BRAIDS: A SURVEY

Joan S. Birman *
e-mail jb@math.columbia.edu

Tara E. Brendle †
e-mail brendle@math.cornell.edu

December 2, 2004

Abstract

This article is about Artin's braid group $\mathbf{B}_n$ and its role in knot theory. We set ourselves two goals: (i) to provide enough of the essential background so that our review would be accessible to graduate students, and (ii) to focus on those parts of the subject in which major progress was made, or interesting new proofs of known results were discovered, during the past 20 years. A central theme that we try to develop is to show ways in which structure first discovered in the braid groups generalizes to structure in Garside groups, Artin groups and surface mapping class groups. However, the literature is extensive, and for reasons of space our coverage necessarily omits many very interesting developments. Open problems are noted and so-labelled, as we encounter them. A guide to computer software is given together with a 10 page bibliography.

Contents

1 Introduction **3**
1.1 $\mathbf{B}_n$ and $\mathbf{P}_n$ via configuration spaces . . . 3
1.2 $\mathbf{B}_n$ and $\mathbf{P}_n$ via generators and relations . . . 4
1.3 $\mathbf{B}_n$ and $\mathbf{P}_n$ as mapping class groups . . . 5
1.4 Some examples where braiding appears in mathematics, unexpectedly . . . 7
1.4.1 Algebraic geometry . . . 7
1.4.2 Operator algebras . . . 8
1.4.3 Homotopy groups of spheres . . . 9
1.4.4 Robotics . . . 10
1.4.5 Public key cryptography . . . 10

2 From knots to braids **12**
2.1 Closed braids . . . 12
2.2 Alexander's Theorem . . . 13
2.3 Markov's Theorem . . . 17

*The first author acknowledges partial support from the U.S.National Science Foundation under grant number 0405586.

†The second author is partially supported by a VIGRE postdoc under NSF grant number 9983660 to Cornell University.

3 Braid foliations **27**
3.1 The Markov Theorem Without Stabilization (special case: the unknot) . . . 27
3.2 The Markov Theorem Without Stabilization, general case 36
3.3 Braids and contact structures 39

4 Representations of the braid groups **45**
4.1 A brief look at representations of Σ_n 45
4.2 The Burau representation and polynomial invariants of knots. 46
4.3 Hecke algebras representations of braid groups and polynomial invariants of knots 47
4.4 A topological interpretation of the Burau representation 52
4.5 The Lawrence-Krammer representation 53
4.6 Representations of other mapping class groups 56
4.7 Additional representations of $\mathbf{B}_n$. 57

5 The word and conjugacy problems in the braid groups **60**
5.1 The Garside approach, as improved over the years 61
5.2 Generalizations: from $\mathbf{B}_n$ to Garside groups 67
5.3 The new presentation and multiple Garside structures 69
5.4 Artin monoids and their groups 70
5.5 Braid groups and public key cryptography 71
5.6 The Nielsen-Thurston approach to the conjugacy problem in $\mathbf{B}_n$ 72
5.7 Other solutions to the word problem 75

6 A potpourri of miscellaneous results **79**
6.1 Centralizers of braids and roots of braids 79
6.2 Singular braids, the singular braid monoid, and the desingularization map . 79
6.3 The Tits conjecture 80
6.4 Braid groups are torsion-free: a new proof 80

1 Introduction

In a review article, one is obliged to begin with definitions. Braids can be defined by very simple pictures such as the ones in Figure 1. Our braids are illustrated as oriented from left to right, with the strands numbered $1, 2, \ldots, n$ from bottom to top. Whenever it is more convenient, we will also think of braids 'vertically', i.e., oriented from top to bottom, with the strands numbered $1, 2, \ldots, n$ from left to right. Crossings are suggested as they are in a picture of a highway overpass on a map. The identity braid has a canonical representation in which two strands never cross. Multiplication of braids is by juxtaposition, concatenation, isotopy and rescaling.

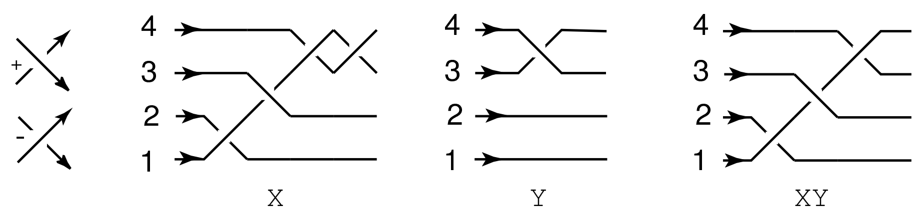


Figure 1: Examples of 4-braids X, Y and their product XY

Pictures like the ones in Figure 1 give an excellent intuitive feeling for the braid group, but one that quickly becomes complicated when one tries to pin down details. What is the ambient space? (It is a slice $\mathbb{R}^2 \times I$ of 3-space.) Are admissible isotopies constrained to $\mathbb{R}^2 \times I$? (Yes, we cannot allow isotopies in which the strands are allowed to loop over the initial points.) Does isotopy mean level-preserving isotopy? (No, it will not matter if we allow more general isotopies in $\mathbb{R}^2 \times I$, as long as strands don't pass through one-another.) Are strands allowed to self-intersect? (No, to allow self-intersections would give the 'homotopy braid group', a proper homomorphic image of the group that is our primary focus.) Can we replace the ambient space by the product of a more general surface and an interval, for example a sphere and an interval? (Yes, it will become obvious shortly how to modify the definition.)

We will bypass these questions and other related ones by giving several more sophisticated definitions. In §1.1, §1.2 and §1.3 we will define the braid group $\mathbf{B}_n$ and pure braid group $\mathbf{P}_n$ in three distinct ways. We will give a proof that two of them yield the same group. References to the literature establish the isomorphism in the remaining case. In §1.4 we will demonstrate the universality of 'braiding' by describing four examples which show how braids have played a role in parts of mathematics which seem far away from knot theory.

1.1 $\mathbf{B}_n$ and $\mathbf{P}_n$ via configuration spaces

We define the topological concept of a braid and of a group of braids via the notion of a configuration space. This approach is nice because it gives, in a concise way, the appropriate equivalence relations and the group law, without any fuss.

The configuration space of n points on the complex plane $\mathbb{C}$ is:

$$\mathcal{C}_{0,\hat{n}} = \mathcal{C}_{0,\hat{n}}(\mathbb{C}) = \{(z_1, \ldots, z_n) \in \mathbb{C} \times \ldots \times \mathbb{C} \mid z_i \neq z_j \text{ if } i \neq j\}.$$

A point on $\mathcal{C}_{0,\hat{n}}$ is denoted by a vector $\vec{z} = (z_1, \ldots, z_n)$. The symmetric group acts freely on $\mathcal{C}_{0,\hat{n}}$, permuting the coordinates in each $\vec{z} \in \mathcal{C}_{0,\hat{n}}$. The orbit space of the action is $\mathcal{C}_{0,n} = \mathcal{C}_{0,\hat{n}}/\Sigma_n$ and the orbit space projection is $\tau : \mathcal{C}_{0,\hat{n}} \to \mathcal{C}_{0,n}$. Choosing a fixed base point $\vec{p} = (p_1, \ldots, p_n)$, we define the pure braid group $\mathbf{P}_n$ on n strands and the braid group $\mathbf{B}_n$ on n strands to be the fundamental groups:

$$\mathbf{P}_n = \pi_1(\mathcal{C}_{0,\hat{n}}, \vec{p}), \qquad \mathbf{B}_n = \pi_1(\mathcal{C}_{0,n}, \tau(\vec{p})).$$

At first encounter $\pi_1(\mathcal{C}_{0,n}, \tau(\vec{p}))$ doesn't look as if it has much to do with braids as we illustrated them in Figure 1, but in fact there is a simple interpretation which reveals the intuitive picture. While the manifold $\mathcal{C}_{0,n}$ has dimension $2n$, the fact that the points $z_1, \ldots, z_n$ are pairwise distinct allows us to think of a point $\vec{z} \in \mathcal{C}_{0,n}$ as a set of n distinct points on $\mathbb{C}$. An element of $\pi_1(\mathcal{C}_{0,n}, \tau(\vec{p}))$ is then represented by a loop which lifts uniquely to a path $\vec{g} : I \to \mathcal{C}_{0,\hat{n}}$, where $\vec{g} = \langle g_1, \ldots, g_n \rangle$ consists of n coordinate functions $g_i : I \to \mathbb{C}$ satisfying $g_i(t) \neq g_j(t)$ if $i \neq j$, $t \in [0,1]$, also $\vec{g}(0) = \vec{g}(1) = \vec{p}$, the base point. The graph of the n simultaneous functions $g_1, \ldots, g_n$ is a geometric (pure) braid. The appropriate equivalence relation on geometric braids is captured by simultaneous homotopy of the n simultaneous paths, rel their endpoints, in the configuration space.

