a(\alpha))^2$, so β^g can be expressed as a sum of the $S_n(\alpha)$'s by Lemma 14.11. Then, by Lemma 13.9, $< \omega, \beta^g >_H$ is $\mu^{-2-2g}N$, where N is the number of times $S_0(\alpha)$ appears in the expansion (by Lemma 14.11) for $(\sum_{a=0}^{r-2}(S_a(\alpha))^2)^g$ as a sum of the $S_n(\alpha)$'s. This N is the number of r-admissible labellings of the edges of Figure 14.18. □

The last result can, in general terms, be anticipated. It was shown in [13] that these $SU_q(2)$ invariants can be regarded as emanating from a topological quantum field theory. These theories will *not* be described here in any detail (but see [4]). Roughly, such a theory is a functor from the category of oriented surfaces and cobordisms to that of vector spaces and linear maps that sends disjoint unions to tensor products. It follows from such an abstract formulation [4] that the invariant of the mapping torus of an automorphism of a surface F is the trace of some linear map. The invariant for $S^1 \times F$ is the trace of the *identity* map, and that is certainly an integer.

When the surface has genus equal to 1 or 2, it is easy to make a count of this integer from the theorem. The results obtained are

$$\mathcal{I}_A(S^1 \times S^1 \times S^1) = (r-1), \quad \mathcal{I}_A(S^1 \times F_2) = \frac{r^3 - r}{6}.$$

Working from the same surgery diagram, with A still a primitive $4r^{\text{th}}$ root of unity, $r \geq 3$, it can also be shown [86] that

$$\mathcal{I}_A(S^1 \times F_g) = (-2r)^{g-1} \sum_{a=0}^{r-2} (A^{2(a+1)} - A^{-2(a+1)})^{2-2g}.$$

It is surprising that this last expression must be an integer as, indeed, it has just been proved to be.

The mapping torus of the automorphism of $S^1 \times S^1$ that reverses the sign of every element of $H_1(S^1 \times S^1)$ has a surgery diagram as shown in Figure 14.19. It is clearly very similar to the diagram of the Borromean rings with zero framings, considered above, for $S^1 \times S^1 \times S^1$. It is an easy exercise using the above methods to show

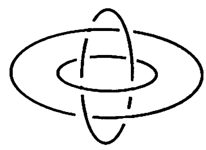


Figure 14.19

that this manifold also has $\mathcal{I}_A = (r - 1)$. That can, in fact, also be deduced from considerations of the topological quantum field theory. Of course this manifold is not equal to $S^1 \times S^1 \times S^1$; its first homology is $\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$. Thus two manifolds can have all $SU_q(2)$ invariants the same and yet be distinguished by their first homology groups.

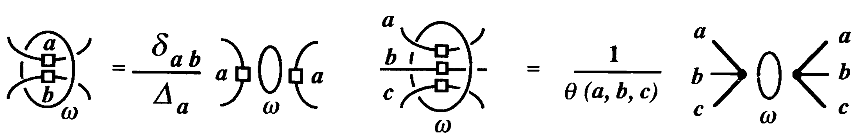


Figure 14.20

Less obvious, whether by calculation or philosophy, is the fact that the 3-manifolds (described by Kauffman [62]), obtained by surgery on the framed links shown in the diagrams at the top of Figure 14.21 and the top of Figure 14.22, also have the same integer invariants. To see this, first note the equalities shown in Figure 14.20, applicable when A is a primitive $4r^{\text{th}}$ root of unity. The first of these identities has, essentially, already been used. It follows at once from Figure 14.15 and Lemma 13.9. The second one follows by using Figure 14.15 twice and then Lemma 13.9, but note that the right-hand side is to be interpreted as zero unless (a, b, c) is an r-admissible triple. The diagram D at the top of Figure 14.21 has three zero-framed components. With appropriate orientations it has linking matrix

$$\begin{pmatrix} 0 & 3 & 0 \\ 3 & 0 & 3 \\ 0 & 3 & 0 \end{pmatrix},$$

and this matrix has signature zero. The matrix represents the first homology of the 3-manifold M obtained by surgery on this diagram, and so this homology group is $\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}$.

The remainder of Figure 14.21 considers the above identities (together with that of Figure 14.14) applied to the diagram, with ω decorating the left and right components and $\Delta_a S_a(\alpha)$ decorating the other component. The result is $(\cdot\ \omega\ \cdot\mu)^2\ \mu^{-1}$ when (a, a, a) is an r-admissible triple and zero otherwise.

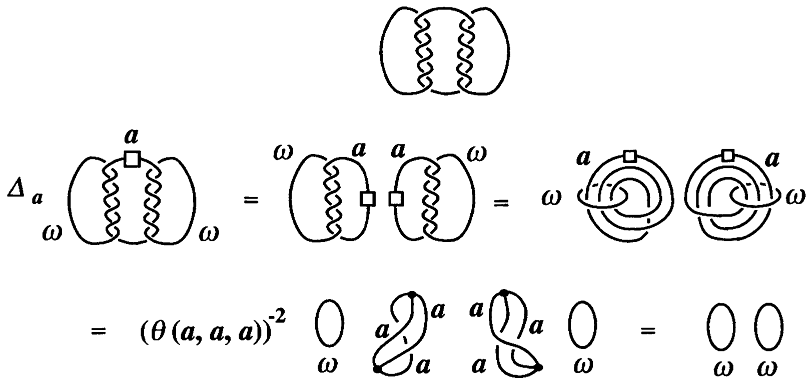


Figure 14.21

Summing this over all a, $0 \leq a \leq r-2$, shows that $\mu^4 < \omega, \omega, \omega >_D$ is the number of a such that (a, a, a) is r-admissible. Thus $\mathcal{I}_A(M)$ is equal to the number of such a.

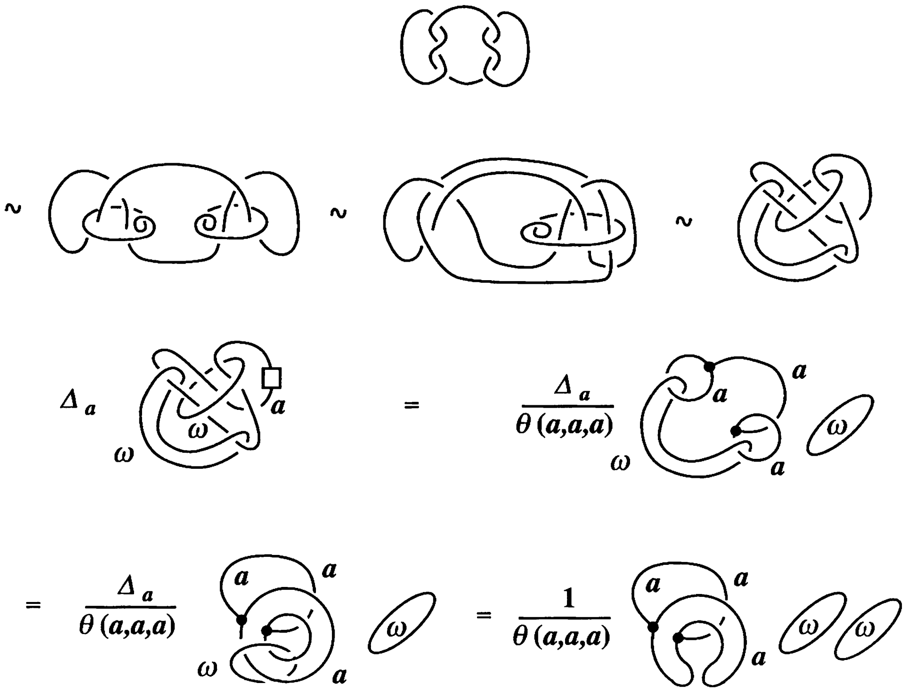


Figure 14.22

The manifold M obtained by surgery on the zero-framed reef knot shown at the top of Figure 14.22 has $\mathbb{Z}$ as its first homology group. The second line of Figure 14.22 shows the effect of changing this to a different diagram by some adroit Kirby moves. The first moves introduce a -1-framed unknot and a $+1$-framed unknot and slide parts of the knot over them, then the first of those unknots is slid over the second; the last move is an isotopy. Figure 14.22 then analyses the resulting diagram, with ω decorating two components and $\Delta_a S_a(\alpha)$ decorating the third component. Again, summing this over all a shows that the result of evaluating ω decorating the original reef knot diagram is the product of $(<\omega>_U)^2\{<\omega>_{U_+} <\omega>_{U_-}\}^{-1}$ with the number of a such that (a, a, a) is r-admissible. Hence again that number is the invariant $\mathcal{I}_A(M)$.

One more example of these "recombination techniques" will conclude this chapter. It is not actually concerned with calculating a 3-manifold invariant, but with calculating the Jones polynomial of a torus knot. The method used here is modelled on a paper by P. M. Strickland [117]. In what follows, A is a generic complex number.

Theorem 14.13. *If p and q are coprime positive integers, then the Jones polynomial of the (p, q)-torus knot is*

$$t^{(p-1)(q-1)/2}(1 - t^2)^{-1}(1 - t^{p+1} - t^{q+1} + t^{p+q}).$$

PROOF. Consider the diagram of Figure 14.23, which shows p arcs traversing a rectangle. Suppose q copies of this are placed side by side and the result is closed up by joining the p points on the left to those on the right, using p crossing-free arcs encircling an annulus $S^1 \times I$ to form a diagram $T(p, q)$ in that annulus. It is desired to evaluate this diagram in the skein of the annulus in terms of the base elements $\{S_n(\alpha)\}$. Then, placing the annulus in the plane will at once give a value for the Jones polynomial of the (p, q)-torus knot. For some fixed k (which will here later be taken to be 1), consider p arcs side by side in a diagram, each labelled with an $f^{(k)}$ so that as usual each arc represents k parallel arcs with the idempotent inserted. Applying the identity of Figure 14.15 $(p - 1)$ times shows that this is of the form of Figure 14.24, where the coefficient $\Lambda(i_1, i_2, \ldots, i_{p-2}, a)$ is the quotient of a product of Δ's by a product of θ's and the summation is over all $(i_1, i_2, \ldots, i_{p-2}, a)$ that produce an admissible triple at each vertex of the diagram.

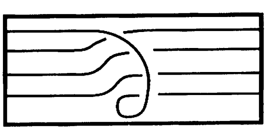


Figure 14.23

$$\sum \Lambda\,(i_1, i_2, \ldots, i_{p-2}, a)$$

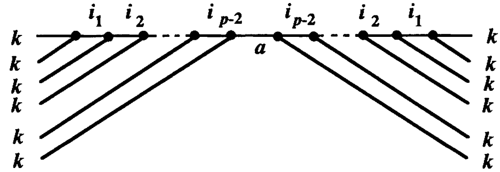


Figure 14.24

The diagram of Figure 14.25 is, of course, a multiple of $f^{(a)}$; let it be denoted, without any summation convention, by

$$\{\Lambda(i_1, i_2, \ldots, i_{p-2}, a)\Lambda(j_1, j_2, \ldots, j_{p-2}, a)\}^{-1/2} M(a)^{\mathbf{i}}_{\mathbf{j}} f^{(a)}.$$

The $\{M(a)^{\mathbf{i}}_{\mathbf{j}}\}$ will be regarded as a matrix $M(a)$ with rows and columns indexed by $\mathbf{i}$ and $\mathbf{j}$, each representing a multi-suffix $(i_1, i_2, \ldots, i_{p-2})$ or $(j_1, j_2, \ldots, j_{p-2})$.

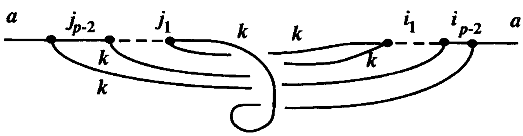


Figure 14.25

Suppose $T(p, q)^{(k)}$ is the diagram $T(p, q)$ decorated by $S_k(\alpha)$. Suppose that in $T(p, q)^{(k)}$, between consecutive occurrences of (the k-weighted) Figure 14.23, the p parallel strings are "combined" as in Figure 14.24. Note that terms with distinct values of a compose to give zero. The terms with a given value of a combine to be the trace of the matrix $(M(a))^q$ multiplying $S_a(\alpha)$. This follows just from the above notation.

Now, $((M(a))^p)^{\mathbf{i}}_{\mathbf{j}} f^{(a)}$ is $\{\Lambda(i_1, i_2, \ldots, i_{p-2}, a)\Lambda(j_1, j_2, \ldots, j_{p-2}, a)\}^{1/2}$ times the diagram of Figure 14.26, and that diagram in turn is zero unless $\mathbf{i} = \mathbf{j}$ and is then, by Lemma 14.1 and Figure 14.8, equal to

$$\{\Lambda(i_1, i_2, \ldots, i_{p-2}, a)\Lambda(j_1, j_2, \ldots, j_{p-2}, a)\}^{-1/2}(-1)^a A^{a^2+2a} f^{(a)}.$$

Hence $(M(a))^p$ is $(-1)^a A^{a^2+2a}$ times an identity matrix, of size dependent on the number of admissible labellings of the diagram of Figure 14.24.

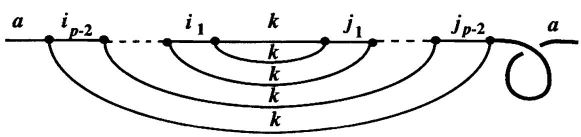


Figure 14.26

This means that the eigenvalues of $M(a)$ are all p^{th} roots of $(-1)^a A^{a^2+2a} = (-A)^{a^2+2a}$ and, of course, the trace is the sum of those eigenvalues. Each eigenvalue can be regarded as $\xi_j \rho$, where ρ is some fixed p^{th} root of $(-A)^{a^2+2a}$ and ξ_j is some p^{th} root of unity. The trace of $M(a)$ is, as explained above, the coefficient of $S_a(\alpha)$ in the expansion of $T(p,1)^{(k)}$ in terms of the $S_n(\alpha)$. Suppose (see below) that ρ can be chosen so that $\sum_j \xi_j$ is an integer; it will be denoted N_a. Then, if p and q are coprime, $\sum_j (\xi_j)^q$ is also equal to N_a (because if a primitive p^{th} root of unity is a zero of a polynomial with integer coefficients, then any other primitive p^{th} root of unity is a zero of the same polynomial). From this it emerges that $T(p,1)^{(k)} = \sum_a N_a(-A)^{(a^2+2a)/p} S_a(\alpha)$ for some integers $\{N_a\}$ and then, using the same integers, $T(p,q)^{(k)} = \sum_a N_a(-A)^{q(a^2+2a)/p} S_a(\alpha)$. Thus it remains to evaluate $T(p,1)^{(k)}$, at least when $k = 1$.

The element $T(p,1)$ of $\mathcal{S}(S^1 \times I)$ contains a copy of Figure 14.23. The process of removing the top crossing of Figure 14.23 by using the defining skein relation and removing a kink leads at once to the recurrence relation

$$T(p,1) = A\alpha T(p-1,1) - A^2 T(p-2,1).$$

Letting $x_p = A^{-p}T(p,1)$, this becomes $x_p = \alpha x_{p-1} - x_{p-2}$. This is, of course, the recurrence relation which has as solution the Chebyshev polynomials $S_n(\alpha)$. Now $x_1 = -A^2\alpha$ and $x_0 = -A^2 - A^{-2}$. Thus $x_p = -A^2 S_p(\alpha) + A^{-2}S_{p-2}(\alpha)$. Hence, for $k = 1$, ρ can indeed be chosen for each a so that $\sum_j \xi_j$ is an integer N_a: when $a = p$, choose $\rho = (-A)^{p+2}$ and then $N_p = (-1)^{p+1}$, and when $a = p-2$, choose $\rho = (-A)^{p-2}$ and then $N_{p-2} = (-1)^p$, otherwise $N_a = 0$. Hence

$$T(p,q)^{(1)} = (-1)^{p+1}(-A)^{q(p+2)}S_p(\alpha) + (-1)^p(-A)^{q(p-2)}S_{p-2}(\alpha).$$

Placing $S^1 \times I$ in the standard way in the plane sends $T(p,q)^{(1)}$ to the element

$$(-1)^{p+1}(-A)^{q(p+2)}\Delta_p + (-1)^p(-A)^{q(p-2)}\Delta_{p-2}$$

in the skein of the plane. This planar diagram is a diagram with writhe pq of the (p,q) torus knot. To obtain the Jones polynomial, the above expression must be multiplied by $(-A)^{-3pq}$ to account for the writhe, then by $(-A^{-2} - A^2)^{-1}$ because the Jones polynomial is the coordinate of the skein element with the zero-crossing unknot as base, and then the substitution $t = A^{-4}$ must be made. Thus, with $t = A^{-4}$, the Jones polynomial of the knot is

$$(A^{-4}-A^4)^{-1}(-A)^{-2pq}\Big[-A^{2q}(A^{2(p+1)}-A^{-2(p+1)})+A^{-2q}(A^{2(p-1)}-A^{-2(p-1)})\Big].$$

This is

$$(A^{-4} - A^4)^{-1}(-A)^{-2pq}A^{2(p+1)}A^{2q}\Big[-1 + A^{-4(p+1)} + A^{-4(q+1)} - A^{-4(p+q)}\Big]$$

$$= -(1 - A^{-8})^{-1}A^{-2(pq-p-q+1)}\Big[-1 + A^{-4(p+1)} + A^{-4(q+1)} - A^{-4(p+q)}\Big],$$

and this is the stated result. □

This chapter has really been about the calculation of $SU_q(2)$ invariants of framed coloured links in S^3. Such a link is a framed link L together with a non-negative integer $n(i)$ assigned to each component L_i. The coloured invariant is then the element of $\mathbb{Z}[A^{-1}, A]$ that results from decorating every L_i with $S_{n(i)}(\alpha)$ and evaluating the result in $\mathcal{S}(\mathbb{R}^2)$. Applications have been found for these link invariants. They are used in [96], [138] and [71] to give information about the tunnel number of a knot. (The *tunnel number* is the minimal number of arcs that can be embedded in the knot exterior X, with each arc meeting ∂X at its end points, so that X less a regular neighbourhood of the arcs is a handlebody.) They have also been employed in [73] to give information about a generalised unknotting operation.

Exercises

1. Prove the formula displayed in Figure 14.14.
2. Let D be a diagram in S^2 of the p-framed $(p, 2)$ torus knot. Calculate $\langle S_n(\alpha)\rangle_D$, where S_n is the usual Chebyshev polynomial.
3. Suppose that D is the usual three-crossing diagram of the trefoil knot and that surgery prescribed by this diagram gives a 3-manifold M. Calculate the invariant $\mathcal{I}_A(M)$ when $A = \exp(\pi i/10)$.
4. Let $A = \exp(\pi i/2r)$. If M and $\overline{M}$ are the same closed connected 3-manifold but with opposite orientations, show that $\mathcal{I}_A(\overline{M})$ is the complex conjugate of $\mathcal{I}_A(M)$.

   Let $M_1 + M_2$ be the connected sum of two oriented closed connected 3-manifolds M_1 and M_2. (This sum is formed by removing a 3-ball from each manifold and identifying the 2-sphere boundaries together so that orientations match up.) Show that
   $$\mu\mathcal{I}_A(M_1 + M_2) = \mathcal{I}_A(M_1)\mathcal{I}_A(M_2).$$
5. Let A be a primitive $4r^{\text{th}}$ root of unity. Find, in the r-admissible situation, expressions for all the $6j$-symbols when $r = 4$.
6. Let A be a primitive $4r^{\text{th}}$ root of unity. Suppose that ω' is another element of $\mathcal{S}(S^1 \times I)$ with the property of invariance under type 2 Kirby moves that is described (for ω) in Lemma 13.5. Let μ' be defined so that $(\mu')^{-2} = \langle\omega'\rangle_U$ and suppose that this is non-zero. Suppose that ω' and μ' are used to define an invariant $\mathcal{I}'_A(M)$ of closed, connected, oriented 3-manifolds M exactly as in Definition 14.4. By considering
   $$\langle \mu\mu'\omega\omega', \mu\mu'\omega\omega', \ldots, \mu\mu'\omega\omega'\rangle_D$$
   for a link diagram D, determine the relationship between $\mathcal{I}'_A(M)$ and $\mathcal{I}_A(M)$.
7. Show that any compact connected oriented 3-manifold with boundary a torus can be obtained by surgery on a framed link in a solid torus. [Glue a solid torus to the boundary and use the surgery result for closed 3-manifolds.] Suppose that X_1 and X_2 are knot exteriors and $h : \partial X_1 \to \partial X_2$ is a homeomorphism. Let $-h$ be the composition of h and $-\text{id} : \partial X_1 \to \partial X_1$, where $-\text{id}$ is a homeomorphism that sends longitude to longitude and meridian to meridian but reverses the directions of them both. Show that for every primitive $4r^{\text{th}}$ root of unity A,
   $$\mathcal{I}_A(X_1 \cup_h X_2) = \mathcal{I}_A(X_1 \cup_{-h} X_2).$$

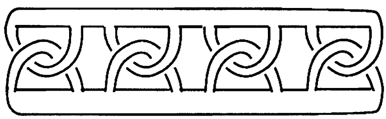
8. Let A be a primitive $4r^{\text{th}}$ root of unity. In Lemma 14.10, a base is described for the space of skeins of a disc, with points in its boundary partitioned into four sets and with an idempotent (of the relevant Temperley–Lieb algebra) adjacent to each set, when that space is regarded as a dual space (as "maps of outside diagrams"). Generalise this from a partition into four sets to a partition into n sets of points, finding a base corresponding to labelled trivalent graphs with r-admissibility for the labels at every vertex.

9. Let A be a primitive $4r^{\text{th}}$ root of unity. Let

$$N = \{\phi \in \mathcal{S}(S^1 \times I) : \langle \phi, \psi \rangle_H = 0 \text{ for all } \psi \in \mathcal{S}(S^1 \times I)\}.$$

Here, again, H is a two-crossing diagram of a non-trivial (Hopf) link. Show that the dimension of the quotient space $\mathcal{S}(S^1 \times I)/N$ is $r - 1$, and find for this space two bases, represented by sets $\{\beta_i\}$ and $\{\gamma_i\}$ of elements of $\mathcal{S}(S^1 \times I)$, such that $\langle \beta_i, \gamma_j \rangle_H = \delta_{i,j}$.

10. Repeat the previous exercise with a g-holed disc replacing the annulus, the bilinear form being given by placing skein space elements in the diagram of two linked g-holed discs, shown below, and evaluating the result in the skein space of the plane. The dimension of the quotient space is to be shown to be the integer $\mathcal{I}_A(S^1 \times F_g)$ obtained in Theorem 14.12. Show that a base is all labelled diagrams in the g-holed disc, of the form of Figure 14.18 with r-admissibility at each vertex.



15

Generalisations of the Jones Polynomial

The Jones polynomial invariant of oriented links has already been expressed by means of a so-called skein formula in Proposition 3.7. and a similar, but different, formula was given for the Conway polynomial in Theorem 8.6. It will now be shown that those are two instances of a more general polynomial invariant in two indeterminates, sometimes called the HOMFLY polynomial ([31], [90], [106]). This is one of *two* two-variable generalisations of the Jones invariant. The other is the Kauffman polynomial invariant ([60], [58], [16], [45]). The main aim of this chapter is to show that these two invariants exist—that is, that they are indeed well defined. These proofs of existence are harder than the one given for the Jones polynomial in Chapter 3.

The simple defining formulae of these invariants are in the statements of the next two theorems and the proofs of the two are very similar. First, however, there follows a preparatory and slightly technical result of planar geometry. It investigates the way in which a lens-shaped region R of the plane is divided into regions by a collection of transversals $\{t_i\}$.

Lemma 15.1. *Suppose that p and q are two arcs in $\mathbb{R}^2$ meeting only at their end points A and B, and let R be the compact region bounded by $p \cup q$. Suppose that $t_1, t_2, \ldots t_n$ are arcs in R, each meeting $p \cup q$ at just its end points, one in p and one in q. Suppose that every $t_i \cap t_j$ is at most one point, that intersections of arcs are transverse and that there are no triple points. The graph, with vertices all intersections of these arcs and edges comprising $p \cup q \cup \bigcup_i t_i$, separates R into a collection of v-gons; amongst these v-gons there is a 3-gon with an edge in p and a 3-gon with an edge in q.*

PROOF. Proceed by induction on the number n of arcs. The result is trivial if $n = 1$, so assume $n > 1$. Amongst the end points of the t_i that lie on p, let X be the nearest to A. If then X is an end of t_j, let B' be the other end of t_j on q. If possible, from $\{t_i : i \neq j\}$ select a t_k with $t_k \cap t_j = X'$, $t_k \cap q = A'$, such that t_k has no point of intersection with a t_i between A' and X'. Select such a t_k with X' as near as possible to B', see Figure 15.1. If there is no such t_k, select p instead, taking $A' = A$ and $X' = X$. Now let p' be an arc starting at A', proceeding along

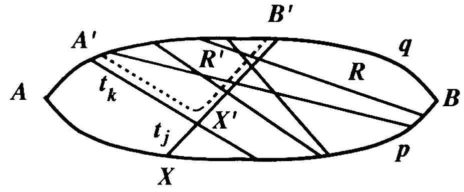


Figure 15.1

t_k (or p if there is no t_k) to X' and then along t_j to B' (it helps not to think of a corner at X'). Let q' be the sub-arc of q from A' to B' and let R' be the region bounded by $p' \cup q'$. If no t_i meets the interior of R', then R' is a 3-gon with an edge in q. Otherwise R' meets $\{t_i : i \neq j, i \neq k\}$ in fewer than n arcs and so, by induction, there is a 3-gon in R' with an edge in q'. The choice made for t_k ensures that A' is not a vertex of this 3-gon (which is important, as X' is *not* a vertex of a ν-gon of R'), and so it is one of the original 3-gons having an edge contained in q. Similarly, there is a 3-gon with an edge in p. □

In the course of the proof of the existence of the HOMFLY and Kauffman polynomials, the idea of an *ascending* link diagram will be used. The idea is as follows: A diagram D of an oriented link is *ordered* if an ordering is chosen for the link components and *based* if a base point is selected in D on each link component. If D is so ordered and based, the associated *ascending* diagram αD is formed from D by changing the crossings so that on a journey around all the components in the given order, always beginning at the base point of each component, *each crossing is first encountered as an under-pass*. That means that the link represented by αD can be thought of as lying in $\mathbb{R}^3$ above the diagram, with each component entirely below those following it in the given order, and with each component ascending as one moves around it away from its base point, but eventually dropping vertically back to that base point. Thus αD represents a trivial link. It is important to remember, given D, that αD depends on those two choices, component order and base points.

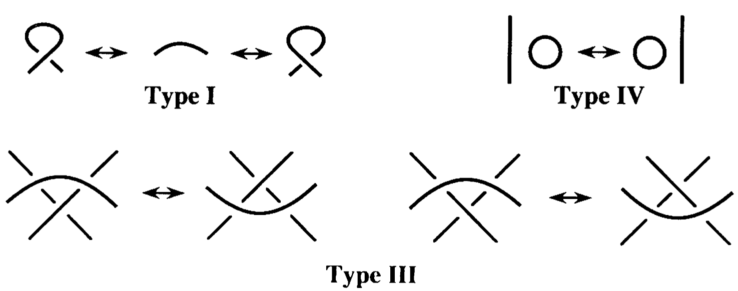


Figure 15.2

Also needed in the proof is the idea of Reidemeister moves that do not increase the number of crossings of a diagram above a certain bound. In principle it is clear what that means, but the moves will be taken to be the usual Type II move together with those of the Figure 15.2, which include two forms of the Type I and Type III moves and a Type IV move. The usual proofs that this is a redundant list involve increasing the number of crossings.

Theorem 15.2. *There is a unique function*

$$P : \{\text{Oriented links in } S^3\} \longrightarrow \mathbb{Z}[l^{\pm 1}, m^{\pm 1}]$$

such that P takes the value 1 on the unknot and, if L_+, L_- and L_0 are links that have diagrams D_+, D_- and D_0 that are the same except near a single point where they are as in Figure 15.3, then

$$l P(L_+) + l^{-1} P(L_-) + m P(L_0) = 0.$$

$P(L)$ is called the HOMFLY polynomial (see [31]) of the oriented link L.

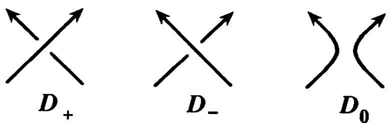


Figure 15.3

PROOF. In outline, the proof consists of defining P on link diagrams (using induction on the number of crossings), ensuring the validity of the skein relation

$$(\star) \qquad l P(D_+) + l^{-1} P(D_-) + m P(D_0) = 0$$

for diagrams related as in Figure 15.3, and verifying invariance under Reidemeister moves. Note first that the equation $(\star)$ determines uniquely any one of $P(D_+)$, $P(D_-)$ and $P(D_0)$ from knowledge of the other two. Note, too, that a solution to $(\star)$ is

$$(P(D_+), P(D_-), P(D_0)) = (x, x, \mu x),$$

where x is arbitrary and $\mu = -m^{-1}(l + l^{-1})$.

Let $\mathcal{D}_n$ be the set of all oriented link diagrams in the plane with at most n crossings (two diagrams are regarded as being identical if they differ by an orientation-preserving homeomorphism of the plane). Suppose inductively that $P : \mathcal{D}_{n-1} \to \mathbb{Z}[l^{\pm 1}, m^{\pm 1}]$ has been defined such that on $\mathcal{D}_{n-1}$

(i) the skein relation $(\star)$ holds for any three diagrams in $\mathcal{D}_{n-1}$ related in the usual way;

(ii) $P(D)$ is unchanged by Reidemeister moves on D that never involve more than $n - 1$ crossings;

(iii) if D is any ascending diagram of a link in $\mathcal{D}_{n-1}$ with $\#D$ components, then $P(D) = \mu^{\#D-1}$.

The induction starts with $\mathcal{D}_0$ in which any diagram is, trivially, ascending, and there is nothing to prove.