The group $\mathbf{B}_n$ is the group which Artin set out to investigate in 1925 in his seminal paper [6] (however he defined it in a less concise way); in the course of his investigations he was led almost immediately to study $\mathbf{P}_n$. Indeed, the two braid groups are related in a very simple way. Let τ_* be the homomorphism on fundamental groups which is induced by the orbit space projection $\tau : \mathcal{C}_{0,\hat{n}} \to \mathcal{C}_{0,n}$. Observe that the orbit space projection is a regular $n!$-sheeted covering space projection, with Σ_n as the group of covering translations. From this is follows that the group $\mathbf{P}_n$ is a subgroup of index $n!$ in $\mathbf{B}_n$, and there is a short exact sequence:

$$1 \to \mathbf{P}_n \xrightarrow{\tau_*} \mathbf{B}_n \to \Sigma_n \longrightarrow 1 \tag{1}$$

1.2 $\mathbf{B}_n$ and $\mathbf{P}_n$ via generators and relations

We give two definitions of the group $\mathbf{B}_n$ by generators and relations. The classical presentation for $\mathbf{B}_n$ first appeared in [6]. We record it now, and will refer back to it many times later. It has generators $\sigma_1, \ldots, \sigma_{n-1}$ and defining relations:

$$\sigma_i\sigma_k = \sigma_k\sigma_i \ \text{ if } \ |i-k| \geq 2, \quad \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}. \tag{2}$$

The elementary braid σ_i is depicted in sketch (i) of Figure 2.

Many years after Artin did his fundamental work, Birman, Ko and Lee discovered a new presentation, which enlarged the set of generators to a more symmetric set. Let $\sigma_{s,t} = (\sigma_{t-1} \cdots \sigma_{s+1})\sigma_s(\sigma_{s+1}^{-1} \cdots \sigma_{t-1}^{-1})$, where $1 \leq s < t \leq n$. We define $\sigma_{s,t} = \sigma_{t,s}$, and adopt the convention that whenever it is convenient to do so we will write the smaller subscript first. The new presentation has generators $\{\sigma_{s,t},\ 1 \leq s < t \leq n\}$ and defining relations:

$$\begin{aligned} \sigma_{s,t}\sigma_{q,r} &= \sigma_{q,r}\sigma_{s,t} && \text{if } \ (t-r)(t-q)(s-r)(s-q) > 0, \\ \sigma_{s,t}\sigma_{r,s} &= \sigma_{r,t}\sigma_{s,t} = \sigma_{r,s}\sigma_{r,t} && \text{if } \ 1 \leq r < s < t \leq n. \end{aligned} \tag{3}$$

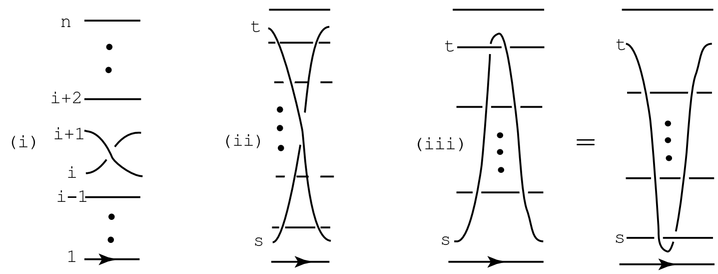


Figure 2: (i) The elementary braid σ_i. (ii) The elementary braid $\sigma_{s,t}$. (iii) The pure braid $A_{s,t} = A_{s,t} = \sigma^2_{s,t}$.

See sketch (ii) of Figure 2 for a picture of $\sigma_{s,t}$, and [23] for a proof that (2) and (3) define the same group. Both presentations will be needed in our work. Note that the new generators include the old ones as a proper subset, since $\sigma_i = \sigma_{i,i+1}$ for each $i = 1, 2, \ldots, n-1$.

By (1) the pure braid group $\mathbf{P}_n$ has index $n!$ in $\mathbf{B}_n$. Let $A_{s,t} = A_{t,s} = \sigma^2_{s,t}$. (See sketch (iii) of Figure 2). The symmetry $A_{s,t} = A_{t,s}$ can be seen by tightening the s^{th} strand at the expense of loosening the t^{th} strand. It is proved in [6] and also in [65] that $\mathbf{P}_n$ has a presentation with generators $A_{r,s}$, $1 \leq r < s \leq n$ and defining relations:

$$
\begin{aligned}
A^{-1}_{r,s} A_{i,j} A_{r,s} &= A_{i,j} \ \text{ if } \ 1 \leq r < s < i < j \leq n \ \text{ or } \ 1 \leq i < r < s < j \leq n \\
&= A_{r,j} A_{i,j} A^{-1}_{r,j} \ \text{ if } \ 1 \leq r < s = i < j \leq n, \\
&= (A_{i,j} A_{s,j}) A_{i,j} (A_{i,j} A_{s,j})^{-1} \ \text{ if } \ 1 \leq r = i < s < j \leq n, \\
&= (A_{r,j}, A_{s,j} A^{-1}_{r,j} A^{-1}_{s,j}) A_{i,j} (A_{r,j}, A_{s,j} A^{-1}_{r,j} A^{-1}_{s,j})^{-1} \ \text{ if } \ 1 \leq r < i < s < j \leq n.
\end{aligned} \tag{4}
$$

The relations in (4) come from the existence of a split short exact sequence, for every $k = 2, \ldots, n$:

$$\{1\} \rightarrow \mathbf{F}_{n-1} \rightarrow \mathbf{P}_n \xrightarrow{\pi^\star_n} \mathbf{P}_{n-1} \rightarrow \{1\}. \tag{5}$$

The map $\pi^\star_n$ is defined by pulling out the last braid strand, and the image of $\mathbf{P}_{n-1}$ under its inverse embeds $\mathbf{P}_{n-1}$ in $\mathbf{P}_n$, as the subgroup generated by pure braids on the first $n-1$ strands. The free subgroup F_{n-1} is generated by the braids $A_{1,n}, A_{2,n}, \ldots, A_{n-1,n}$. The pure braid group $\mathbf{P}_2$ is infinite cyclic and generated by $A_{1,2}$. Inducting on n, the structure of $\mathbf{P}_n$ via a sequence of semi-direct products is uncovered.

1.3 $\mathbf{B}_n$ and $\mathbf{P}_n$ as mapping class groups

Our announced goal in this review was to concentrate on areas where there have been new developments in recent years. While it has been known for a very long time that Artin's braid group is naturally isomorphic to the mapping class group of an n-times punctured disc, people have asked us many times for a simple proof of this fact. We do not know of a simple one in the literature, therefore we present one here. In this case 'simple' does not mean intuitive and based upon first principles, rather it means using machinery which is normally available to a graduate student who has the tools learned in a first year graduate course in topology, and is preparing to begin research.

Let $S = S_{g,b,n}$ denote a 2-manifold of genus g with b boundary components and n punctures, and let $\mathrm{Diff}^+(S)$ denote the groups of all orientation preserving diffeomorphisms of S. Observe that we may assign the compact open topology to $\mathrm{Diff}^+(S)$, making it into a topological group. The mapping class group $\mathcal{M} = \mathcal{M}_{g,b,n}$ of S is $\pi_0(\mathrm{Diff}^+(S))$, that is, the quotient of $\mathrm{Diff}^+(S)$ modulo its subgroup of all diffeomeorphisms of S which are isotopic to the identity rel ∂S. We allow diffeomeorphisms in $\mathrm{Diff}^+(S)$ to permute the punctures, writing $\mathrm{Diff}^+(S_{g,b,\hat{n}})$ if they are to be fixed pointwise. Our interest in this article is mainly in the special case of $\mathcal{M}_{0,1,n}$.

Theorem 1 *There are natural isomorphisms:*

$$\mathbf{B}_n \cong \mathcal{M}_{0,1,n} \quad \text{and} \quad \mathbf{P}_n \cong \mathcal{M}_{0,1,\hat{n}}$$

Proof: We begin with an intuitive description of how to pass from diffeomorphisms to geometric braids and back again. Choose any $h \in \mathrm{Diff}^+(S_{0,1,n})$. While h is in general not isotopic to the identity, its image $i(h)$ in $\mathrm{Diff}^+(S_{0,1,0})$ under the inclusion map $i : \mathrm{Diff}^+(S_{0,1,n}) \to \mathrm{Diff}^+(S_{0,1,0})$ is, because $\mathrm{Diff}^+(S_{0,1,0}) = \{1\}$. Let h_t denote the isotopy. If the punctures in $S_{0,1,n}$ are at $(p_1, p_2, \ldots, p_n)$, then the n paths $(h_t(p_1), h_t(p_2), \ldots, h_t(p_n))$ defined by the traces of the points $(p_1, p_2, \ldots, p_n)$ under the isotopy sweep out a braid in $S_{0,1,0} \times [0,1]$, and the equivalence class of this braid is the image of the mapping class $[h]$ in the braid group $\mathbf{B}_n$.