Now extend the definition of P over $\mathcal{D}_n$ in the following way: If D is an n-crossing diagram, select an ordering of its components, select a base point on each component and let αD be the associated ascending diagram. Define $P(\alpha D) = \mu^{\#D-1}$, where $\#D$ is the number of link components of D. The crossings of αD can be changed one at a time to achieve D. If D_+ and D_- are the diagrams before and after such a crossing change, and D_0 is the diagram with the crossing annulled, the value of P on the diagram after the change can be calculated, using ($\star$), from $P(D_0)$ and the value on the diagram before the change. The value of $P(D_0)$ is known by induction. The value for $P(D)$ is then *defined* to be the value thus calculated from $P(\alpha D)$ by changing crossings of αD to produce D. It is easy to see that $P(D)$ does not depend on the ordering of the sequence of crossing changes chosen to get from αD to D (consider transposing one crossing change and the next one in the sequence). However, the problem is to show that $P(D)$ does not depend on component order and choice of base points.

It is fairly easy to deal with base points. Suppose, keeping fixed the order of link components, the base point b of a certain link component of D is moved from just before a crossing to b', a point just after the crossing. Let βD be the ascending diagram using b' instead of b. If the other segment involved at the crossing is from a different component, then $\alpha D = \beta D$. Otherwise βD is constructed from αD by simply changing this crossing. However, the diagram D_0 obtained by annulling this crossing is also an ascending diagram with $\#D + 1$ link components and is, of course, in $\mathcal{D}_{n-1}$. Thus by the induction, $P(D_0) = \mu^{\#D}$ and, as $P(\alpha D) = \mu^{\#D-1}$, the skein formula gives $P(\beta D) = \mu^{\#D-1}$. This means that if one had defined $P(\beta D) = \mu^{\#D-1}$, next calculated that $P(\alpha D) = \mu^{\#D-1}$ and then calculated $P(D)$, one would have obtained the same value for $P(D)$ as before. Hence the definition of $P(D)$ is independent of choice of base points.

At this stage P is well defined on n-crossing diagrams with an ordering of their components. For such diagrams the identity ($\star$) is satisfied (assuming D_+ and D_- have the "same" orderings), for ($\star$) may be regarded as the first step in a calculation of $P(D_+)$ from $P(\alpha D)$. Reidemeister moves, that never involve more than n-crossings, on an ordered diagram D will now be considered. Note that if the diagram before a Reidemeister move has an ordering on its components, then this clearly induces an ordering on the components of the diagram after the move. A move is to be interpreted with respect to such associated orderings.

Suppose a crossing of D is to be removed by a Type I move on some component. It can be assumed (as position of base points is immaterial) that in selecting base points, the base point on the component in question is immediately before the crossing. Then this crossing is not changed in obtaining D from αD. Thus the calculation of P for D is exactly the same as the calculation for the diagram after the move.

With reference to a Reidemeister move of Type II, consider the two triples of diagrams shown in Figure 15.4(a). The two diagrams labelled D are the same, as are the two labelled D_0. Thus, by ($\star$), P takes the same value on the two labelled

D_+. The same is true for the two triples of Figure 15.4(b), using in addition invariance of P under Type I moves. Hence when considering removing two crossings of D by a Type II move, choose base points away from the area concerned, and note that the above remarks on Figure 15.4 imply that, without loss of generality, both crossings may be changed before even considering the move. Thus it may be assumed that the crossings are such that neither has to be changed in obtaining D from αD. As before, the calculations of P on the diagrams before and after the move are the same.

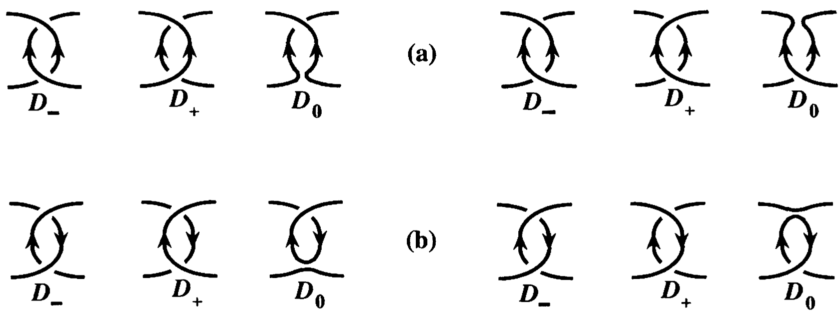


Figure 15.4

The same general idea works for Type III moves. Consider the diagrams of Figure 15.5 where it is supposed that the components are in some way oriented. The task of showing that $P(D_1) = P(D_1')$ is equivalent to the task of showing that $P(D_2) = P(D_2')$. This is because $D_3 = D_3'$, and D_4 and D_4' are related by a Type II move; $(\star)$ gives the usual relationship between $P(D_1)$, $P(D_2)$, and one of $P(D_3)$ and $P(D_4)$ (according to the orientation situation), and it also gives exactly the same relation between the $P(D_i')$. In this way the three crossings under consideration in the diagrams before and after a contemplated Type III move may be adjusted so that (choosing base points well out of the way) no crossing needs to be changed to achieve the ascending diagram. As before, the calculations of P before and after the move are the same. Finally, note that the fourth type of move, introduced (temporarily) above, clearly does not effect calculations of P.

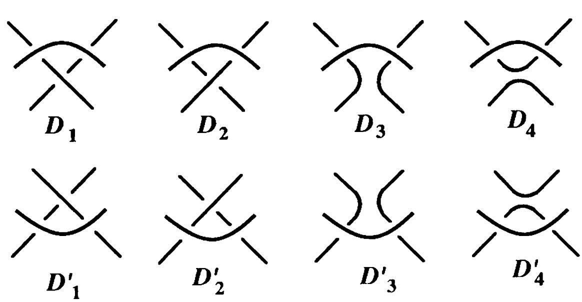


Figure 15.5

One thing remains to be proved. Suppose D is any n-crossing diagram with an ordering of its components and αD is the associated ascending diagram (with respect to any base points). Suppose that βD is an ascending diagram constructed from D with reference to a *different* ordering. Give the components of βD the *original* ordering, so that $P(\beta D)$ is defined by calculating it from $P(\alpha D)$. It is required to show that $P(\beta D) = \mu^{\#D-1}$. If this is so, then a calculation of a value of $P(D)$ could be started from knowing $P(\beta D) = \mu^{\#D-1}$, calculating $P(\alpha D) = \mu^{\#D-1}$ and then calculating the value of $P(D)$ as prescribed in the definition (with D having its original α-given order). That would mean that $P(D)$ would be well defined, independent of the ordering of its components. It would also complete the induction as βD is an arbitrary ascending n-crossing diagram.

To make this final check, consider the ascending diagram βD. Any component with no crossing that bounds a disc whose interior is disjoint from the diagram may be moved away from the rest of the diagram (into the unbounded complementary region) using the fourth type of move. Now, consider an innermost loop of the diagram (it may help to forget the over-crossing information for a while), a loop being a sub-arc of the diagram starting and stopping at the same crossing. *If* this loop contains no crossing (except at its ends) it can be removed using a Type I move (in this case, "innermost" and the remark on zero-crossing components imply there is no component totally within the area bounded by the loop). That move leaves the value of P unchanged. However, the new diagram has $n - 1$ crossings and is still ascending. Thus, by the induction, $P(\beta D) = \mu^{\#D-1}$. *Otherwise* other arcs of βD traverse the loop; these transversals are simple arcs, as the loop is innermost and each meets the loop at two points. One transversal and part of the loop bound a 2-gon, which is probably crossed by many transversals. Amongst such 2-gons, and similar 2-gons bounded by pairs of the transversals, choose an innermost one. Let the two arcs involved, denoted p and q, meet at points A and B and bound the region R. (Again, the innermost condition implies there is no component entirely within R.) Any of the remaining transversals that meets R, meets each of p and q, and within R, transversals meet each other in at most one point (as R is innermost). This is the situation described in Lemma 15.1, so considering the pattern of ν-gons in R formed by the complementary regions of βD, there is a 3-gon having an edge in p. Assuming all the base points are outside R, the fact that βD is ascending means that the 3-gon has cross-overs at its three vertices that are appropriate for a Type III Reidemeister move. Thus change the diagram by such a move, moving the part of p across the 3-gon. This changes R to a new region, and the procedure can be repeated; at each stage the diagram is still ascending. Eventually there are no 3-gons in the new region R, in which case that region can be removed completely by a Type II move. Thus βD can be changed by Reidemeister moves, which never involve more than n crossings, to an ascending diagram with $n - 2$ crossings. Thus $P(\beta D) = \mu^{\#D-1}$ by induction. This means that choosing the base points outside R was valid, as position of base points is irrelevant in ascending diagrams with this value of P.

This completes the proof of the induction hypothesis. Thus P is, finally, well defined on $\mathcal{D}_n$ for all n, the skein formula ($\star$) is always satisfied, and as any collection of Reidemeister moves remains within $\mathcal{D}_n$ for *some* n, $P(D)$ is unchanged by all Reidemeister moves. Thus a link invariant is produced by taking $P(L) = P(D)$, where D is any diagram for the link L.

There is only one such P as described in the theorem, for the properties of the statement of the theorem always allow $P(L)$ to be calculated from an ascending diagram. An ascending diagram with c components is an unlink and so is represented by a zero-crossing diagram of c components. It then follows (an easy exercise) from the statement of the theorem that the value of P on that diagram is μ^{c-1}. □

There is nothing sacrosanct about the notation used here for the HOMFLY polynomial. The skein formula simply expresses a linear relation between the values of P on three oriented diagrams related in the usual way. It is equally valid to regard P as having values in the Laurent polynomials in three (projective) variables x, y and z, with the skein relation being $xP(L_+) + yP(L_-) + zP(L_0) = 0$. A helpful custom has been established to the effect that any use of the HOMFLY polynomial is accompanied by a declaration of notational conventions.

In earlier chapters the skein formulae for the Jones polynomial and the Conway polynomial have already been considered, so the general procedures that might be applied to the HOMFLY polynomial are not unfamiliar. The details of the proof of the above theorem explain how $P(L)$ can be calculated by reference to an ascending diagram. Although the length of such a calculation depends exponentially on the number of crossings of a diagram, it is easy to calculate with diagrams of only a few crossings. It is also easy to make trivial errors in such calculations; several computer programs have been written to obviate this and to manage an inhuman number of crossings. Further exploration of the HOMFLY polynomial will be postponed to the following chapter. Instead, the preceding existence proof will first be adapted to give a proof of the existence of the Kauffman polynomial. That is the other two-variable polynomial invariant that generalises the Jones polynomial; it should not to be confused with the Kauffman bracket. Parts of the two existence proofs are the same, including the final tricky component re-ordering section. Thus emphasis will be placed on places where the proofs differ. One difference is that the Kauffman polynomial is really not defined on links but on *framed* links. The HOMFLY polynomial *can* be regarded as referring to framed links but that can seem, initially, to be an unnecessary sophistication. For the Kauffman polynomial it *is* necessary. In what follows, framings will be interpreted by means of diagrams, the framing on a component being the sum of the signs of the crossings at which that component crosses itself.

The work of the second existence proof consists in defining a two-variable Laurent polynomial invariant $\Lambda(D) \in \mathbb{Z}[a^{\pm 1}, z^{\pm 1}]$ for unoriented link diagrams D. This is the burden of the next theorem. If a diagram D happens to have an orientation, it should be forgotten when evaluating $\Lambda(D)$. Given this, the Kauffman polynomial $F(L)$ of an oriented link L has the following simple definition:

Definition 15.3. The Kauffman polynomial is the function

$$F : \{\text{Oriented links in } S^3\} \longrightarrow \mathbb{Z}[a^{\pm 1}, z^{\pm 1}]$$

defined by $F(L) = a^{-w(D)}\Lambda(D)$, where D is a diagram with writhe $w(D)$ of the oriented link L and Λ is the function of the next theorem.

In the course of the proof of the next theorem, it will be useful to use a *self-writhe* $\overline{w}(D)$ of a link diagram D, defined to be the sum of the signs of crossings at which all link components cross themselves (not other components). This self-writhe can be considered as the sum of the framings of all the components; note that it does *not* depend on a choice of orientation on the link. For an oriented link diagram D, the difference $w(D) - \overline{w}(D)$ is twice the sum of the linking numbers between all pairs of link components. (It might have been better, and equally valid, had the Kauffman polynomial been defined using $\overline{w}(D)$ instead of $w(D)$, for $a^{-\overline{w}(D)}\Lambda(D)$ is just an invariant of unoriented links.)

The proof of the next theorem will again use a definition involving induction on the number of crossings in a diagram and reference to ascending diagrams. However, a slight generalisation of an ascending diagram will be needed. This consists of the idea of a link diagram having an *untying function* (to be thought of as a "height") in the following sense:

Definition 15.4. Suppose D is a diagram for a link L with ordered components. An untying function for D is a real-valued function h on D, two-valued at the crossings, that corresponds to a continuous function $h : L \to \mathbb{R}$, with the following properties:

(i) If component c_i precedes component c_j in the ordering, then $h(x_i) < h(x_j)$ for any $x_i \in c_i$ and $x_j \in c_j$.

(ii) On each link component c_i, the function h is monotonically strictly increasing from some base point $b_i \in c_i$ to some top point $t_i \in c_i$, in both directions around c_i.

(iii) At a crossing the value of h on the over-pass exceeds that on the under-pass.

Note that any ascending (oriented) diagram has an untying function in which the top points of components always just precede the base points. Note, too, that if D has an untying function, it represents the unlink. This is because it represents a link L in which h is the height function of L above the plane of the diagram D (just "lift D up" to the height specified by h). Then L attains each height at most twice (by the monotonicity of h), so that the union of line segments joining points of L of equal height gives a collection of disjoint discs bounded by L.

Theorem 15.5. *There exists a function*

$$\Lambda : \{\text{Unoriented links diagrams in } S^2\} \longrightarrow \mathbb{Z}[a^{\pm 1}, z^{\pm 1}]$$

that is defined uniquely by the following:

(i) $\Lambda(U) = 1$, *where U is the zero-crossing diagram of the unknot;*

(ii) $\Lambda(D)$ *is unchanged by Reidemeister moves of Types II and III on the diagram* D;

(iii) $\Lambda($  $) = a\Lambda($  $)$;

(iv) *If* D_+, D_-, D_0 *and* D_∞ *are four diagrams exactly the same except near a point where they are as shown in Figure 15.6, then*

$$(\star\star) \qquad \Lambda(D_+) + \Lambda(D_-) = z(\Lambda(D_0) + \Lambda(D_\infty)).$$

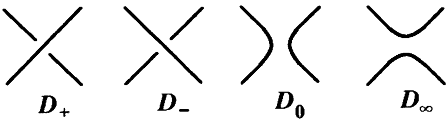


Figure 15.6

PROOF. Note that, when considering a crossing in an unoriented diagram, it has no claim to be termed D_+ rather than D_- in the above notation. However, this never matters, since D_+ and D_- feature symmetrically in the formula $(\star\star)$; the treatment of D_0 and D_∞ is likewise symmetric. Observe that the equation $(\star\star)$ determines uniquely any one of $\Lambda(D_+)$, $\Lambda(D_-)$, $\Lambda(D_0)$ and $\Lambda(D_\infty)$ from knowledge of the other three. Observe also that a solution to $(\star\star)$ is $(\Lambda(D_+), \Lambda(D_-), \Lambda(D_0), \Lambda(D_\infty)) = (ax, a^{-1}x, x, \delta x)$, where x is arbitrary and $\delta = (a + a^{-1})z^{-1} - 1$.

Now follow the pattern of the proof of Theorem 15.2. Let $\mathcal{D}_n$ be the set of all *unoriented* link diagrams in the plane with at most n crossings. Suppose inductively that $\Lambda : \mathcal{D}_{n-1} \to \mathbb{Z}[a^{\pm 1}, z^{\pm 1}]$ has been defined such that on $\mathcal{D}_{n-1}$

(a) the skein relation $(\star\star)$ holds for any four diagrams in $\mathcal{D}_{n-1}$ related as in Figure 15.6;

(b) $\Lambda($  $) = a\Lambda($  $)$ and $\Lambda($  $) = a^{-1}\Lambda($  $)$;

(c) $\Lambda(D)$ is unchanged by Reidemeister moves of Types II, III and IV on D that never involve more than $n - 1$ crossings (see Figure 15.2);

(d) if D is any diagram of a link in $\mathcal{D}_{n-1}$ with $\#D$ link components that has an untying function, then $\Lambda(D) = a^{\overline{w}(D)}\delta^{\#D-1}$.

The induction starts with $\mathcal{D}_0$ in which any diagram has, trivially, an untying function.

Now extend the definition of Λ over $\mathcal{D}_n$ in the following way: If D is an n-crossing diagram, select an orientation on each component, an ordering of the components and a base point on each component. Let αD be the associated ascending diagram (it is just to define this that the orientation is needed). Define $\Lambda(\alpha D) = a^{\overline{w}(\alpha D)}\delta^{\#D-1}$, where $\#D$ is the number of link components of D (and of αD). The value for $\Lambda(D)$ is *defined* to be the value calculated from $\Lambda(\alpha D)$ by changing one by one the crossings of αD to produce D, using $(\star\star)$ and inductive knowledge of $\Lambda(D_0)$ and $\Lambda(D_\infty)$ at each crossing change. It is easy to see that $\Lambda(D)$ does not depend on the ordering of the sequence of crossing changes chosen to change αD to D. The problem is to show that $\Lambda(D)$ does not depend on component order, component orientation and choice of base points.

Suppose, keeping fixed the orientations and the order of link components, the base point b of a certain link component of D is moved from just before a crossing to b', a point just after the crossing. Let βD be the ascending diagram using b' instead of b. If the other segment involved at the crossing is from a different component, then $\alpha D = \beta D$. Otherwise βD is constructed from αD by simply changing this crossing. However, the diagram D_0, obtained from αD by annulling that crossing in the way consistent with orientation, is also an ascending diagram with $\#D + 1$ link components, and the diagram D_∞, obtained by annulling in the other way, is easily seen to have an untying function. Both D_0 and D_∞ are in $\mathcal{D}_{n-1}$. Hence $\Lambda(D_0) = a^{\overline{w}(D_0)}\delta^{\#D}$ and $\Lambda(D_\infty) = a^{\overline{w}(D_\infty)}\delta^{\#D-1}$. However, as in all four diagrams there are zero linking numbers between components, $\overline{w}(D_\infty) = \overline{w}(D_0)$, and $\overline{w}(\alpha D)$ and $\overline{w}(\beta D)$ are $\overline{w}(D_0) + 1$ and $\overline{w}(D_0) - 1$ (which is which depending on the sign of the crossing). Hence the skein relation ($\star\star$) shows that $\Lambda(\beta D) = a^{\overline{w}(\beta D)}\delta^{\#D-1}$, as this valued substituted in ($\star\star$) gives

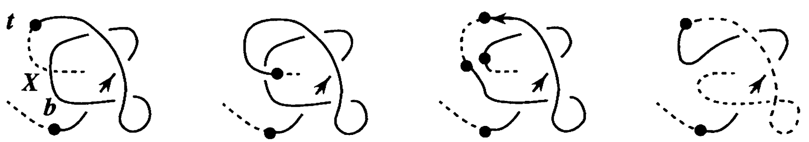
$$a^{\overline{w}(D_0)}\{a\delta^{\#D-1} + a^{-1}\delta^{\#D-1}\} = za^{\overline{w}(D_0)}\{\delta^{\#D} + \delta^{\#D-1}\},$$

and that accords with the definition of δ. This value for $\Lambda(\beta D)$ is that calculated using b as the base point; it is, of course, equal to the value it would have, by definition, if b' were the base point. Thus $\Lambda(D)$ does not depend on choice of base points.

At this stage Λ is well defined on n-crossing diagrams with an ordering of their components and an orientation on each component. For such diagrams the ($\star\star$) identity of statement (a) of the induction hypothesis is satisfied exactly as in the proof of Theorem 15.2. To see that the formulae of statement (b) are satisfied, let D be the diagram on the left-hand side of such a formula and D' that on the right. Placing the base point just before the crossing shown ensures that the crossing is unchanged in the ascending diagram αD. But then $\overline{w}(\alpha D) = w(\alpha D') \pm 1$, the choice of sign depending on the sign of the crossing. Thus by definition, $\Lambda(\alpha D) = a^{\pm 1}\Lambda(\alpha D')$, and the factor $a^{\pm 1}$ persists throughout the calculations to show that $\Lambda(D) = a^{\pm 1}\Lambda(D')$.

Reidemeister moves other than of Type I, and never involving more than n-crossings, on ordered oriented diagrams must now be considered. The invariance of $\Lambda(D)$ under a Reidemeister move of Type II is shown in exactly the same way as in Theorem 15.2. The diagrams of Figure 15.4 should be considered without arrows, and it should be noted that the values of Λ on the two diagrams labelled D_0 in Figure 15.4(b) are the same by means of two applications of the formula (b) that has just been proved. Similarly, invariance under a Reidemeister move of Type III follows as before; it is simply required to use all of the D_i and D_i' shown in Figure 15.5. Again invariance under a Type IV move follows trivially.

It is necessary, for (d), to check that $\Lambda(D) = a^{\overline{w}(D)}\delta^{\#D-1}$ for an n-crossing oriented ordered diagram D with an untying function h. This is true if the top points on all components immediately precede the base points, as then the diagram is ascending with respect to its ordering and orientation. Thus proceed by a sub-induction on the total number of self-crossings of all components from top points to base points in the directions of the orientations. On a component c, let X be the first self-crossing on c after top point t (before the base point b is reached). If, on travelling from t, the crossing is an over-pass, h can be changed to be increasing from t to just beyond X and still be an untying function. Then $\Lambda(D) = a^{\overline{w}(D)}\delta^{\#D-1}$ by the sub-induction. Otherwise X is encountered as an under-pass. It must be an under-passing of part of c from b to t; this follows from the monotonicity properties of h. The situation is illustrated in Figure 15.7(a), where h is to be thought of as decreasing along broken lines and increasing along unbroken lines. Calculate $\Lambda(D)$ using ($\star\star$) applied to the crossing X. Changing the crossing gives a diagram D', and this diagram has an untying function with the top point moved nearer to base point b. Thus, by the sub-induction, $\Lambda(D') = a^{\overline{w}(D')}\delta^{\#D-1}$. The diagrams D_0 and D_∞ have $n-1$ crossings and are as in Figure 15.7(c) and Figure 15.7(d); these diagrams have untying functions as indicated by the broken and unbroken lines. Thus $\Lambda(D_0)$ and $\Lambda(D_\infty)$ are known by induction on n. Of course, $\overline{w}(D_\infty) = \overline{w}(D_0)$, and $\overline{w}(D)$ and $\overline{w}(D')$ are $\overline{w}(D_0)+1$ and $\overline{w}(D_0)-1$. Further, D_0 has one more component than the other diagrams. It then follows at once from ($\star\star$) that $\Lambda(D) = a^{\overline{w}(D)}\delta^{\#D-1}$, and the induction argument is complete.



(a) (b) (c) (d)

Figure 15.7

Consider an ordered oriented based diagram D and the associated ascending diagram αD. Suppose the component ordering is kept fixed but that the orientation on some component c is reversed. Let βD be the resulting ascending diagram. Then βD has an untying function (and untying functions ignore orientation). Thus, with respect to the original orientation on βD, $\Lambda(\beta D) = a^{\overline{w}(\beta D)}\delta^{\#D-1}$. This means that the definition of $\Lambda(D)$ does not depend on choice of the orientations used in defining ascending diagrams.

At this stage any possible ambiguity in $\Lambda(D)$ depends only on the chosen ordering of components. However, that this choice, too, is irrelevant follows exactly as in Theorem 15.2. So with the induction on n complete, the theorem is proved, for uniqueness follows easily as before. □

A variant in the signs for the Kauffman polynomial is sometimes useful. (The resulting polynomial is sometimes called the "Dubrovnic" polynomial [60].) This is based on a function

$$\Lambda^* : \{\text{Unoriented links diagrams in } S^2\} \to \mathbb{Z}[a^{\pm 1}, z^{\pm 1}]$$

that is defined in exactly the same way as is Λ in Theorem 15.5 (with α in place of a and ω in place of z) except that (iv) is replaced by

$$\Lambda^\star(D_+) - \Lambda^\star(D_-) = \omega(\Lambda^\star(D_0) - \Lambda^\star(D_\infty)).$$

If an oriented link L is represented by diagram D, define $F^\star(L) = \alpha^{-w(D)}\Lambda^\star(D)$. It is, however, fairly easy to verify that if L has $\#L$ components, then

$$F^\star(L) = (-1)^{\#L-1}[F(L)]_{(a,z)=(i\alpha,-i\omega)},$$

where $i^2 = -1$. Thus this variant contains no additional information.

Exercises

1. Evaluate the HOMFLY and Kauffman polynomials for each of the three knots with crossing number 6.

2. Suppose that the HOMFLY polynomial exists and satisfies the criteria of the statement of Theorem 15.2. Show that if L is the trivial link with $\#L$ components then $P(L) = \mu^{\#L-1}$, where $\mu = -m^{-1}(l + l^{-1})$.

3. Suppose that an oriented link L' is obtained from oriented link L by reversing the direction of one of the components of L. Show, by considering specific examples, that there is no simple multiplicative formula relating $P(L)$ and $P(L')$ of the type that exists for the Jones and Kauffman polynomials.

4. Show that there exists a version $P^\star(L)$ of the HOMFLY polynomial invariant of oriented links that is a function of indeterminates x, y, and z, with the property that $P^\star(\text{unknot}) = 1$ and

$$xP^\star(L_+) + yP^\star(L_-) + zP^\star(L_0) = 0.$$

Here, as usual, L_+, L_- and L_0 are three oriented links identical except within a ball where, respectively, they have a positive crossing, a negative crossing and no crossing. Show that $P^\star(L)$ is homogeneous in x, y, and z and determine the relationship between $P^\star(L)$ and $P(L)$.

5. Show that the existence of the HOMFLY polynomial implies that $X(L)$, an invariant of an oriented link L, that is a function of l, m, a and z, can be defined by

$$X(\text{unknot}) = a \text{ and } lX(L_+) + l^{-1}X(L_-) + mX(L_0) + z = 0.$$

Here L_+, L_- and L_0 are related in the usual way.

6. Show that an invariant $Y(L)$ of an oriented link L, with $Y(L)$ a function of an indeterminate x, can be defined by

$$Y(\text{unknot}) = 1 \text{ and } Y(L_-)Y(L_0) = x\Big(Y(L_+) - Y(L_-) - Y(L_0)\Big).$$

7. Show that if L is a split link, then under the substitution $(a, z) = (q, q^{-1} + q)$, the polynomial $F(L)$ is zero.

8. For an oriented diagram D of an oriented link L, define $\widehat{P}(D)$ by $\widehat{P}(D) = l^{w(D)}P(L)$. Show that

(i) if D is changed by regular isotopy, then $\widehat{P}(D)$ does not change;
(ii) if D_+, D_- and D_0 are oriented diagrams related in the usual way, then

$$\widehat{P}(D_+) + \widehat{P}(D_-) + m\widehat{P}(D_0) = 0;$$

(iii) if D' is D with a positive kink removed, then $\widehat{P}(D) = l\widehat{P}(D')$.

9. What is the value of the "Dubrovnik" polynomial $F^*(L)$ of the oriented link L when $\alpha = 1$?

10. For a knot K, let $Q(K)$ be the polynomial in z obtained by substituting $a = 1$ in $F(L)$. If K has an alternating diagram with n crossings, show that the degree of $Q(L)$ is at most $n - 1$.

16

Exploring the HOMFLY and Kauffman Polynomials

Elementary properties of the Jones polynomial have already been discussed in Chapter 3. Versions of some of those results which hold equally well for the HOM FLY and Kauffman polynomials are given below. Where proofs are essentially the same as those relating to the Jones polynomial, they are left as an exercise.

Proposition 16.1. *If L is an oriented link and $\overline{L}$ is its reflection, then*
(i) *changing the signs of both variables leaves $P(L)$ and $F(L)$ unchanged;*
(ii) $\overline{P(L)} = P(\overline{L})$ *where* $\overline{l} = l^{-1}$ *and* $\overline{m} = m$;
(iii) $\overline{F(L)} = F(\overline{L})$ *where* $\overline{a} = a^{-1}$ *and* $\overline{z} = z$.

Proposition 16.2. *If L_1 and L_2 are oriented links, then*
(i) $P(L_1 + L_2) = P(L_1)P(L_2)$;
(ii) $F(L_1 + L_2) = F(L_1)F(L_2)$;
(iii) $P(L_1 \sqcup L_2) = -(l + l^{-1})m^{-1}P(L_1)P(L_2)$;
(iv) $F(L_1 \sqcup L_2) = ((a + a^{-1})z^{-1} - 1)F(L_1)F(L_2)$.

Note that in the above "$\sqcup$" denotes the separated (or split) union of links. Note too that the sum of oriented links is not well defined; it depends upon which components are used to produce the sum. The above results are true whatever components are selected, and so by varying that selection, different links are obtained with the same polynomials.

Proposition 16.3. *Both $P(L)$ and $F(L)$ are unchanged by mutation of L.*

Mutation was discussed in Chapter 3, with a famous example shown in Figure 3.3. That example thus produces different *knots* with the same polynomials. In [54], T. Kanenobu gave infinitely many distinct knots all with the same HOMFLY polynomial (and hence also the same Jones polynomial).

Proposition 16.4. *If the oriented link L' is obtained from the oriented link L by reversing the orientation of one component K, then*

$$F(L') = a^{4\,\mathrm{lk}(K,\,L-K)}F(L).$$

Changing the orientation of <u>all</u> *components of L leaves both $P(L)$ and $F(L)$ unchanged.*

The HOMFLY polynomial and the Kauffman polynomial are independent invariants in the sense that they distinguish different pairs of knots. Thus neither polynomial can be seen to be trivially contained in the other by means of some subtle change of variables. Examples are shown in Figure 16.1. The knots 8_8 and 10_{129} have the same HOMFLY polynomial but distinct Kauffman polynomials (even when taking the variable $a = 1$). Knots 11_{255} and 11_{257} have the same Kauffman polynomial but distinct HOMFLY polynomials (and even have distinct Alexander polynomials).