It's a little bit harder to understand the inverse isomorphism, from the braid group to the mapping class group. One chooses a geometric braid and imagines it as being located in a slice of 3-space, with the bottom endpoints of the n braid strands (which are oriented top to bottom) as being pinned to the distinguished points $p_1, \ldots, p_n$ on the punctured disc. If one is very careful the n braid strings can be laid down on the punctured disc so that they become n non-intersecting simple arcs, each of which begins and ends at a base point. One then constructs a homeomorphism of the punctured disc to itself in such a way that the trace of the isotopy to the identity is the given set of n non-intersecting simple arcs.

To prove the theorem, we begin by establishing the isomorphism between $\mathbf{P}_n$ and $\mathcal{M}_{0,1,\hat{n}}$. A good general reference for the underlying mathematics is Chapter 6 of the textbook [38]. As previously noted $\mathrm{Diff}^+(S_{0,1,0})$ is a topological group. Also, $\mathrm{Diff}^+(S_{0,1,\hat{n}})$ is a closed subgroup of $\mathrm{Diff}^+(S_{0,1,0})$. The evaluation map $\mathcal{E} : \mathrm{Diff}^+(S_{0,1,0}) \to \mathcal{C}_{0,\hat{n}}$ is defined by $\mathcal{E}(h) = (h(p_1), \ldots, h(p_n))$. It is clear that $\mathcal{E}$ is continuous with respect to the compact open topology on $\mathrm{Diff}^+(S_{0,1,\hat{n}})$ and the subspace topology for $\mathcal{C}_{0,\hat{n}} \subset \mathbb{C} \times \mathbb{C} \times \cdots \times \mathbb{C}$. The topological group $\mathrm{Diff}^+(S_{0,1,0})$ acts n-transitively on the disc in the sense that if $(p_1, \ldots, p_n)$ are n distinct points and $(w_1, \ldots, w_n)$ are n others then there is an $h \in \mathrm{Diff}^+(S_{0,1,0})$ such that $h(p_i) = w_i,\ i = 1, \ldots, n$. Observe that if $h \in \mathrm{Diff}^+(S_{0,1,\hat{n}})$ then $(h(p_1), \ldots, h(p_n)) = (p_1, \ldots, p_n)$ and if $h, h' \in \mathrm{Diff}^+(S_{0,1,0})$ with $\mathcal{E}(h) = \mathcal{E}(h')$, then h, h' are in the same left coset of $\mathrm{Diff}^+(S_{0,1,\hat{n}})$ in $\mathrm{Diff}^+(S_{0,1,0})$. In this situation it is shown in [127], part I, Sections 7.3 and 7.4, that the 3-tuple $(\mathcal{E},\ \mathrm{Diff}^+(S_{0,1,0}),\ \mathcal{C}_{0,\hat{n}})$ is a fiber space, with total space $\mathrm{Diff}^+(S_{0,1,0})$, base space $\mathcal{C}_{0,\hat{n}}$, projection $\mathcal{E}$ and fiber $\mathrm{Diff}^+(S_{0,1,\hat{n}})$. (It is a good exercise for a graduate student to prove this directly by constructing the required local product structure in an explicit manner.) The long exact sequence of homotopy groups of a fibration then gives the following exact sequence of groups and homomorphisms, where we focus on the range that is of interest:

$$\ldots \to \pi_1(\mathrm{Diff}^+(S_{0,1,0})) \stackrel{\mathcal{E}_*}{\to} \pi_1(\mathcal{C}_{0,\hat{n}}) \stackrel{\partial_*}{\to} \pi_0(\mathrm{Diff}^+(S_{0,1,\hat{n}})) \stackrel{i_*}{\to} \pi_0(\mathrm{Diff}^+(S_{0,1,0})) \stackrel{\mathcal{E}_*}{\to} \ldots$$

The two end groups are trivial. The left middle group is $\mathbf{P}_n$ and the right middle group is $\mathcal{M}_{0,1,\hat{n}}$. The isomorphism of Theorem 1 is ∂_*. Tracing through the mathematics one finds that in fact its inverse is the map that we described right after we stated the theorem. The assertion about $\mathbf{P}_n$ is therefore true. The proof for $\mathbf{B}_n$ can then be completed by comparing the short exact sequence (1), which says that $\mathbf{B}_n$ is a finite extension of $\mathbf{P}_n$ with quotient the symmetric group with a related short exact sequence for the mapping class groups:

$$1 \to \mathcal{M}_{0,1,\hat{n}} \stackrel{j_*}{\longrightarrow} \mathcal{M}_{0,1,n} \to \Sigma_n \longrightarrow 1 \tag{6}$$

Thus $\mathcal{M}_{0,1,n}$ is a finite extension of $\mathcal{M}_{0,1,\hat{n}}$ with quotient the symmetric group Σ_n. Comparing corresponding groups in the short exact sequences (1) and (6), we see that the first two and last two are isomorphic. The 5-Lemma then shows that the middle ones are too. This completes the proof of Theorem 1.$\|$

1.4 Some examples where braiding appears in mathematics, unexpectedly

We discuss, briefly, a variety of examples, outside of knot theory, where 'braiding' is an essential aspect of a mathematical or physical problem.

1.4.1 Algebraic geometry

Configuration spaces and the braid group appear in a natural way in algebraic geometry. Consider the complex polynomial

$$(X - z_1)(X - z_2)\ldots(X - z_n) = X^n + a_1X^{n-1} + \ldots + a_{n-1}X + a_n$$

of degree n with n distinct complex roots $z_1, \ldots, z_n$. The coefficients $a_1, \ldots, a_n$ are the elementary symmetric polynomials in $\{z_1, \ldots, z_n\}$, and so we get a continuous map $\mathbb{C}^n \to \mathbb{C}^n$ which takes roots to coefficients. Two points have the same image if and only if they differ by a permutation, so we get the same identification as in the quotient map $\tau : \mathcal{C}_{0,\hat{n}} \to \mathcal{C}_{0,n}$, in quite a different way. Since we are requiring that our polynomial have n distinct roots, a point $\{a_1, \ldots, a_n\}$ is in the image of $\vec{z}$ under the root-to-coefficient map if and only if the polynomial $X^n + a_1X^{n-1} + \ldots + a_n$ has n distinct roots, i.e. if and only if its coefficients avoid the points where the discriminant

$$\Delta = \prod_{i<j}(z_i - z_j)^2,$$

expressed as a polynomial in $\{a_1, \ldots, a_n\}$, vanishes. Thus $\mathcal{C}_{0,n}(\mathbb{C})$ can be interpreted as the complement in $\mathbb{C}^n$ of the algebraic hypersurface defined by the equation $\Delta = 0$, where Δ is rewritten as a polynomial in the coefficients $a_1, \ldots, a_n$. (For example, the polynomial $X^2 + a_1X + a_2$ has distinct roots precisely when $a_1^2 - 4a_2 = 0$). In this setting the base point

$\tau(\vec{p})$ is regarded as the choice of a complex polynomial of degree n which has n distinct roots, and an element in the braid group is a choice of a continuous deformation of that polynomial along a path on which two roots never coincide. There is a substantial literature in this area, from which we mention only one paper, by Gorin and Lin [84]. We chose it because it contains a description of the commutator subgroup $\mathbf{B}_n'$ of the braid group, and many people have asked the first author for a reference on that over the years. While there may be other references they are unknown to us.

1.4.2 Operator algebras

Our next example, taken from the work of Vaughan Jones [87],[88], is interesting because it shows how 'braiding' can appear in disguise, so that initially one misses the connection. We consider the theory, in operator algebras, of 'type II_1 factors', ordered by inclusion. Let M denote a Von Neumann algebra, i.e. an algebra of bounded operators acting on a Hilbert space h. The algebra M is called a <u>factor</u> if its center consists only of scalar multiples of the identity. The factor is type II_1 if it admits a linear functional, called a trace, $tr : M \rightarrow \mathbb{C}$, which satisfies the following three conditions: (i) $tr(xy) = tr(yx)\ \forall\, x, y \in M$, (ii) $tr(1) = 1$, and $tr(xx^\star) > 0$, where $x^\star$ is the adjoint of x. In this situation it is known that the trace is unique, in the sense that it is the only linear functional satisfying the first two conditions. An old discovery of Murray and Von Neumann was that factors of type II_1 provide a type of 'scale' by which one can measure the dimension of h. The notion of dimension which occurs here generalizes the familiar notion of integer-valued dimensions, because for appropriate M and h it can be any non-negative real number or ∞.