8_8 10_{129} 11_{255} 11_{257}

Figure 16.1

Skein formulae have already been given for the Jones polynomial, in Proposition 3.7, and for the Conway-normalised Alexander polynomial in Theorem 8.6. Those results just mean that certain substitutions of variables in the HOMFLY polynomial give the Jones polynomial on the one hand or the Alexander polynomial on the other. The precise results, in the notation used here, are as follows:

Proposition 16.5. *For an oriented link L, the Conway-normalised Alexander polynomial $\Delta_L(t)$ and the Jones polynomial $V(L)$ are related to the HOMFLY polynomial $P(L)$ by*

$$\Delta_L(t) = P(L)_{(i,\ i(t^{1/2}-t^{-1/2}))} \qquad \text{and} \qquad V(L) = P(L)_{(it^{-1},\ i(t^{-1/2}-t^{1/2}))},$$

where $i^2 = -1$.

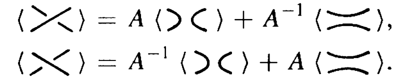
The Alexander polynomial is not contained within the Kauffman polynomial as the example above shows. However, the Jones polynomial *is* hidden in the Kauffman polynomial in two ways.

Proposition 16.6. *For an oriented link L,*

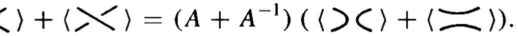
$$V(L) = F(L) \text{ when } (a, z) = (-t^{-3/4},\ (t^{-1/4} + t^{1/4})),$$

$$(V(L))^2 = (-1)^{\#L-1} F(L) \text{ when } t = -q^{-2},\ (a, z) = (q^3, q^{-1} + q).$$

PROOF. Underlying the Jones polynomial is the Kauffman bracket. Reverting to the notation for that, given in Definition 3.1,

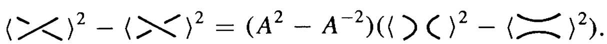


Adding these equations gives



However,  $= -A^3$ , and the Kauffman bracket is invariant under regular isotopy (Reidemeister Type II and III moves). However, these are the defining rules for the polynomial $\Lambda(D)$ of Theorem 15.5 with the variables changed to $(a, z) = (-A^3, (A + A^{-1}))$. The substitution $t = A^{-4}$ gives the first result.

Subtracting the square of one of the above two equations from the square of the other gives



Of course  $= A^6$ , and so the square of the Kauffman bracket is an instance of the $\Lambda^\star(D)$ polynomial, defined at the very end of Chapter 15, that satisfies

$$\Lambda^\star(D_+) - \Lambda^\star(D_-) = \omega(\Lambda^\star(D_0) - \Lambda^\star(D_\infty)).$$

Now, translating the notation by $\omega = (A^2 - A^{-2})$, $\alpha = A^6$, $t = -q^{-2} = A^{-4}$ and (from Chapter 15) $(a, z) = (i\alpha, -i\omega)$ the second result follows. □

Many of the best simple applications of the skein theoretic polynomials refer to the Jones and Alexander polynomials and have already been considered (particularly in Chapter 5). However, the following result is a significant direct application of the HOMFLY polynomial; it is one of the best applications of that polynomial to geometric questions. Versions of it first appeared in [27] and [97]. The result gives, particularly in its corollary, some information about the complexity that must exist in a diagram of a specific link. As convenient notation, let $E_l(P(L))$ and $e_l(P(L))$ be the maximum and minimum exponents of l that appear in the HOMFLY polynomial $P(L)$ of an oriented link L.

Theorem 16.7. *Suppose that an oriented link L is represented by a diagram D with writhe $w(D)$, having $s(D)$ Seifert circuits and $n(D)$ crossings. Then the degrees of m that occur in $P(L)$ are bounded above by $n(D) - s(D) + 1$, and*

$$-w(D) - s(D) + 1 \le e_l(P(L)) \le E_l(P(L)) \le -w(D) + s(D) - 1.$$

PROOF. The first inequality asserts that $E_m(P(L)) \le n(D) - s(D) + 1$, where $E_m(P(L))$ is the maximum exponent of m occurring in $P(L)$. This will be proved by induction on $n(D)$. if $n(D) = 0$, then $s(D) = \#D$ and $P(D) = \mu^{\#D-1}$ where $\mu = -m^{-1}(l + l^{-1})$, and the result follows. The skein relation for $P(D)$ is $lP(D_+) + l^{-1}P(D_-) + mP(D_0) = 0$, where D_+, D_- and D_0, being diagrams related in the usual way, have the same number of Seifert circuits. The inequality is by induction true on D_0, so using induction again on the number of crossings that need to be changed to achieve an ascending diagram, it is just necessary to prove the result for ascending diagrams. That is, it is required to prove for an ascending diagram D that $-(\#D - 1) \le n(D) - s(D) + 1$, or that $s(D) \le n(D) + \#D$. However, this inequality is true for any link diagram, as can be seen in the following way, again by induction on $n(D)$. It is clear when $n(D) = 0$. Let D' be the result of annulling a crossing of D. Then the inequality is true for D', and clearly $s(D) = s(D')$, $\#D = \#D' \pm 1$ and $n(D) = n(D') + 1$. That gives the inequality for D.

To consider the second inequality, define for an oriented link diagram D the Laurent polynomial $X(D) = l^{w(D)} P(D)$. It is required to show that $E_l(X(D)) \le s(D) - 1$. Proceed again by induction on n; the inequality is clearly true when $n = 0$. The skein relation for $X(D)$ is $X(D_+) + X(D_-) + mX(D_0) = 0$. Again, D_+, D_- and D_0 all have the same number of Seifert circuits and, as the required inequality is true by induction for D_0, it is sufficient to prove it for ascending diagrams. Suppose D is ascending so that $P(D) = \mu^{\#D-1}$. It is required to show that $w(D) + \#D \le s(D)$. Suppose that in D, some crossing of a component with itself is annulled to give another ascending diagram D' (the self-crossing following a base point will do). By induction on the number of crossings, $w(D') + \#D' \le s(D') = s(D)$. However, $\#D' = \#D + 1$ and $w(D') = w(D) \pm 1$, and so the inequality is true for D. If there is no crossing at which a component crosses itself, let D'' be obtained from D by annulling a *negative* crossing where one component crosses another. This can be done, as all linking numbers are zero. Choose the two components as close as possible, in the ordering of the components of D as an ascending diagram, and then D'' is also ascending. By induction on the number of crossings, $w(D'') + \#D'' \le s(D'') = s(D)$. Now $w(D'') = w(D) + 1$ and $\#D'' = \#D - 1$, and so the inequality holds for D.

The inequality $-w(D) - s(D) + 1 \le e_l(P(L))$ can be proved similarly or deduced from the above by reflection of the diagram. □

Corollary 16.8. *The l-breadth of $P(L)$ satisfies* $E_l(P(L)) - e_l(P(L)) \le 2(s(D) - 1)$.

The significance of this corollary is that for an oriented link L it gives a lower bound on the number of Seifert circuits in any diagram that might represent L. The minimum number of such Seifert circuits is known to be equal to another invariant, the "braid index" of L, which is defined to be the minimal n for which L can be described as the closure of an n-string braid (see [136]), so the corollary gives a lower bound for the braid index.

Applications of the Kauffman polynomial have been explored by Thistlethwaite in [119] and [120] and by Kidwell [64]. Some, though not all, of those results now follow from the technique of Chapter 5. One result of [64] is that if $Q(L)$ is the polynomial in z obtained by the substitution $a = 1$ in $F(L)$, and L has a diagram with $n(D)$ crossings, then

$$\text{degree } Q(L) \le n(D) - b(D),$$

where $b(D)$ is the maximum number of consecutive over-passes that occur anywhere in the diagram. This $b(D)$ is called the bridge length of the diagram D; if D is alternating then $b(D) = 1$.

It may be helpful to have a rough idea of the appearance of these polynomial invariants for knots. There follow two tables giving the values of the HOMFLY polynomial for knots up to eight crossings (as depicted in Chapter 1) and the Kauffman polynomial for knots up to 8_7. The HOMFLY polynomial of a *knot*, in the notation of the last chapter, is of the form $\sum_{i\geq 0} p_i(l^2)m^i$, where $p_i(l^2)$ is a Laurent polynomial in l^2, $p_i(l^2)$ is zero if i is odd and if i is sufficiently large. These simple facts are exploited in the table, which gives (in notation due to Thistlethwaite), for each knot listed, the coefficients in the polynomial. The numbers in the ith brackets give the coefficients in $p_{2(i-1)}(l^2)$, the bold face number being the coefficient of l^0. Thus, for example, the knot 7_7 has polynomial $(l^{-4} + 2l^{-2} + 2) + (-2l^{-2} - 2 - l^2)m^2 + m^4$, and this is abbreviated in the table to (1 2 **2**) (−2 **−2** − 1) (**1**).

The Kauffman polynomial of a *knot*, in the notation of the last chapter, is of the form $\sum_{i\geq 0} q_i(a)z^i$, where $q_i(a)$ is a Laurent polynomial in a. However, $q_i(a)$ contains only odd powers of a if i is odd and only even powers if i is even. If i is sufficiently large, q_i is zero. The table given here for Kauffman polynomials uses these elementary facts. Again in notation due to Thistlethwaite, the conventions are as follows: For a knot listed, the numbers in the i^{th} bracket give the coefficients in $q_i(a)$. If i is even, the coefficients listed are of the even powers of a (for the others are zero), the bold face number being the coefficient of a^0. If i is odd, the coefficients listed are of the odd powers of a, the star denoting the divide between negative and positive powers. Many of the listings occupy two lines. Thus, for example, the knot 6_1 has polynomial

$$(-a^{-2} + a^2 + a^4) + (2a + 2a^3)z + (a^{-2} - 4a^2 - 3a^4)z^2$$
$$+ (a^{-1} - 2a - 3a^3)z^3 + (1 + 2a^2 + a^4)z^4 + (a + a^3)z^5,$$

and this is encoded as

(-1 **0** 1 1) (⋆ 2 2) (1 **0** -4 -3) (1 ⋆ -2 -3) (**1** 2 1) (⋆ 1 1).

A glance at the tables of the HOMFLY and Kauffman polynomials (Tables 16.1 and 16.2) reveals that for any knot the first entry is the same in the two tables. That is an instance of the following result, the proof of which is left as an easy exercise.

Proposition 16.9. *Suppose L is an oriented link with #L components. Then (1 − #L) is the lowest power both of m in P(L) and of z in F(L), and*

$$\left[z^{\#L-1}F(L)\right]_{(a,z)=(l,0)} = \left[(-m)^{\#L-1}P(L)\right]_{m=0}.$$

Perhaps the elegance of this result is a quirk of notation, but it serves to focus attention on the polynomial $p_0(l^2)$. That invariant has been used by P. Traczyk [122] to provide a necessary condition that a knot should have a certain type of symmetry. The symmetry envisaged is that the knot might be (set wise) invariant

TABLE 16.1. HOMFLY Polynomial Table

3_1	(**0** -2 -1)	(**0** 1)		
4_1	(-1 **-1** -1)	(**1**)		
5_1	(**0** 0 3 2)	(**0** 0 -4 -1)	(**0** 0 1)	
5_2	(**0** -1 1 1)	(**0** 1 -1)		
6_1	(-1 **0** 1 1)	(**1** -1)		
6_2	(**2** 2 1)	(**-1** -3 -1)	(**0** 1)	
6_3	(1 **3** 1)	(-1 **-3** -1)	(**1**)	
7_1	(**0** 0 0 -4 -3)	(**0** 0 0 10 4)	(**0** 0 0 -6 -1)	(**0** 0 0 1)
7_2	(**0** -1 0 -1 -1)	(**0** 1 -1 1)		
7_3	(-2 -2 1 0 **0**)	(1 3 -3 0 **0**)	(-1 1 0 **0**)	
7_4	(-1 0 2 0 **0**)	(1 -2 1 **0**)		
7_5	(**0** 0 2 0 -1)	(**0** 0 -3 2 1)	(**0** 0 1 -1)	
7_6	(**1** 1 2 1)	(**-1** -2 -2)	(**0** 1)	
7_7	(1 2 **2**)	(-2 **-2** -1)	(**1**)	
8_1	(-1 **0** 0 -1 -1)	(**1** -1 1)		
8_2	(**0** -3 -3 -1)	(**0** 4 7 3)	(**0** –1 -5 -1)	(**0** 01)
8_3	(1 0 **-1** 0 1)	(-1 **2** -1)		
8_4	(-2 **-2** 0 1)	(1 **3** -2 -1)	(**-1** 1)	
8_5	(-2 -5 -4 **0**)	(3 8 4 **0**)	(-1 -5 -1 **0**)	(1 0 **0**)
8_6	(**2** 1 -1 -1)	(**-1** -2 2 1)	(**0** 1 -1)	
8_7	(-2 -4 **-1**)	(3 8 **3**)	(-1 -5 **-1**)	(1 **0**)
8_8	(-1 -1 **2** 1)	(1 2 **-2** -1)	(-1 **1**)	
8_9	(-2 **-3** -2)	(3 **8** 3)	(-1 **-5** -1)	(**1**)
8_{10}	(-3 -6 **-2**)	(3 9 **3**)	(-1 -5 **-1**)	(1 **0**)
8_{11}	(**1** -1 -2 -1)	(**-1** -1 2 1)	(**0** 1 -1)	
8_{12}	(1 1 **1** 1 1)	(-2 **-1** -2)	(**1**)	
8_{13}	(**0** -2 -1)	(-1 **-1** 2 1)	(**1** -1)	
8_{14}	(**1**)	(**-1** -1 1 1)	(**0** 1 -1)	
8_{15}	(**0** 0 1 -3 -4 -1)	(**0** 0 -2 5 3)	(**0** 0 1 -2)	
8_{16}	(**0** -2 -1)	(**2** 5 2)	(**-1** -4 -1)	(**0** 1)
8_{17}	(-1 **-1** -1)	(2 **5** 2)	(-1 **-4** -1)	(**1**)
8_{18}	(1 **3** 1)	(1 **1** 1)	(-1 **-3** -1)	(**1**)
8_{19}	(-1 -5 -5 0 0 **0**)	(5 10 0 0 0 **0**)	(-1 -6 0 0 **0**)	(1 0 0 **0**)
8_{20}	(**-1** -4 -2)	(**1** 4 1)	(**0** -1)	
8_{21}	(**0** -3 -3 -1)	(**0** 2 3 1)	(**0** 0 -1)	