The starting point of Jones' work was the following question: if M_1 is a type II_1 factor and if $M_0 \subset M_1$ is a subfactor, is there any restriction on the real numbers which occur as the ratio $\lambda = \dim_{M_0}(h)/\dim_{M_1}(h)$? The question has the flavor of questions one studies in Galois theory. On the face of it, there was no reason to think that λ could not take on any value in $[1, \infty]$, so Jones' answer came as a complete surprise. He called λ the <u>index</u> $|M_1 : M_0|$ of M_0 in M_1, and proved a type of rigidity theorem about type II_1 factors and their subfactors:

The Jones Index Theorem: $\lambda \subset [4, \infty] \cup \{4cos2\pi/p\}$, where $p \in \mathbb{Z}, p \geq 3$. Moreover, each real number in the continuous part of the spectrum $[4, \infty]$ and in the discrete part $\{\{4cos2\pi/p\}, p \in \mathbb{Z}, p \geq 3\}$ is realized.

What does all this have to do with braids? To answer the question, we sketch the idea of the proof, which is to be found in [87]. Jones begins with the type II_1 factor M_1 and the subfactor M_0. There is also a tiny bit of additional structure: It turns out that in this setting there exists a map $e_1 : M_1 \rightarrow M_0$, known as the <u>conditional expectation</u> of M_1 on M_0. The map e_1 is a projection, i.e. $e_1^2 = e_1$. His first step is to prove that the ratio λ is independent of the choice of the Hilbert space h. This allows him to choose an appropriate h so that the algebra M_2 generated by M_1 and e_1 makes sense. He then investigates M_2 and proves that it is another type II_1 factor, which contains M_1 as a subfactor, moreover $|M_2 : M_1| = |M_1 : M_0| = \lambda$. Having in hand another II_1 factor M_2 and its subfactor M_1, there is also a trace on M_2 (which by the uniqueness of the trace) coincides with the trace on M_1 when it is restricted to M_1, and another conditional expectation $e_2 : M_2 \rightarrow M_1$.

This allows Jones to iterate the construction, to build algebras $M_1, M_2, \ldots$ and from them a family of algebras $\{J_n,\ n = 1, 2, 3, \ldots\}$, where J_n is generated by $1, e_1, \ldots, e_{n-1}$.

Rewriting history a little bit in order to make the subsequent connection with braids a little more transparent, we now replace the projections e_k, which are not units, by a new set of generators which are units, defining: $g_k = te_k - (1-e_k)$, where $(1-t)(1-t^{-1}) = 1/\lambda$. The g_k's generate J_n because the e_k's do, and we can solve for the e_k's in terms of the g_k's. So $J_n = J_n(t)$ is generated by $1, g_1, ..., g_{n-1}$ and we have a tower of algebras, $J_1(t) \subset J_2(t) \subset \ldots$, ordered by inclusion. The parameter t, which replaces the index λ, is the quantity now under investigation. It's woven into the construction of the tower. The algebra $J_n(t)$ has defining relations:

$$g_i g_k = g_k g_i \ \text{ if } \ |i-k| \geq 2, \quad g_i g_{i+1} g_i = g_{i+1} g_i g_{i+1}, \quad g_i^2 = (t-1)g_i + t, \tag{7}$$
$$1 + g_i + g_{i+1} + g_i g_{i+1} + g_{i+1} g_i + + g_i g_{i+1} g_i = 0.$$

Of course there are braids lurking in the background. If we rename the $g_i's$, replacing g_i by σ_i, and declare the $\sigma_i's$ to be generators of a group, then the first two relations are defining relations in the group algebra $\mathbb{C}\mathbf{B}_n$. The algebra $J_n(t)$ is thus a homomorphic image of the group algebra of the braid group. Loosely speaking, braids are encountered in Operator Algebras because they encode the way in which each type II_1 factor M_i acts on its subfactor M_{i-1}. Braiding is thus involved in defining the associated extensions. We shall see later, in §4.7 that a similar action, via group extensions, can be used to define representations of $\mathbf{B}_n$.

To see the connection with knots and links, recall that since M_n is type II_1 it supports a unique trace, and since J_n is a subalgebra it does too, by restriction. This trace is known as a Markov trace, i.e. it satisfies the important property:

$$tr(wg_n) = f(t)tr(w) \quad \text{if} \quad w \in J_n, \tag{8}$$

where $f(t)$ is a fixed function of t. Thus, for each fixed value of f the trace is multiplied by a fixed scalar when one passes from one stage of the tower to the next, if one does so by multiplying an arbitrary element of J_n by the new generator g_n of J_{n+1}. The Jones trace is nothing more or less than the 1-variable Jones polynomial [89] associated to the knot or link which is obtained from the closed braid. We will have more to say about all this in §4.3.

1.4.3 Homotopy groups of spheres

As before, let $\mathbf{P}_{n+1}$ denote the pure braid group on $n+1$ strands. For each $i = 1, \ldots, n+1$ there is a natural homomorphism $p_i : \mathbf{P}_{n+1} \to \mathbf{P}_n$, defined by pulling out the i^{th} strand. The group of Brunnian braids is

$$BR_{n+1} = \cap_{i=1}^{i=n+1} \text{kernel}(p_i),$$

i.e. a braid is in BR_{n+1} if and only if, on pulling out any strand, it becomes the identity braid on n strands. Brunnian braids have received some attention in knot theory.

Braids have played a role in homotopy theory for many years, most particularly in the work of F. Cohen and his students (see for example [10]), but during the past few years the connection was sharpened when it was discovered that there is an embedding of a free group F_n in P_{n+1} with the property that a well-defined quotient of $BR_{n+1} \cap F_n$ (a little bit too complicated to describe here) is isomorphic to $\pi_{n+1}S^2$. It remains to be seen whether new knowledge about the unidentified higher homotopy groups of spheres can be obtained through the methods of [10].

1.4.4 Robotics

Our fourth example is an application of configuration spaces to robotics. It shows the braid group popping up in an unexpected way (until you realize how natural it is). Robots, or AGVs (automatic guided vehicles), are required to travel across a factory floor that contains many obstacles, en route to a goal position (e.g. a loading dock or an assembly workstation). The problem is to design a control system which insures that the AGVs not collide with the obstacles, or with each other, and complete the task with efficiency with regard to various work functionals. Here is how configuration spaces appear: The underlying space in this simple example is the workspace floor X, from which a finite set $\mathcal{O}$ of obstacles are to be removed. The configuration space of n non-colliding AGVs is then precisely $\mathcal{C}_{0,n}(X - \mathcal{O})$. More generally, $X - \mathcal{O}$ is replaced by a finite graph Y , and the the braid group $\mathbf{B}_n$ by the braid group $\pi_1(\mathcal{C}_{0,n}(Y))$ of the graph. There is a vast literature on this subject; we suggest [78] by R. Ghrist, as a starter.

1.4.5 Public key cryptography

In this example braids are important for rather different reasons than they were in our earlier examples. In our earlier examples the underlying phenomenon which was being investigated involved actual braiding, albeit sometimes in a concealed way. In the example that we now describe particular properties of the braid groups $\mathbf{B}_n,\ n = 1, 2, 3, \ldots$, rather than the actual interweaving of braid strands, are used in a clever way to construct a new method for encrypting data.

The problem which is the focus of 'public key cryptography' will be familiar to everyone: the security of our online communications, for example our credit card purchases, our ATM transactions, our cell phone conversations and a host of other transactions that have become a part of everyday life in the 21^{st} century, The basic problem is to encrypte or translate a secret message into a code that can be sent safely over a public system such as the internet, and decoded at the receiving end by the use of a secret piece of information known only to the sender and the recipient, the 'key'. The problem that must then be solved is to establish a private key that will be known only to the sender and the recipient, who will then be able to exchange information over an insecure channel. In recent years much work has been done on certain codes which are based upon the assumption that the word problem has polynomial growth as braid index n is increased, whereas the conjugacy problem does not. But in §5 we will review recent work on the word and conjugacy problems in the braid groups, and show that such an assumption seems problematic at best. See §5, and in particular the discussion in §5.5 below.

Acknowledgements: We thank Tahl Nowik, who suggested, during a course that the first author gave on mapping class groups, that techniques she had used for other purposes could be adapted to give the proof we presented here of Theorem 1. We also thank Robert Bell, David Bessis, Nathan Broaddus, John Cannon, Ruth Charney, Fred Cohen, Patrick Dehornoy, Roger Fenn, Daan Krammer, Lee Mosher, Luis Paris, Richard Stanley, Morwen Thistlethwaite and Bert Wiest for their help in chasing down facts and references, and helping us fill gaps in our knowledge. We would particularly like to thank all the students who attended Math 661 in the Fall of 2003 at Cornell University, who were gracious guinea pigs for large parts of this article, especially Heather Armstrong, whose careful attention and diligence significantly improved the manuscript, and Bryant Adams, who suggested the proof of Lemma 2.2 to us.

2 From knots to braids

In this chapter we will explore, for the benefit of readers who are new to the subject, the foundations of the close relationship between knots and braids. We will first describe the straightforward process of obtaining a knot or a link from a given braid by 'closing' the braid. This leads us directly to formulate two fundamental questions about knots and braids. First, is it always possible to transform a given knot into a closed braid? This question will be answered in the affirmative in Theorem 2, first proved by Alexander in 1928 in [2]. The correspondence between knots and braids is clearly not one-to-one (for example, conjugate braids yield equivalent knots), leading naturally to the second question: which closed braids represent the same knot type? That question is addressed in Theorem 4, first formulated by A. Markov in [104], which gives 'moves' relating any two closed braid representatives of a knot or link, while simultaneously preserving the closed braid structure.

Together, Theorems 2 and 4 form the cornerstone of any study of knots via closed braids, so we feel obliged to prove them. Among the many proofs that have been published of both over the years, we have chosen ones that we like but which do not seem to have appeared in any of the review articles that we know. The proof that we give of Alexander's Theorem is due to Shuji Yamada [136], with subsequent improvements by Pierre Vogel [135]. The algorithm is elementary enough to be accessible to a beginner, and has the advantage for experts of being suitable for programming. The proof that we present of Markov's theorem is due to Pawel Traczyk [129]. It is relatively brief, as it assumes Reidemeister's well-known theorem about the equivalence relation on any two diagrams of a knot, Theorem 3 below, building on methods introduced in the proof of Theorem 2.

2.1 Closed braids

For simplicity, let us begin with a planar diagram of a given geometric braid. To obtain a knot or link, one simply 'closes up' the ends of the braid as in Figure 3. The pre-image

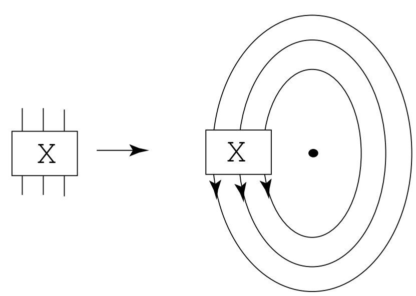


Figure 3: The operation of closing a braid X to form a closed braid

in $\mathbb{R}^3$ of the 'center point' shown in Figure 3 under the usual projection map is called the <u>axis</u> of the braid. (If one wishes to consider the knot in S^3, then we include the point at infinity so that the braid axis is an embedded S^1.) We then orient the resulting knot or link in such a way that the strands of the braid are all travelling counterclockwise about the braid axis. The knot or link type resulting from performing this operation on a braid X is known as the <u>closure</u> of X and will be denoted by $b(X)$. The same notation may also refer to the particular diagram as in Figure 3.

Equivalently, consider a knot $K \subset S^3$. Suppose there exists $A = h(S^1)$ where h is an embedding and Z is unknotted in S^3 and contained in the complement of K. Suppose further that we choose the point at infinity $\{\infty\}$ to be in A and, using standard cylindrical coordinates (ρ, θ, z) on $\mathbb{R}^3$, identify the resulting copy of $\mathbb{R} \cong A - \{\infty\}$ with the z-axis in $\mathbb{R}^3 \cong S^3 - \{\infty\}$. If we always have $d\theta/dt > 0$ as we travel about the knot K with an appropriate cylindrical parametrization, then we say that K is a closed braid with respect to the axis A. The closed braid diagram of Figure 3 is then obtained by projection parallel to the direction defined by A onto a plane that is orthogonal to A.

2.2 Alexander's Theorem

As we just observed, it is a simple matter to obtain a knot or link from a braid. The classical theorem of J. Alexander allows us to reverse this process, though not in a unique way:

Theorem 2 (Alexander's Theorem [2]) *Every knot or link in S^3 can be represented as a closed braid.*

Proof: Alexander's original proof was algorithmic, i.e. it gave an algorithm for transforming a knot or link into closed braid form. While it is straightforward, we do not know of any computer program based upon it. We shall give instead a rather different and newer algorithm originally due to Yamada [136], as later improved by Vogel [135]. We like it for two reasons: (1) It has a beautiful corollary (see Corollary 2.1 below) which reveals structure about knot diagrams that had not even been conjectured by any of the experts before 1987, even though there was abundant evidence of its truth; (2) It leads, very easily, to an efficient computer program for putting knots into braid form. In this regard we note that when Jones was writing the manuscript [88], which resulted in his award of the Fields medal, he computed closed braid representatives for the 249 knots of crossing number less than or equal to 10, constructing the first table known to us of closed braid representatives of knots. His list remains extremely useful to the workers in the area in 2004. Yamada's work was not yet known when he did that work, and there did not seem to him to be a good way to program the Alexander method for a computer, so he calculated them one at a time by hand. The amount of work that was involved can only be appreciated by the reader who is willing to try a few examples.

In order to prove Alexander's Theorem, we shall first present the Yamada-Vogel algorithm for transforming a knot into closed braid form in full, followed immediately by an illustrative example (Example 2.1). We shall then prove Alexander's Theorem by showing that it is always possible to perform the steps of the Yamada-Vogel algorithm for any given knot or link and that the algorithm always leads to a closed braid.

The Yamada-Vogel algorithm draws on Seifert's well-known algorithm for using a diagram of an oriented knot or link K to construct a Seifert surface for K (see [123] or [101], e.g., for a thorough treatment of Seifert's algorithm), and we will need some related terminology. Let C and C' be two oriented disjoint simple closed curves in S^2. Then C and C' cobound an annulus A. We say that C and C' are coherent (or coherently oriented) if C and C' represent the same element of $H_1(A)$. Otherwise we say that C and C' are incoherent. Following Traczyk [129], we define the height of a knot diagram D, denoted $h(D)$, to be the number of distinct pairs of incoherently oriented Seifert circles which arise from applying Seifert's algorithm to D. The height function gives us a useful characterization of a closed braid: *a diagram D represents a closed braid if and only if $h(D) = 0$.* (Recall that D lives in S^2.)

The Yamada-Vogel Algorithm

1. Let D be a diagram of an oriented knot K. Smooth all crossings of D as in Seifert's algorithm to obtain n Seifert circles $C_1, \ldots, C_n$. Record each original crossing with a signed arc: $(+)$ for a positive crossing (often called a right-handed crossing), $(-)$ for a negative (or left-handed) crossing (see Figure 2(iv)). The resulting diagram is the Seifert picture S corresponding to the diagram D. Note that any two circles joined by a signed arc in any Seifert picture are necessarily coherent. For an example that illustrates the construction of a Seifert diagram, see the passage from the bottom left to the top left sketches in Figure 4.

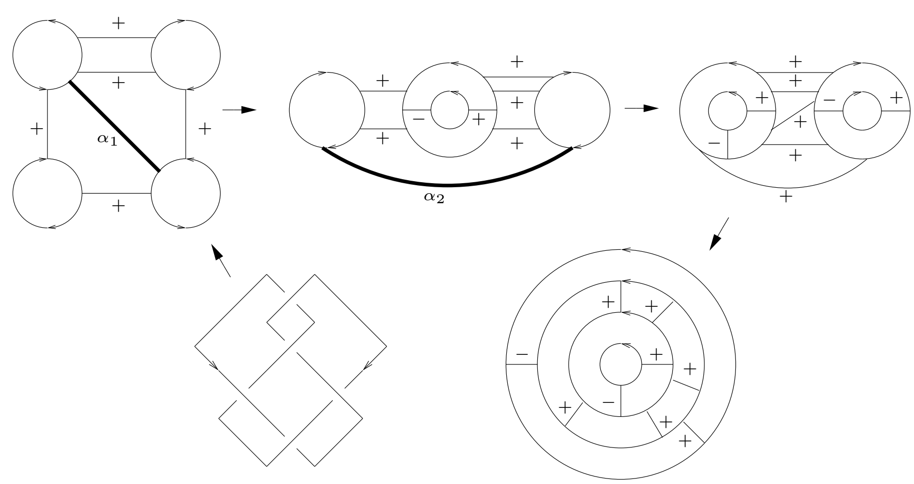


Figure 4: The Yamada-Vogel algorithm performed on the knot 5_2.