TABLE 16.2. Kauffman Polynomial Table

3_1	(**0** -2 -1)	(★ 0 1 1)	(**0** 1 1)	
4_1	(-1 **-1** -1)	(-1 ★ -1)	(1 **2** 1)	(1 ★ 1)
5_1	(**0** 0 3 2)	(★ 0 0 -2 -1 1)	(**0** 0 -4 -3 1)	(★ 0 0 1 1)
	(**0** 0 1 1)			
5_2	(**0** -1 1 1)	(★ 0 0 -2 -2)	(**0** 1 -1 -2)	(★ 0 1 2 1)
	(**0** 0 1 1)			
6_1	(-1 **0** 1 1)	(★ 2 2)	(1 **0** -4 -3)	(1★ -2 -3)
	(**1** 2 1)	(★ 1 1)		
6_2	(**2** 2 1)	(★ 0 -1 -1)	(**-3** -6 -2 1)	(★ -2 0 2)
	(**1** 3 2)	(★ 1 1)		
6_3	(1 **3** 1)	(-1 -2 ★ -2 -1)	(-3 **-6** -3)	(1 1 ★ 1 1)
	(2 **4** 2)	(1 ★ 1)		
7_1	(**0** 0 0 -4 -3)	(★ 0 0 0 3 1 -1 1)	(**0** 0 0 10 7 -2 1)	(★ 0 0 0 -4 -3 1)
	(**0** 0 0 -6 -5 1)	(★ 0 0 0 1 1)	(**0** 0 0 1 1)	
7_2	(**0** -1 0 -1 -1)	(★ 0 0 0 3 3)	(**0** 1 0 3 4)	(★ 0 1 -1 -6-4)
	(**0** 0 1 -3 -4)	(★ 0 0 1 2 1)	(**0** 0 0 1 1)	
7_3	(-2 -2 1 0 **0**)	(-2 1 3 0 0 0 ★)	(-1 6 4 -3 0 **0**)	(1 -1 -4 -2 0 0 ★)
	(1 -3 -3 1 0 **0**)	(1 2 1 0 0 ★)	(1 1 0 0 **0**)	
7_4	(-1 0 2 0 **0**)	(4 4 0 0 0 ★)	(2 -3 -4 1 **0**)	(-4 -8 -2 2 0 ★)
	(-3 0 3 0 **0**)	(1 3 2 0 0 ★)	(1 1 0 0 **0**)	
7_5	(**0** 0 2 0 -1)	(★ 0 0 -1 1 1 -1)	(**0** 0 -3 0 1 -2)	(★ 0 0 -1 -4 -2 1)
	(**0** 0 1 -1 0 2)	(★ 0 0 1 3 2)	(**0** 0 0 1 1)	
7_6	(**1** 1 2 1)	(★ 1 2 0 -1)	(**-2** -4 -4 -2)	(★ -4 -6 -1 1)
	(**1** 1 2 2)	(★ 2 4 2)	(**0** 1 1)	
7_7	(1 2 **2**)	(2 3 ★ 1)	(-2 -6 **-7** -3)	(-4 -8 ★ -3 1)
	(1 2 **4** 3)	(2 5 ★ 3)	(1 **1**)	
8_1	(-1 **0** 0 -1 -1)	(★ 0 -3 -3)	(1 **0** 0 7 6)	(1 ★ -1 5 7)
	(**1** -2 -8 -5)	(★ 1 -4 -5)	(**0** 1 2 1)	(★ 0 1 1)
8_2	(**0** -3 -3 -1)	(★ 0 1 1 -1 -1)	(**0** 7 12 3 -1 1)	(★ 0 3 -1 -2 2)
	(**0** -5 -12 -5 2)	(★ 0 -4 -2 2)	(**0** 1 3 2)	(★ 0 1 1)
8_3	(1 0 **-1** 0 1)	(-4 ★ -4)	(-3 1 **8** 1 -3)	(-2 8 ★ 8 -2)
	(1 -2 **-6** -2 1)	(1 -4 ★ -4 1)	(1 **2** 1)	(1 ★ 1)
8_4	(-2 **-2** 0 1)	(-1 ★ 1 2)	(7 **10** -1 -3 1)	(4 ★ -3 -5 2)
	(-5 **-11** -3 3)	(-4 ★ -1 3)	(1 **3** 2)	(1 ★ 1)
8_5	(-2 -5 -4 **0**)	(4 7 3 0 ★)	(1 -2 4 15 8 **0**)	(2 -8 -10 0 0 ★)
	(3 -7 -15 -5 **0**)	(4 1 -3 0 ★)	(3 4 1 **0**)	(1 1 0 ★)
8_6	(**2** 1 -1 -1)	(★ -1 -3 -1 1)	(**-3** -2 6 3 -2)	(★ -1 5 2 -4)
	(**1** 0 -6 -4 1)	(★ 1 -2 -1 2)	(**0** 1 3 2)	(★ 0 1 1)
8_7	(-2 -4 **-1**)	(-1 0 2 2 ★ 1)	(-2 4 12 **6**)	(1 -1 -2 -3 ★ -3)
	(2 -3 -12 **-7**)	(2 0 -1 ★ 1)	(2 4 **2**)	(1 1 ★)

under a $2\pi/p$ rotation about some axis, for some prime p; the condition is in terms of the coefficients modulo p of $p_0(l^2)$.

Table 16.3 shows the values taken by the HOMFLY, Jones and Kauffman polynomials for an arbitrary link L when various specific values are substituted for the variables of the polynomial. The items of information shown here are not always independent of one another. This follows from Propositions 16.5 and 16.6. The values obtained are all simple functions of classical invariants, of the number of components, $\#L$, of L, of homology data of various covers branched over the link, of the Arf invariant of L, and some intricacies of sign. Immediate additional information can, of course, be obtained by changing the signs of variables and also by taking complex conjugates. It is thought that it may not be possible to produce any more such valuations. This is because complexity theory (see [50]) suggests that bounds on the length of the evaluation process at other choices of the variables may not be expressible as a polynomial in the number of crossings of a link diagram.

That many of these specific evaluations should exist was first suggested by Jones. Once it is suspected what one of these evaluations ought to be, it is usually not too hard to give a proof of the result. That has already been done here in the case of the Arf invariant in Chapter 10. The nature of the proofs is always the same: just check that the postulated evaluation satisfies the relevant skein formula for the given values of the polynomial variables. The proof is then an exercise in the

TABLE 16.3. Evaluations of Polynomials

	P(L) (l, m)	**V(L)** t	**F(L)** (a, z)	**Value**
A	$(l, -l-l^{-1})$	$e^{4\pi i/3}$	$(1, 1)$	1
		1	$(1, -2)$	$(-2)^{\#L-1}$
		$e^{2\pi i/3}$	(i, z)	$(-1)^{\#L-1}$
B	$(i, -2)$	-1		$\Delta_L(-1)$
			$(1, 2)$	$(\det L)^2$
C	$(1, \sqrt{2})$	i		$(-\sqrt{2})^{\#L-1}(-1)^{\mathrm{Arf}(L)}$
				0 if $\mathrm{Arf}(L)$ undefined
D	$(1, 1)$			$(i\sqrt{2})^{d_2(T(L))}$
E	$(e^{\pi i/6}, 1)$	$e^{\pi i/3}$		$\delta_3 i^{\#L-1}(i\sqrt{3})^{d_3(D(L))}$
			$(1, -1)$	$(-3)^{d_3(D(L))}$
F			$(1, \frac{\sqrt{5}-1}{2})$	$\delta_5\sqrt{5}^{\,d_5(D(L))}$
G			$(-q, q^{-1}+q)$	$\frac{1}{2}(-1)^{\#L-1}\sum_{X\subset L} q^{4\mathrm{lk}(X, L-X)}$

understanding of some classical invariant (the Arf invariant or the homology of a branched cover).

Some explanation of the notation used in the table of evaluations is required. The results of row A of the table are elementary. They do however assert that the number of components of a link is incorporated in its polynomial invariants. Row B refers to $\Delta_L(-1)$, the value at $t = -1$ of the Conway-normalised Alexander polynomial, $\det L = |\Delta_L(-1)|$ (see Chapter 9). For a knot K, $\det K$ is the order of $H_1(D(K); \mathbb{Z})$, where $D(K)$ is the double cover of S^3 branched over K (for a link, the determinant is zero if the homology group is infinite). Row C has already been discussed in Chapter 10. In the remaining rows, $D(L)$ is again the double branched cover and $T(L)$ is the threefold cyclic cover of S^3 branched over oriented link L. The prefix d_r denotes the dimension, as a $\mathbb{Z}/r\mathbb{Z}$-vector space, of the first homology with $\mathbb{Z}/r\mathbb{Z}$ coefficients of the space in question. Details of proofs relating to row D and row E can be found in [88]. The coefficients δ_3 and δ_5 are both ± 1 and can be evaluated in terms of Legendre symbols (see [91]). Row G is a result to be found in [89]; the summation is over all sublinks X of L, including the empty sublink and L itself, and $\text{lk}(X, L - X)$ is the sum of the linking numbers of every component of X with every component of $L - X$.

The representation of links as closed braids (see Chapter 1) was the original starting point for the invention of the Jones polynomial [53]. Fundamental were the theorems of Alexander and Markov that, combined, constitute the following proposition. Modern proofs can be found in [7] and [98].

Proposition 16.10. *Any oriented link in S^3 is the closure $\widehat{\xi}$ of some ξ belonging to the braid group B_n, for some n. Oriented links $\widehat{\xi}$ and $\widehat{\eta}$ are equivalent if ξ and η differ by a sequence of (Markov) moves of the following two types and inverses of such moves:*

(i) *Change an element of B_n to a conjugate element in that group;*

(ii) *Change $\xi \in B_n$ to $i_n(\xi)\sigma_n^{\pm 1} \in B_{n+1}$, where $i_n : B_n \to B_{n+1}$ is the inclusion (that disregards the $(n+1)$th string).*

The braid approach has also been used in an entirely different way to give another existence proof for the HOMFLY polynomial [52], [126]. (A version also works for the Kauffman polynomial [126].) This method, which involves "R-matrices" and the Yang–Baxter equations, is of particular interest as it employs some of the same mathematics as is used in quantum statistical mechanics [6]. The method is amenable to considerable extension, abstraction and generalisation; so much so, in fact, that it has lead to the birth of the subject of quantum groups, now a branch of abstract algebra. That subject is now the main topic of several books, for example [127] and [56]. The method can also be interpreted in terms of a "states model" for a sequence of values of the HOMFLY polynomial. This gives a complete, immediate (though complicated) definition of $P(L)_{(iq^{-(m+1)},\, i(q-q^{-1}))}$ as a Laurent polynomial in q, without recourse to any existence theorem. A brief outline follows.

Let V be a free module with base $e_1, e_2, \ldots, e_m$ over a commutative ring K. As usual, let $V^{\otimes n}$ denote the n fold tensor product $V \otimes V \otimes \cdots \otimes V$. Suppose

$R : V \otimes V \to V \otimes V$ is an automorphism; in suffix notation R maps $e_i \otimes e_j$ to $R_{i,j}^{p,q} e_p \otimes e_q$, summing over the repeated suffices. Let $R_i : V^{\otimes n} \to V^{\otimes n}$ be $1 \otimes 1 \otimes \cdots \otimes 1 \otimes R \otimes 1 \otimes \cdots \otimes 1$, where the R operates on the tensor product of the ith and $(i+1)$th copies of V. This R is called a Yang–Baxter operator if it satisfies the (quantum) Yang–Baxter equations

$$R_1 R_2 R_1 = R_2 R_1 R_2.$$

Suppose that $\mu : V \to V$ is represented by the diagonal matrix $\operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_m)$, and

$$R(\mu \otimes \mu) = (\mu \otimes \mu)R, \quad \sum_j R_{i,j}^{k,j} \mu_j = \alpha\beta\delta_i^k \text{ and } \sum_j (R^{-1})_{i,j}^{k,j} \mu_j = \alpha^{-1}\beta\delta_i^k,$$

where α and β are fixed units in $\mathcal{K}$. If this occurs, μ is called an enhancement of R. Often it is not too difficult to find such a μ once a solution is known for the Yang–Baxter equations. Given such R and μ, a representation of the braid group can be found as follows: Define $\phi : B_n \to \operatorname{Aut} V^{\otimes n}$ by $\phi(\sigma_i) = R_i$. The Yang–Baxter equations imply that ϕ is compatible with the relations of the braid group (quoted in Chapter 1), and so ϕ gives a well-defined group homomorphism. Define $T : \bigcup_n B_n \to \mathcal{K}$ by

$$T(\xi) = \alpha^{-w(\xi)} \beta^{-n} \operatorname{Trace}(\phi(\xi)\mu^{\otimes n}),$$

where $\xi \in B_n$ and $w : B_n \to \mathbb{Z}$ is the homomorphism defined by $w(\sigma_i) = 1$.

Theorem 16.11. *If an oriented link L is the closure of $\xi \in B_n$, let $T(L)$ be defined to be $T(\xi)$. This is a well-defined link invariant.*

Proof. Because T is essentially a trace function, if $\xi, \eta \in B_n$ then $T(\eta^{-1}\xi\eta) = T(\xi)$. Using the properties of μ, it is easy to show that $T(\xi\sigma_n) = T(\xi\sigma_n^{-1}) = T(\xi)$. The result then follows from Proposition 16.10. □

Proposition 16.12. *Suppose the minimal polynomial equation satisfied by the automorphism $R : V \otimes V \to V \otimes V$ is $\sum_{i=p}^{q} k_i R^i = 0$ for some $k_i \in \mathcal{K}$. Then $\sum_{i=p}^{q} k_i \alpha^i T(L_i) = 0$ whenever L_i are links identical except near a point where L_i has a "tassel" of i crossings.*

Proof. The "tassel" of i crossings can be taken to be an occurrence of σ_1^i in a braid word representing L_i. Now if $\eta \in B_n$, the result follows from

$$T(\sigma_1^i \eta) = \alpha^{-i-w(\eta)} \beta^{-n} \operatorname{Trace}(R^i \phi(\eta)\mu^{\otimes n}).$$ □

The example that leads to the HOMFLY polynomial is as follows: Let $\mathcal{K}$ be $\mathbb{Z}[q^{-1}, q]$ and let $m \geq 1$. Let $E_{i,j}$ be the endomorphism of V that maps e_i to e_j and maps the other base elements to zero. A solution to the Yang–Baxter equations is given by

$$R = -q \sum_i E_{i,i} \otimes E_{i,i} + \sum_{i \neq j} E_{i,j} \otimes E_{j,i} + (q^{-1} - q) \sum_{i < j} E_{i,i} \otimes E_{j,j}.$$

It is arduous but straightforward to check directly that this is a solution. It is, however, not hard to see that

$$R^{-1} = -q^{-1}\sum_i E_{i,i} \otimes E_{i,i} + \sum_{i \neq j} E_{i,j} \otimes E_{j,i} + (q - q^{-1})\sum_{i>j} E_{i,i} \otimes E_{j,j}.$$

so that $R - R^{-1} = (q^{-1} - q)1_{V\otimes V}$. This is then the minimal polynomial equation for R. Let $\mu = \mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_m)$ where $\mu_i = q^{2i-m-1}$, and let $\alpha = -q^m$ and $\beta = 1$. A routine check shows that this provides an enhancement for R. Thus by Theorem 16.11, these data provide an oriented link invariant $T(L)$ which, by Proposition 16.12, satisfies

$$q^m T(L_+) - q^{-m} T(L_-) + (q^{-1} - q)T(L_0) = 0,$$

where L_+, L_- and L_0 are related in the usual way. Re-normalising to get the polynomial of the unknot to be one, gives for $\xi \in B_n$,

$$P(\widehat{\xi})_{(iq^m, i(q^{-1}-q))} = (-q)^{-mw(\xi)} \frac{\mathrm{Trace}(\phi(\xi)\mu^{\otimes n})}{\mathrm{Trace}\,\mu}.$$

As m varies, the evaluations of the HOMFLY polynomial at these special values of the variables do, of course, determine the whole two-variable polynomial. It is interesting to note that the Alexander polynomial $\Delta_L(t) = P(L)_{(i, i(t^{1/2}-t^{-1/2}))}$ does not feature as one of the special values (as $m \geq 1$); from this standpoint $\Delta_L(t)$ occurs only by way of creating the entire two-variable polynomial from the whole sequence of special values. A full version of this Yang–Baxter equation approach to the HOMFLY and Kauffman polynomials is given in [127]. More complicated R-matrices lead to descriptions of those invariants for "coloured" links that are linear combinations of invariants for satellites and parallels.