2. If $h(D) = 0$, the knot K is already in closed braid form, and we are done. If $h(D) > 0$, we can find a reducing arc α, i.e., an arc joining an incoherent pair C_i, C_j such that α intersects S only at its endpoints. Reducing arcs are illustrated as heavy black arcs in the example in Figure 4. A component of $S^2 \setminus S$ which admits a reducing arc is called a defect region. Perform a reducing move along α, as shown in Figure 5, to obtain a new Seifert picture S' in which a pair of coherent Seifert circles, C_a and C_z, joined by two oppositely signed arcs, replaces the incoherent pair C_i, C_j. The corresponding move on the original diagram D is a Reidemeister move of type II in which we slide C_i over C_j in a small neighborhood of the arc α to obtain a new diagram D' with two new crossings. Note that if we instead slide C_i under C_j, we obtain the same two new Seifert circles but the signs of the two new signed arcs are now switched.

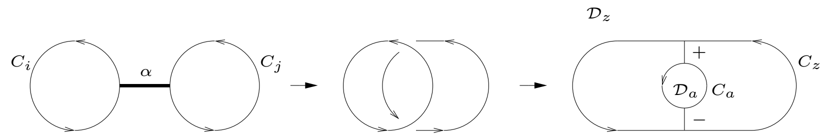


Figure 5: The local picture of a reducing move.

3. Continue performing reducing moves on incoherent pairs until a diagram with height zero is obtained.

Example 2.1 We apply the Yamada-Vogel algorithm to the diagram of the knot 5_2, pictured in Figure 4. (The reader can check that this is the first example in the knot tables of a knot diagram with height greater than zero.) The first stage shows the Seifert picture associated to the original diagram, consisting of 4 Seifert circles, with 5 signed arcs (all positive) recording the original crossings. We see that the original knot diagram has height 2. The figure shows a choice of reducing arc, α_1, joining one of the two pairs of incoherent circles.

In the third sketch of Figure 4, we see the new Seifert picture resulting from performing the reducing move along α_1. Note that we have introduced two new crossings of opposite sign. We also see a new reducing arc, α_2, joining the only remaining pair of incoherent circles. We see in the fourth sketch the Seifert picture with height zero resulting from the second reducing move performed along α_2. At this point, we are done, but in the final sketch we see a different planar projection of the same Seifert picture which allows us easily to read off a braid word associated to the knot: beginning with the positive signed arc in the 'twelve o'clock' position and reading counter-clockwise, we see that the knot 5_2 is equivalent to $b(X)$, where $X = \sigma_2\sigma_1^{-1}\sigma_2\sigma_3^{-1}\sigma_2\sigma_1\sigma_2\sigma_3\sigma_2$.

We may learn several things from this simple example. Applying the braid relations to the word defined by X, we see that $X = \sigma_2\sigma_1^{-1}\sigma_2\sigma_3^{-1}\sigma_2\sigma_1\sigma_2\sigma_3\sigma_2 = \sigma_2\sigma_1^{-1}\sigma_2\sigma_3^{-1}\sigma_1\sigma_2\sigma_1\sigma_3\sigma_2 = \sigma_2\sigma_1^{-1}\sigma_2\sigma_1\sigma_3^{-1}\sigma_2\sigma_3\sigma_1\sigma_2 \quad = \sigma_2\sigma_1^{-1}\sigma_2\sigma_1\sigma_2\sigma_3\sigma_2^{-1}\sigma_1\sigma_2$. Since this braid only involves σ_3 once, we may 'delete a trivial loop to get the 8-crossing 3-braid $\sigma_2\sigma_1^{-1}\sigma_2\sigma_1\sigma_2\sigma_2^{-1}\sigma_1\sigma_2$, so the algorithm did not give us minimum braid index. The algorithm also does not give shortest words because our 8-crossing braid may be shortened to the 6-crossing 3-braid $\sigma_2\sigma_1^{-1}\sigma_2\sigma_1^2\sigma_2$. In fact, 6 is minimal, because the crossing number of a 3-braid knot must be even, and this knot has no diagram with fewer than 5 crossings, so 6 is minimal. So we may deduce one more fact: when we use the Yamada-Vogel algorithm to change a knot which is not in closed braid form to one which is, the crossing number goes up. ♠

To prove Alexander's Theorem, we first need to show that a reducing move strictly decreases the height of a diagram. This lemma is sometimes stated as 'obvious' in the literature, but the question arises frequently enough to warrant a short but thorough argument.

Lemma 2.1 *Suppose a reducing move is performed which transforms a diagram D to a diagram D'. Then $h(D') = h(D) - 1$.*

Proof. Let $C_1, \ldots, C_n$ be the Seifert circles in the Seifert picture S corresponding to the diagram D. Let C_i, C_j be an incoherent pair. The union $C_i \cup C_j$ separates the 2-sphere into three components: an annulus A cobounded by C_i and C_j, and two disks, $\mathcal{D}_i$ and $\mathcal{D}_j$ bounded by C_i and C_j, respectively. Suppose that A admits a reducing arc α. A reducing move along α preserves circles $C_p, p \neq i, j$ and replaces C_i and C_j with two new circles, one of which necessarily bounds a disk $\mathcal{D}_a$ containing no other Seifert circles in the new Seifert picture S'. We denote this circle by C_a (see Figure 5). The other new circle, denoted C_z, bounds a disk $\mathcal{D}_z$ containing all Seifert circles originally contained in the annulus A in S.

To simplify the bookkeeping, we shall write $(C_r, C_s) = 1$ if the pair C_r, C_s is coherent, or else $(C_r, C_s) = -1$ if C_r, C_s are incoherent. Obviously, if $\{p, q\} \cap \{i, j\} = \emptyset$, then (C_p, C_q) is unchanged by the reducing move, so we need only consider the effect of the reducing move on (C_p, C_x), where $x = i$ or $x = j$ and $p \neq i, j$. Now if C_p is contained in the annulus A in S (and hence in $\mathcal{D}_z$ in S'), then clearly $(C_p, C_z) = (C_p, C_a) = (C_p, C_i) = (C_p, C_j)$. Also, if $C_p \subset \mathcal{D}_i$ in S, then $(C_p, C_z) = (C_p, C_i)$ and $(C_p, C_a) = (C_p, C_j)$. Similarly, if $C_p \subset \mathcal{D}_j$ in S, then $(C_p, C_z) = (C_p, C_j)$ and $(C_p, C_a) = (C_p, C_i)$. Therefore the number of distinct incoherent pairs C_r, C_s in S with $\{r, s\} \neq \{i, j\}$ is equal to the total number of distinct incoherent pairs in S' of the form C_r, C_s with $\{r, s\} \neq \{a, z\}$. By construction, however, we have replaced $(C_i, C_j) = -1$ with $(C_z, C_a) = 1$. Thus $h(D') = h(D) - 1$. $\|$

The previous lemma tells us that the Yamada-Vogel algorithm will always lead to a diagram of height zero, i.e., a closed braid, as long as it is always possible to perform Step 3. Therefore the following lemma, whose proof was suggested to us by Bryant Adams, will conclude the proof of Alexander's Theorem.

Lemma 2.2 [136] *Let D be a knot or link diagram. If $h(D) > 0$, then the Seifert picture S associated to D contains a defect region.*

Proof: Each component $\mathcal{R}$ of $S^2 \backslash S$ is a surface of genus 0 with $k \geq 1$ boundary components. Each boundary component of $\mathcal{R}$ is a union of some number of signed arcs (possibly zero) and subarcs of Seifert circles. We call the collection of Seifert circles in S which form part or all of a boundary component of $\mathcal{R}$ the <u>exposed circles</u> of $\mathcal{R}$.

Let us now examine the possible ways in which $\mathcal{R}$ could fail to be a defect region. Certainly, if $\mathcal{R}$ has one exposed circle, then it is not a defect region. If $\mathcal{R}$ has two exposed circles, then $\mathcal{R}$ is either a disk or an annulus. In this case, if $\mathcal{R}$ is a disk, then its two exposed circles are joined by at least one signed arc; hence the two circles are coherent and $\mathcal{R}$ is not a defect region. If $\mathcal{R}$ is an annulus, either its two exposed circles are incoherent, in which case it is a defect region, or else they are coherent and $\mathcal{R}$ is not a defect region. If $\mathcal{R}$ has three or more exposed circles, then there is necessarily one incoherent pair among them, and $\mathcal{R}$ is a defect region.

Suppose that no component of $S^2 \setminus S$ is a defect region and hence that each component is of one of the three types of non-defect regions described above. It is clear that we must have at least one region of the second type, since otherwise $h(D)$ is clearly zero. We can think of such a region as lying between two nested, coherent circles joined by at least one signed arc.

Let us now start with such a region, and try to build a diagram with no defect regions. We cannot add any circles in the annulus cobounded by the two nested circles, since this necessarily gives rise to at least one component with three or more exposed circles. In fact, our only option is to add coherent circles which nest with the original two circles (and as many signed arcs between adjacent pairs as we like). However, such a diagram has height zero. Therefore, $h(D) > 0$ implies that a defect region exists. This completes the proof of Lemma 2.1, and so also of Theorem 2. $\|$

The braid index of a knot or link K is the minimum number n such that there exists a braid $X \in \mathbf{B}_n$ whose closure $b(X)$ represents K. (We note that it is also common to refer to the index of a braid or a closed braid, meaning simply the number of its strands or the number of times it travels around its axis, respectively.) It is clear that the minimum number of Seifert circles in any diagram of a knot or link K is bounded above by the braid index of K. It is equally clear from the Yamada-Vogel algorithm that the reverse inequality holds. Thus we obtain the following corollary, which is due to Yamada [136]. It seems remarkable that it was not noticed long before 1987.

Corollary 2.1 [136] *The minimum number of Seifert circles in any diagram of a knot or link K is equal to the braid index of K.*

It also follows that we have a measure of the complexity of the process of transforming a knot into closed braid form as follows.

Corollary 2.2 ([129], [135]) *Let N denote the length of any sequence of reducing moves required to transform a diagram D into closed braid form. Then we have:*

$$N = h(D) \leq \frac{(n-1)(n-2)}{2}. \tag{9}$$

where n is the number of Seifert circles associated to D.

Open Problem 1 It is an open problem to determine, among all regular diagrams for a given knot or link, the minimum number of Seifert circles that are needed. By Corollary 2.1 this is the same as the minimum braid index, among all closed braid representatives of a given knot or link. We know of only one general result relating to this problem, namely the Morton-Franks-Williams inequality of [111] and [74]. It will be discussed briefly in §4.3. The literature also contains an assorted collection of ad-hoc techniques for determining the braid index of individual knots. For example, see the methods used in [31] to prove that the 6-braid template in Figure 20 below actually has braid index 6, which rests on the fact that non-trivial braid-preserving flypes always have braid index at least 3. ♣

2.3 Markov's Theorem

To introduce the main goal of this section, we begin by recalling for the reader Reidemeister's theorem, which dates from the earliest days of knot theory. It was assumed and used (as a folk theorem) long before anybody wrote down a formal statement and proof.

Theorem 3 (Reidemeister's Theorem) *Let D, D' be any two (in general not closed braid) diagrams of the same knot or link K. Then there exists a sequence of diagrams $D = D_1 \to D_2 \to \cdots \to D_k = D'$ such that any D_{i+1} in the sequence is obtained from D_i by one of the three Reidemeister moves, depicted in Figure 6.*

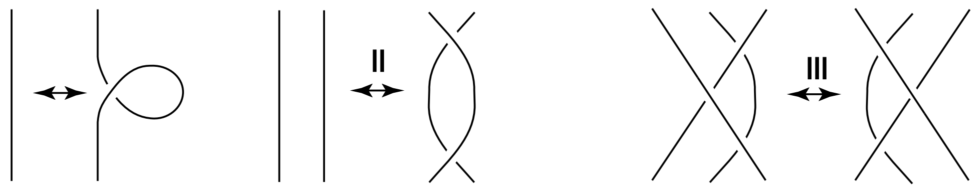


Figure 6: The 3 Reidemeister moves. The 1, 2 or 3 strands in the left sketch of each have arbitrary orientations, also we give only one of the possible choices for the signs of the crossings, for each move.

Proof: We refer the reader to [43] for a complete proof.

Alexander's Theorem, proved in the last section, guarantees us that closed braid representatives of a knot exist, but as previously noted, they are certainly not unique. Markov's Theorem, first stated in [104] with a sketch of a proof, gives us a certain amount of control over different closed braid representatives of the same knot. It asserts that any two are related by a finite sequence of elementary moves and serves as the analogue for closed braids of the Reidemeister Theorem for knots.

One of the moves of the Markov Theorem is braid isotopy. From the point of view of a topologist, braid isotopy means isotopy of the closed braid, through braids, in the complement of the braid axis. Morton has proved that if two braids have closures that are braid isotopic, then they are conjugate in $\mathbf{B}_n$ [108]. The other two moves that we need are mutually inverse, and are illustrated in Figure 7 as a move on certain $(w+2)$-braids. We call them destabilization and stabilization, where the former decreases braid index by one and the latter increases it by one. The weight w that is attached to one of the braid strands in Figure 7 denotes that many 'parallel' strands, where parallel means in the framing defined by the given projection. The braid inside the box which is labelled P is an arbitrary $(w+1)$-braid. Later, it will be necessary to distinguish between positive and negative destabilizations, so we illustrate both now.

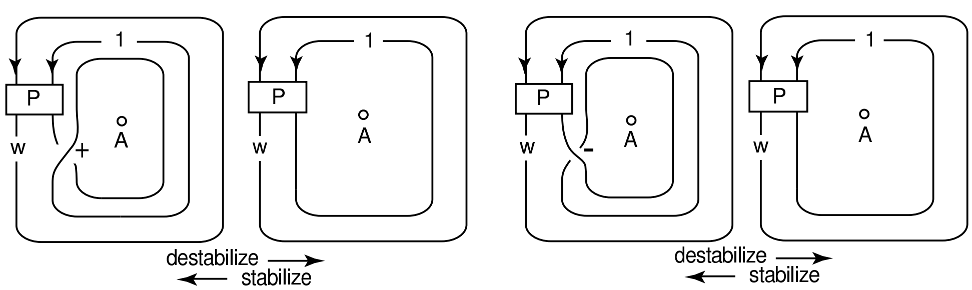


Figure 7: The destabilization and stabilization moves.

Theorem 4 (Markov's Theorem) *Let X, X' be closed braid representatives of the same oriented link type K in oriented 3-space. Then there exists a sequence of closed braid representatives of K:*

$$X = X_1 \rightarrow X_2 \rightarrow \cdots \rightarrow X_r = X'$$

taking such that each X_{i+1} is obtained from X_i by either (i) *braid isotopy, or* (ii) *a single stabilization or destabilization.*

We call the moves of Theorem 4 Markov moves, and say that closed braids that are related by a sequence of Markov moves are Markov-equivalent.

Forty years after Markov's theorem was announced, the first detailed proof was published in [18]. At least 5 essentially different proofs exist today. See for example [110], in which Morton gives his beautiful threading construction for knots and braids which also yields an alternate proof of Alexander's Theorem. Here we shall present a proof due to Pawel Traczyk [129]. It begins with Reidemeister's theorem, and uses the circle of ideas that were described in the previous section, and so it is particularly appropriate for us.

Proof: We are given closed braids X, X' which represent the same oriented knot type K. Without loss of generality we may assume that X and X' are defined by closed braid diagrams Y, Y' of height $h(Y) = h(Y') = 0$. By Theorem 3 we know there is a sequence of knot diagrams $Y = Y_1 \to Y_2 \to \cdots \to Y_k = Y'$, where in general $h(Y_i) \geq 0$ for $i = 2, \ldots, k-1$, such that any two diagrams in the sequence are related by a single Reidemeister move of type I, II or III. The first step in Traczyk's proof is to reduce the proof to sequences of knot diagrams which are related by Yamada-Vogel reducing moves:

Lemma 2.3 *It suffices to prove Theorem 4 for closed braid diagrams Y, Y' which are related by sequences $Y = Y_1 \to Y_2 \to \cdots \to Y_q = Y'$ with the properties* (i) $h(Y) = h(Y') = 0$, (ii) $h(Y_i) > 0$ *for* $i = 2, \ldots, q-1$, *and* (iii) Y_{i+1} *is obtained from* Y_i *by a single Yamada-Vogel reducing move or the inverse of a reducing move.*

Proof: We may always assume that the diagrams $Y_2, \ldots, Y_{q-1}$ have height > 0, for if not we simply replace the given sequences by the subsequences joining any two intermediate diagrams of height zero.

We say that a Reidemeister move is braid-like if the strands that are involved in it are locally oriented in a coherent fashion, as they would be if the diagram is a closed braid. In particular, any Reidemeister move of type I is braid-like. To begin the proof of the lemma, we establish a somewhat weaker result: we claim that we can get from Y to Y' via a finite sequence of the following four types of moves and their inverses:

- a braid-like Reidemeister move of type I, denoted type I^b,
- a braid-like Reidemeister move of type II, denoted type II^b,
- a braid-like Reidemeister move of type III, denoted type III^b,
- a Yamada-Vogel reducing move, denoted type $\mathcal{Y}$.