From the above example, Jones [52] produced a "states model" for each of the above values of the HOMFLY polynomial. His result is as follows: Fix $n \geq 0$, let D be a diagram of an oriented link L, and let $D^\star$ be D less the crossings of D. A map s : {segments of $D^\star$} $\to \{-n, -n+2, -n+4, \ldots, n-2, n\}$ is a *labelling* of D if near each crossing the values of s conform to one of the three types shown in Figure 16.2.

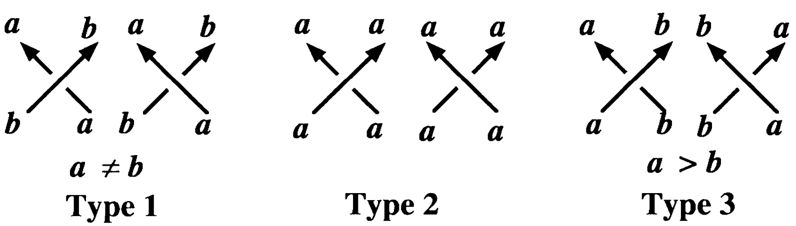


Figure 16.2

Let $|C_i^{\pm}s|$ be the number of $\pm$crossings of type i with respect to s for $i \in \{1, 2, 3\}$, and let $|C_i s| = |C_i^+ s| + |C_i^- s|$. Then

$$(-1)^{1+\mathrm{rot}(D)}(q^{-n} + q^{-n+2} + \cdots + q^{n})P(L)_{(iq^{-(n+1)}, i(q-q^{-1}))}$$
$$= q^{(n+1)w(D)} \sum_s (-1)^{|C_3^- s|+|C_2 s|} q^{|C_2^- s|-|C_2^+ s|-\int s\, d\theta}(q - q^{-1})^{|C_3 s|}.$$

In this formula, $\int s\, d\theta$ is the integer obtained by smoothing D (that is, changing $D_{\pm}$ to D_0) at all crossings of types 2 and 3 (so that each resulting link component has a constant label) and subtracting the sum of the labels on clockwise components from the sum of those on anti-clockwise components. The term rot(D) is $\int 1\, d\theta$, where 1 is the labelling with constant value 1; the summation is over all labellings s. This formula could be used as a definition of $P(L)_{(iq^{-(n+1)}, i(q-q^{-1}))}$ and then the (L_+, L_-, L_0)-formula and invariance under Reidemeister moves could be checked directly. Note that, although this "states model" was certainly derived by means of consideration of representations of the braid group, the final formulation contains no mention of braids.

The structure and significance of the HOMFLY and Kauffman polynomials are frequently interpreted in the language of *Vassiliev invariants*, sometimes called *invariants of finite type*. A final remark will give the rough idea of these, but for many more details, see [5] and [10]. Suppose V is any invariant of oriented links taking values in some abelian group. This V be can extended to be an invariant of singular links in the following way: A singular link is an immersion of simple closed curves in S^3 with finitely many transverse double points. These self-intersections are required to remain transverse in any isotopy demonstrating the equivalence of such singular links. If the definition of V has been extended over singular links with $n - 1$ double points, define it on a singular link $L_{\times}$ with n singularities by

$$V(L_{\times}) = V(L_+) - V(L_-),$$

where $V(L_{\times})$, $V(L_+)$ and $V(L_-)$ are identical except near a point where they are as in Figure 16.2. Note that $V(L_+)$ and $V(L_-)$ each has $n - 1$ double points. Then V is called a Vassiliev invariant of order n, or an invariant of finite type n, if $V(L) = 0$ for every L with $n + 1$ or more singularities.

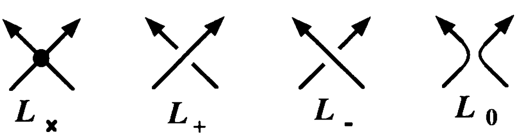


Figure 16.3

Recall the Conway polynomial invariant, $\nabla_L(z) \in \mathbb{Z}[z]$, of oriented links defined by $\nabla_{\mathrm{unknot}}(z) = 1$ and

$$\nabla_{L_+}(z) \quad \nabla_{L_-}(z) \quad z\nabla_{L_0}(z).$$

Extend this over singular links by the above method. Then if $L_\times$ is a link with r singularities, $\nabla_{L_\times}(z) = z\nabla_{L_0}(z)$ where L_0 is a link with $r-1$ singularities. Thus by induction on r, if L has r singularities then $\nabla_L(z)$ has a factor of z^r. This implies at once that the coefficient of z^n in the Conway polynomial of a link is a Vassiliev invariant of order n.

Now suppose one considers the HOMFLY polynomial and makes the substitution $(l, m) = (it^{N/2}, i(t^{-1/2} - t^{1/2}))$. The characterising skein relation becomes

$$t^{N/2}P(L_+) - t^{-N/2}P(L_-) = (t^{1/2} - t^{-1/2})P(L_0).$$

Note that this becomes the Jones polynomial when $N = 2$. Now make the further substitution $t = \exp x$. Here $\exp x$ should be thought of as the classical power series expansion. Of course, $\exp x/2$ and $\exp(-x/2)$ have power series expansions; power series can be multiplied and added to give power series. Thus $P(L)$ has a power series expansion in powers of x. It follows immediately that $P(L_+) - P(L_-) = xS(x)$ for some power series $S(x)$. Hence the proof used above for the Conway polynomial shows at once that the coefficient of x^n in the power series expansion of $P(L)$ is a Vassiliev invariant of order n.

Vassiliev invariants have attracted much attention, partly because they seem to give a structured view of the polynomial invariants discussed here. They also have associated with them a pleasing blend of linear algebra and diagrammatic combinatorics and an interaction with Lie algebras. This is described in some detail in [5]. They can also be interpreted in terms of the configuration space of immersions of closed curves into S^3 ([129], [130]).

Exercises

1. Generalise to the theories of the HOMFLY and Kauffman polynomials the "numerator" and "denominator" ideas that work so neatly for the Conway polynomial (see Exercise 4 of Chapter 8).

2. Prove Proposition 16.9 concerning the "first" terms in the HOMFLY and Kauffman polynomials.

3. Suppose that in a diagram of an oriented knot K, some crossings labelled $1, 2, \ldots, n$, with crossing i having sign ϵ_i, are changed one by one to obtain the unknot. Let K_i be the knot created from K by the first i changes so that $K = K_0$ and K_n is the unknot. Let L_i be the oriented two-component link obtained from K_{i-1} by nullifying the crossing i (that is, by replacing it with no crossing in the way that respects orientations). Let the total twisting $\tau(K)$ of K be defined by $\tau(K) = \sum_{i=1}^n \epsilon_i \operatorname{lk}(L_i)$, where $\operatorname{lk}(L_i)$ is the linking number of the two components of L_i. If the HOMFLY polynomial of K is written $P(K) = \sum_i p_i(l)m^i$, prove that the derivatives of the Laurent polynomial $p_0(l)$ satisfy
(i) $p_0'(\sqrt{-1}) = 0$,
(ii) $p_0''(\sqrt{-1}) = 8\tau(K)$.
Deduce that $\tau(K)$ is well defined independent of the chosen crossing changes

4. In the notation of the previous question, show that $p_2(\sqrt{-1}) = -\tau(K)$.

5. The HOMFLY skein of the solid torus $S^1 \times D^2$, denoted $\mathcal{S}_H(S^1 \times D^2)$, is the free module over $\mathbb{Z}[l^{\pm 1}, m^{\pm 1}]$ generated by all oriented links in $S^1 \times D^2$ quotiented by all relations of the form $lL_+ + l^{-1}L_- + mL_0 = 0$, where L_+, L_- and L_0 are related in the usual way. Show that embedding two solid tori in one (by taking the product with S^1 of two discs embedded in one disc) induces a product structure on $\mathcal{S}_H(S^1 \times D^2)$ that turns it into a commutative algebra. Show that this algebra is generated by the closure in the solid torus of all braids of the form $\sigma_1\sigma_2 \dots \sigma_r$ with either orientation. Consider the order-2 homeomorphism of $S^1 \times D^2$ to itself that rotates the solid torus through an angle π about an axis meeting it in two intervals (h reverses the S^1 direction but preserves the orientation of the solid torus). Show that h induces a linear map of $\mathcal{S}_H(S^1 \times D^2)$ which sends one of the above generators to itself but with reversed direction. Show that this fact can be used to construct different links with the same HOMFLY polynomial by rotating a solid torus containing some components of a link and changing their directions.

6. Suppose that K_1 and K_2 are two oriented knots that are related by a mutation. For $i = 1, 2$, let $K_i^{(2)}$ be the 2-parallel of K_i (that is, the link consisting of K_i and a longitude of K_i with parallel orientation). Prove that $P(K_1) = P(K_2)$. Extend this result to anti-parallels, where now the other orientation is chosen for the longitude.

7. Prove the theorem of Alexander that asserts that any oriented link in S^3 is the closure of some element of some braid group B_n.

8. Suppose that V is a free module of *dimension 2* over $\mathbb{Z}[q^{-1}, q]$ and that $R : V \otimes V \to V \otimes V$ is defined by

$$R = -q \sum_i E_{i,i} \otimes E_{i,i} + \sum_{i \neq j} E_{i,j} \otimes E_{j,i} + (q^{-1} - q) \sum_{i<j} E_{i,i} \otimes E_{j,j} .$$

Show that R satisfies the Yang–Baxter equation

$$R_1 R_2 R_1 = R_2 R_1 R_2.$$

References

[1] C. C. Adams. *The Knot Book*, W. H. Freeman (1994).

[2] C. C. Adams. Toroidally alternating knots and links, *Topology* **33** (1994) 353–369.

[3] J. W. Alexander. Topological invariants of knots and links, *Trans. Amer. Math. Soc.* **30** (1928) 275–306.

[4] M. F. Atiyah. *The Geometry and Physics of Knots*, Cambridge University Press (1990).

[5] D. Bar-Natan. On the Vassiliev knot invariants, *Topology* **34** (1995) 423–472.

[6] R. J. Baxter. *Exactly Solved Models in Statistical Mechanics*, Academic Press London (1982).

[7] J. S. Birman. Braids, links and mapping class groups, *Ann. of Math. Studies* **82** Princeton University Press (1974).

[8] J. S. Birman and D. R. J. Chillingworth. On the homeotopy group of a non-orientable surface, *Proc. Cambridge Philos. Soc.* **71** (1972) 437–448.

[9] J. S. Birman and A. Libgober. Braids, *Contemporary Math.*, **78** Amer. Math. Soc., Providence, R.I., (1988).

[10] J. S. Birman and X.-S. Lin. Knot polynomials and Vassiliev invariants, *Invent. Math.* **111** (1993) 225–270.

[11] C. Blanchet. *Invariants on three-manifolds with spin structure*, *Comment. Math. Helv.* **67** (1992) 406–427.

[12] C. Blanchet, N. Habegger, G. Masbaum and P. Vogel. Three-manifold invariants derived from the Kauffman bracket, *Topology* **31** (1992) 685–699.

[13] C. Blanchet, N. Habegger, G. Masbaum and P. Vogel. Topological quantum field theories derived from the Kauffman bracket, *Topology* **34** (1995) 883–927.

[14] F. Bonahon and L. C. Siebenmann. The characteristic tori splitting of irreducible compact 3-orbifolds, *Math. Ann.* **278** (1987) 441–479.

[15] F. Bonahon and L. C. Siebenmann. New geometric splittings of classical knots, *London Math. Soc. Lecture Notes* **75** (to appear), Cambridge University Press.

[16] R. D. Brandt, W. B. R. Lickorish and K. C. Millett. A polynomial invariant for unoriented knots and links, *Invent. Math.* **84** (1986) 563–573.

[17] G. Burde and H. Zieschang. *Knots*, de Gruyter (1986).

[18] A. J. Casson and C. M. Gordon. On slice knots in dimension three, *Geometric Topology. Proc. Symp. Pure Math.* XXXII Amer. Math. Soc. Providence, R.I. (1978) 39–53.

[19] D. R. J. Chillingworth. A finite set of generators for the homeotopy group of a non-orientable surface, *Proc. Cambridge Philos. Soc.* **65** (1969) 409–430.

[20] J. H. Conway. "An enumeration of knots and links," *Computational problems in abstract algebra*, Ed. J. Leech, Pergamon Press, (1969), 329–358.
[21] D. Cooper. The universal abelian cover of a link, *London Math. Soc. Lecure Notes* **48** (1982) 51–66.
[22] R. H. Crowell. Genus of alternating link types, *Ann. of Math.* **69** (1959) 258–275.
[23] R. H. Crowell and R. H. Fox. *Introduction to Knot Theory*, Graduate Texts in Mathematics **57** Springer-Verlag (1977).
[24] M. Dehn. Die Groupe der Abbildungsklassen, *Acta Math.* **69** (1938) 135–206.
[25] S. K. Donaldson. The Seiberg-Witten equations and 4-manifold topology, *Bull. Amer. Math. Soc.* **33** (1996) 45–70.
[26] C. H. Dowker and M. B. Thistlethwaite. On the classification of knots, *C. R. Math. Rep. Acad. Sci. Canada* **IV** (1982) 129–131.
[27] J. Franks and R. Williams. Braids and the Jones polynomial, *Trans. Amer. Math. Soc.* **303** (1987) 97–107.
[28] D. S. Freed and R. E. Gompf. Computer calculations of Witten's 3-manifold invariants, Comm. Math. Phys. **141** (1991) 79–117.
[29] M. H. Freedman and R. C. Kirby. A geometric proof of Rohlin's theorem, *Proc. Symp. Pure Math.* **32** (1978) 85–97.
[30] M. H. Freedman and F. S. Quinn. *Topology of 4-manifolds*, Princeton Math. Ser. **39**, Princeton University Press, N.J. (1990).
[31] P. Freyd, D. Yetter, J. Hoste, W. B. R. Lickorish, K. Millett and A. Ocneanu. A new polynomial invariant of knots and links, *Bull. Amer. Math. Soc.* **12** (1985) 239–246.
[32] D. Gabai. Genera of arborescent links, *Mem. Amer. Math. Soc.* **339** (1986) 1–98.
[33] N. D. Gilbert and T. Porter. *Knots and Surfaces*, Oxford University Press (1994).
[34] L. Goeritz. Knotten und quadratische Formen, *Math. Z.* **36** (1933) 647–654.
[35] C. M. Gordon. "Some aspects of classical knot theory," *Lecture Notes in Mathematics* **685**, Springer-Verlag (1978) 1–60.
[36] C. M. Gordon and R. A. Litherland. On the signature of a link, *Invent. Math.* **47** (1978) 53–69.
[37] C. M. Gordon and J. Luecke. Knots are determined by their complements, *J. Amer. Math. Soc.* **2** (1989) 371–415.
[38] B. Hartley and T. O. Hawkes. *Rings, Modules and Linear Algebra*, Chapman and Hall (1970).
[39] R. Hartley. The Conway potential function for links, *Comment. Math. Helv.* **58** (1983) 365–378.
[40] R. Hartley. Identifying non-invertible knots, *Topology* **22** (1983) 137–145.
[41] C. Hayashi. Links with alternating diagrams on closed surfaces of positive genus, *Math. Proc. Cambridge Philos. Soc.* **117** (1995) 113–128.
[42] G. Hemion. *The Classification of Knots and 3-Dimensional Spaces*, Oxford University Press (1992).
[43] J. Hempel. 3-manifolds, *Ann. of Math. Studies* **86** Princeton University Press (1976).
[44] J. A. Hillman. Alexander ideals of links, *Lecture Notes in Math.* **895** Springer-Verlag (1981).
[45] C. F. Ho. A new polynomial of knots and links - preliminary report, *Abstracts Amer. Math. Soc.* **6** (1985) 300.
[46] C. D. Hodgson and J. H. Rubinstein. Involutions and isotopies of lens spaces, *Lecture Notes in Math.* **1144** Springer-Verlag (1985) 60–96.
[47] J. F. P. Hudson. *Lecture Notes on Piecewise Linear Topology*, Benjamin New York (1969)

[48] S. P. Humphries. Generators for the mapping class group, *Topology of low-dimensional manifolds*, Ed. R. A. Fenn, Springer-Verlag (1979) 44–47.

[49] W. H. Jaco. Lectures on three-manifold topology, *C.B.M.S. Regional Conference Series* **43** Amer. Math. Soc. Providence R.I. (1980).

[50] F. Jaeger, D. L. Vertigan and D. J. A. Welsh. On the computational complexity of the Jones and Tutte polynomials, *Math. Proc. Cambridge Philos. Soc.* **108** (1990) 35–53.

[51] L. C. Jeffrey. Chern–Simons Witten invariants of lens spaces and torus bundles, and the semi-classical approximation, *Comm. Math. Phys.* **147** (1992) 563–604.

[52] V. F. R. Jones. Notes on a talk to the Atiyah seminar (1986).

[53] V. F. R. Jones. Hecke algebra representations of braid groups and link polynomials, *Ann. of Math.* **126** (1987) 335–388.

[54] T. Kanenobu. Infinitely many knots with the same polynomial, *Proc. Amer. Math. Soc.* **97** (1986) 158–161.

[55] J. Kania-Bartoszynska. Examples of different 3-manifolds with the same invariants of Witten and Reshetikhin, *Topology* **32** (1993) 47–54.

[56] C. Kassel. *Quantum Groups*, Springer-Verlag (1995).

[57] L. H. Kauffman. Formal Knot Theory, *Math. Notes* **30** Princeton University Press (1983).

[58] L. H. Kauffman. On knots, *Ann. of Math. Studies* **115** Princeton University Press (1987).

[59] L. H. Kauffman. States models and the Jones polynomial, *Topology* **26** (1987) 395–407.

[60] L. H. Kauffman. An invariant of regular isotopy, *Trans. Amer. Math. Soc.* **318** (1990) 417–471.

[61] L. H. Kauffman. *Knots and physics*, World Scientific (1991).

[62] L. H. Kauffman and S. Lins. Temperley–Lieb recoupling theory and invariants of 3-manifolds, *Ann. of Math. Studies*, **134** Princeton University Press (1994).

[63] A. Kawauchi. The invertibility problem for amphicheiral excellent knots, *Proc. Japan Acad.* **55** (1979) 399–402.

[64] M. E. Kidwell. On the degree of the Brandt-Lickorish-Millett polynomial, *Proc. Amer. Math. Soc.* **100** (1987) 755–762.

[65] R. C. Kirby. A calculus for framed links in S^3, *Invent. Math.* **45** (1978) 35–56.

[66] R. C. Kirby. The topology of 4-manifolds, *Lecture Notes in Math.* **1374**, Springer Verlag (1989).

[67] R. C. Kirby. Problems in low dimensional topology, *Geometric Topology*, Ed. W. H. Kazez, Amer. Math. Soc. (1997).

[68] R. C. Kirby and P. Melvin. The 3-manifold invariants of Witten and Reshetikhin-Turaev for $sl(2, \mathbb{C})$, *Invent. Math.* **105** (1991) 473–545.

[69] A. N. Kirillov and N. Y. Reshetikhin. "Representations of the algebra $U_q(Sl(2))$, q-orthogonal polynomials and invariants of links," Infinite dimensional Lie algebras and groups *Adv. Ser. Math. Phys.* 7, World Scientific (1989) 285–339.

[70] P. Kirk and C. Livingston. *Twisted knot polynomials, inversion, mutation and concordance*, Indiana University preprint (1996).

[71] T. Kohno. *Tunnel numbers of knots and Jones–Witten invariants. Braid groups, knot theory and statistical mechanics II*, *Adv. Ser. Math. Phys.* **17**, World Scientific (1994) 275–293.

[72] P. B. Kronheimer and T. S. Mrowka. Gauge theory for embedded surfaces I, *Topology* **32** (1993) 773–826.

[73] M. Lackenby. Fox's congruence classes and the quantum-$SU(2)$ invariants of links in 3-manifolds, *Comment. Math. Helv.* **71** (1996) 664–677.

[74] S. Lang. *Algebraic Number Theory*, Springer-Verlag (1986).

[75] J. Levine. Polynomial invariants of knots of codimension 2, *Ann. of Math.* **84** (1966) 537–544.

[76] B.-H. Li and T.-J. Li. Generalized Gaussian sums and Chern–Simons–Witten–Jones invariants of lens spaces, *J. Knot Theory and its Ramifications* **5** (1996) 183–224.

[77] W. B. R. Lickorish. A representation of orientable combinatorial 3-manifolds, *Ann. of Math.* **76** (1962) 531–540.

[78] W. B. R. Lickorish. A finite set of generators for the homeotopy group of a 2-manifold, *Proc. Cambridge Philos. Soc.* **60** (1964) 769–778.

[79] W. B. R. Lickorish. On the homeomorphisms of a non-orientable surface, *Proc. Cambridge Philos. Soc.* **61** (1965) 61–64.

[80] W. B. R. Lickorish. A finite set of generators for the homeotopy group of a 2-manifold. (Corrigendum), *Proc. Cambridge Philos. Soc.* **62** (1966) 679–681.

[81] W. B. R. Lickorish. The irreducibility of the three-sphere, *Michigan Math. J.* **36** (1989) 345–349.

[82] W. B. R. Lickorish. Invariants for 3-manifolds from the combinatorics of the Jones polynomial, *Pacific J. Math.* **149** (1991) 337–347.

[83] W. B. R. Lickorish. *Three-manifolds and the Temperley–Lieb algebra*, Math. Ann. **290** (1991) 657–670.

[84] W. B. R. Lickorish. Calculations with the Temperley-Lieb algebra, *Comment. Math. Helv.* **67** (1992) 571–591.

[85] W. B. R. Lickorish. Distinct 3-manifolds with all $SU(2)_q$ invariants the same, *Proc. Amer. Math. Soc.* **117** (1993) 285–292.

[86] W. B. R. Lickorish. The skein method for three-manifold invariants, *J. Knot Theory and its Ramifications* **2** (1993) 171–194.

[87] W. B. R. Lickorish. Skeins and handlebodies, *Pacific J. Math.* **159** (1993) 337–349.

[88] W. B. R. Lickorish and K. C. Millett. Some evaluations of link polynomials, *Comment. Math. Helv.* **61** (1986) 349–359.

[89] W. B. R. Lickorish and K. C. Millett. An evaluation of the F-polynomial of a link, *Lecture Notes in Math.* **1350**, Springer Verlag (1987) 104–108.

[90] W. B. R. Lickorish and K. C. Millett. A polynomial invariant of oriented links, *Topology* **26** (1987) 107–141.

[91] A. S. Lipson. An evaluation of a link polynomial, *Math. Proc. Cambridge Philos. Soc.* **100** (1986) 361–364.

[92] C. N. Little. Non-alternate ± knots, *Trans. Roy. Soc. Edinburgh* **39** (1900) 771–778.

[93] C. Livingston. *Knot theory*, Carus Mathematical Monographs **24**, Math. Assoc. Amer. (1993).

[94] W. W. Menasco. Closed incompressible surfaces in alternating knot and link complements, *Topology* **23** (1984) 37–44.

[95] W. W. Menasco and M. B. Thistlethwaite. The classification of alternating links, *Ann. of Math.* **138** (1993) 113–171.

[96] K. Morimoto, M. Sakuma and Y. Yokota. Identifying tunnel number one knots, *J. Math. Soc.* Japan **48** (1996) 667–688.

[97] H. R. Morton. Seifert circles and knot polynomials, *Math. Proc. Cambridge Philos. Soc.* **99** (1986) 107–109.

[98] H. R. Morton. Threading knot diagrams, *Math. Proc. Cambridge Philos. Soc.* **99** (1986) 247–260.

[99] H. Murakami. A weight system derived from the Conway potential function, *J. London Math. Soc.* (1997) (to appear).

[100] J. Murakami. A state model for the multi-variable Alexander polynomial, *Pacific J. Math.* **157** (1993) 109–135.

[101] K. Murasugi. On a certain numerical invariant of link types, *Trans Amer. Math. Soc.* **117** (1965) 387–422.

[102] K. Murasugi. On the signature of links, *Topology* **9** (1970) 283–298.

[103] Y. Nakanishi. A note on unknotting number, *Math. Sem. Notes Kobe Univ.* **9** (1981) 99–108.

[104] J. R. Neil. Combinatorial calculations of the various normalisations of the Witten invariants for 3-manifolds, *J. Knot Theory and its Ramifications* **1** (1992) 407–499.

[105] C. D. Papakyriakopoulos. On Dehn's lemma and the asphericity of knots, *Ann. of Math.* **66** (1957) 1–26.

[106] J. H. Przytycki and P. Traczyk. Invariants of links of Conway type, *Kobe J. Math.* **4** (1987) 115–139.

[107] K. Reidemeister. *Knotentheorie*, Springer-Verlag New York (1948).

[108] K. Reidemeister. *Knot Theory (Translation of Knotentheorie)*, BSC Associates Moscow, Idaho (1983).

[109] N. Y. Reshetikhin and V. G. Turaev. Invariants of 3-manifolds via link polynomials and quantum groups, *Invent. Math.* **103** (1991) 547–597.

[110] R. Riley. Homomorphisms of knot groups on finite groups, *Math. Comp.* **25** (1971) 603–619.

[111] R. A. Robertello. An invariant of knot cobordism, *Comm. Pure and App. Math.* **18** (1965) 543–555.

[112] D. Rolfsen. *Knots and Links*, Publish or Perish (1976).

[113] C. P. Rourke and B. J. Sanderson. Introduction to piecewise-linear topology, *Ergebnisse der mathematik* **69** Springer-Verlag (1972).

[114] C. P. Rourke and D. P. Sullivan. On the Kervaire construction, *Ann. of Math.* **94** (1971) 397–413.

[115] H. Schubert. Knotten und Vollringe, *Acta Math.* **90** (1953) 131–286.

[116] R. A. Stong. The Jones polynomial of parallels and applications to crossing number, *Pacific J. Math.* **164** (1994) 383–395.

[117] P. M. Strickland. On the quantum group invariants of cables, Preprint (University of Liverpool) (1990).

[118] P. G. Tait. On knots I, II, III., *Scientific papers* I. Cambridge University Press London (1898) 273–437.

[119] M. B. Thistlethwaite. Kauffman's polynomial and alternating links, *Topology* **27** (1988) 311–318.

[120] M. B. Thistlethwaite. On the Kauffman polynomial of an adequate link, *Invent. Math.* **93** (1988) 285–298.

[121] W. P. Thurston. *Geometry and topology of 3-manifolds*, Princeton University notes (1979).

[122] P. Traczyk. Periodic knots and the skein polynomial, *Invent. Math.* **106** (1991) 73–84.

[123] A. G. Tristram. Some cobordism invariants for links, *Proc. Cambridge Philos. Soc.* **66** (1969) 251–264.

[124] H. F. Trotter. Homology of group systems with applications to knot theory, *Ann. of Math.* **76** (1962) 464–498.

[125] H. F. Trotter. Non-invertible knots exist, *Topology* **2** (1964) 275–280.

[126] V. G. Turaev. The Yang-Baxter equations and invariants of links, *Invent. Math.* **92** (1988) 527–553

[127] V. G. Turaev. *Quantum invariants for knots and 3-manifolds*, de Gruyter Berlin (1994).

[128] V. G. Turaev and H. Wenzl. Quantum invariants of 3-manifolds associated with classical simple Lie algebras, *International J. Math.* **4** (1993) 323–358.

[129] V. A. Vassiliev. Cohomology of knot spaces, "Theory of singularities and its applications" *Amer. Math. Soc.*, Providence (1990)

[130] V. A. Vassiliev. Complements of discriminants of smooth maps: topology and applications, *Amer. Math. Soc. Translations of Math.* **98** (1992).

[131] B. Wajnryb. A simple presentation for the mapping class group of an orientable surface, *Israel J. Math.* **45** (1983) 157–174.

[132] F. Waldhausen. On irreducible 3-manifolds which are sufficiently large, *Ann. of Math.* **87** (1968) 56–88.

[133] H. Wenzl. On sequences of projections, *C. R. Math. Rep. Acad. Sci.* **IX** (1987) 5–9.

[134] W. Whitten. Knot complements and groups, *Topology* **26** (1987) 41–44.

[135] E. Witten. Quantum field theory and Jones' polynomial, *Comm. Math. Phys.* **121** (1989) 351–399.

[136] S. Yamada. The minimum number of Seifert circuits equals the braid index of a link, *Invent. Math.* **89** (1987) 347–356.

[137] S. Yamada. A topological invariant of spatial regular graphs, *Knots* **90**, Ed. A. Kawauchi, de Gruyter, (1992) 447–454.

[138] Y. Yokota. On quantum SU(2) invariants and generalised bridge numbers of knots, *Math. Proc. Cambridge Philos. Soc.* **117** (1995) 545–557.

[139] Y. Yokota. Skeins and quantum $SU(N)$ invariants of 3-manifolds, *Math. Ann.* **307** (1997) 109–138.

[140] E. C. Zeeman. Unknotting combinatorial balls, *Ann. of Math.* **78** (1963) 501–526.

Index