To prove the claim, it is enough to show that non-braid-like Reidemeister moves, which we denote by the symbols I^{nb}, II^{nb} and III^{nb}, can be achieved via a finite sequence of moves of type I^b, II^b, III^b and $\mathcal{Y}$. To prove this, we examine the cases I^{nb}, III^{nb} and II^{nb} in that order:

1. As previously noted, any type I Reidemeister move is of type I^b.

2. A type III^{nb} Reidemeister move involves three arcs of the knot or braid. There are many different cases, depending on the local orientations and the signs of the 3 crossings, but they are all similar. One of the possible cases is given by the first and last

sketches of Figure 8, where an arc passing under a crossing formed by the other two arcs is locally oriented opposite to the other two strands. The replacement sequence that is given in Figure 8 shows that our type III^{nb} Reidemeister move can be achieved by a sequence consisting of a type II^{nb} move, an isotopy, a type III^{b} move and finally another type II^{nb} move. We leave the other type III^{nb} cases to the reader, and we have reduced to the case of moves of type II^{nb}.

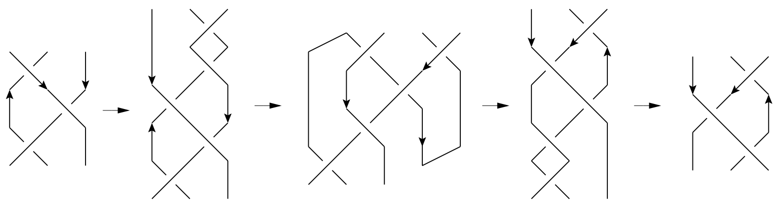


Figure 8: Replacing moves of type III^{nb}

3. A move of type II^{nb} may be regarded as a move of type $\mathcal{Y}^{\pm}$ if the arcs that are involved belong to distinct Seifert circles, so we only need to handle the case where they are subarcs of the same Seifert circle. This is done in Figure 9, where it is shown that the move can be replaced by two moves of type I^{b} (which create two new Seifert circles) followed by a move of type $\mathcal{Y}$ and another of type $\mathcal{Y}^{-1}$. This proves the claim.

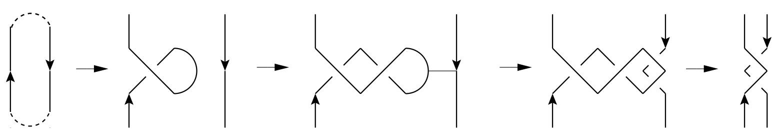


Figure 9: Replacing moves of type II^{nb}

We are thus reduced to the case in which each diagram in the sequence taking Y to Y' is either type I^{b}, II^{b}, III^{b} or $\mathcal{Y}^{\pm}$. To complete the proof of Lemma 2.3, let t be a braid-like Reidemeister move to be performed on diagram Y_i. Suppose that $h_i = h(Y_i) > 0$. Then we can find a sequence of reducing moves $r_1, \dots, r_{h_1}$ such that $r_{h_i} \circ \cdots \circ r_1(Y_i)$ is a braid and such that the associated reducing arcs $\alpha_1, \dots, \alpha_{h_i}$ are each disjoint from the region in which t is to be performed. Thus each reducing move r_j commutes with t, and we can replace t with its 'conjugate' $r_1^{-1} \circ \cdots \circ r_{h_i}^{-1} \circ t \circ r_{h_i} \circ \cdots \circ r_1$, so that t is now performed at height 0, i.e., on a braid.

If t is of type II^{b} or III^{b}, then we are done, since a braid-like move of type II or type III performed on a braid is a braid isotopy. If t is of type I^{b}, it is a stabilization (up to isotopy) only if it is performed on the braid strand 'nearest' to the braid axis. However, it is not hard to see how to realize a type I^{b} move on an arbitrary strand in terms of Markov moves: simply push the strand under the others via type II^{b} moves, and perform the required stabilization in a neighborhood of the braid axis. Note that to pass the resulting 'kink' back under a neighboring strand in a braid requires first a type III^{b} move followed by a type $\mathcal{Y}^{-1}$ move (the 'kink' is always its own Seifert circles, so two distinct circles are necessarily involved).

Thus we can return the strand with the 'kink' in it back to its original position by repeated applications of this two-step process. Since we have just seen that any type II^b or type III^b move can be realized by a finite sequence of type $\mathcal{Y}^{\pm}$ moves and braid isotopies, this means that a type I^b move can also be replaced by a finite sequence of type $\mathcal{Y}^{\pm}$ moves and braid isotopies. We can handle inverse moves of type I^b in a similar fashion

We have thus replaced our original sequence relating Y to Y' by a new one which is in general much longer, but which consists entirely of Markov moves (performed, by definition, on braid diagrams, i.e., on diagrams of height zero) and moves of type $\mathcal{Y}^{\pm}$. To be precise, our original sequence from Y to Y' may be replaced by a sequence of the form $Y = Y_0, \ldots, Y_{a_1}, \ldots, Y_{a_2}, \ldots, Y_{a_n} = Y'$, where $h(Y_{a_i}) = 0$ for all i and in each subsequence $Y_{a_i}, \ldots, Y_{a_{i+1}}$, either (1) each diagram in the subsequence has height zero and adjacent diagrams are related by a single Markov move, or (2) all the intermediate diagrams have strictly positive height and adjacent diagrams are related by a single move of type $\mathcal{Y}$ or $\mathcal{Y}^{-1}$. Therefore in order to prove Markov's theorem, it suffices to consider only sequences of the second type, and the proof of Lemma 2.3 is complete. $\|$

Remark 2.1 The astute reader will have noticed the following: we have eliminated Reidemeister moves completely (they will not appear in the arguments that follow), nevertheless they played an important role already. We started with a sequence relating the given braids X and X' that consisted entirely of Reidemeister moves. We replaced it with a sequence of reducing moves and braid-like Reidemeister moves. The latter are in general not applied to diagrams of height zero, but we changed them to apply to diagrams of height zero. That is the moment when Traczyk's braid-like Reidemeister moves were changed to Markov moves. The modified sequence from X to X' has changed to a series of subsequences, each of which starts with a closed braid and ends with a closed braid, after which the ending braid is modified (by Markov moves) to a new closed braid, which is the initial closed braid in the next subsequence. As will be seen, Markov moves will not be used explicitly again until the proof of Lemma 2.8, where they are used (without the help of Reidemeister's Theorem) to relate very special closed braid diagrams.

Consider now a sequence of diagrams $Y = Y_1, \ldots, Y_n = Y'$ satisfying the criteria of Lemma 2.3. We note that, as in the above proof, we shall not in general distinguish between a diagram and its associated Seifert picture. Thus we shall make reference to 'a Seifert circle in the diagram Y_i', for example, meaning a Seifert circle in the Seifert picture associated to Y_i. In fact, we can think of each circle in a Seifert picture as forming part of the associated diagram, except in a small neighborhood of signed arcs, which correspond to crossings. Since reducing arcs avoid signed arcs, there is no ambiguity when referring to 'a reducing arc in a diagram'.

We now wish to consider the graph of the height function on our sequence. The graph will begin and end at height zero; each 'step' in between will either take us up 1 or down 1 since we have reduced to the case where all moves are reducing moves (or their inverses). We will examine local maxima in the height function. Let $Y(r), \hat{Y}, Y(s)$ be three consecutive diagrams in our sequence such that the height function has a local maximum at $\hat{Y}$. In other words, we have two reducing moves r, s with corresponding arcs α_r, α_s in $\hat{Y}$ such that reducing $\hat{Y}$ along α_r (resp. α_s) results in the diagram $Y(r)$ (resp. $Y(s)$), and it makes sense discuss $\alpha_r \cup \alpha_s$. We will call such a triple $\{Y(r), \hat{Y}, Y(s)\}$ a peak in the height function of our sequence. We define the height of the peak to be $h(\hat{Y})$ and define the height of the sequence to be the maximum value attained by the height function on the sequence, in other words, the maximum over the height of all the peaks in the sequence. In order to prove Theorem 4, we are going to induct on the height of the sequence.

Lemma 2.4 *We may assume that the reducing arcs involved in any peak in the height function of our sequence are disjoint. Further, the adjustments in our sequence of reducing moves which are required preserve the height of the sequence.*

Proof: Let $\{Y(r), \hat{Y}, Y(s)\}$ be a peak in the height function with associated reducing arcs α_r and α_s. We may always assume that the arcs intersect transversally and minimally. Suppose that $|\alpha_r \cap \alpha_s| = n \geq 2$. By smoothing out one or more of the points of intersection, we find a new reducing arc $\alpha_{r'}$ with the same endpoints as α_r such that $\alpha_r \cap \alpha_{r'} = \emptyset$ and $|\alpha_{r'} \cap \alpha_s| < n$. We can then replace the given peak $\{Y(r), \hat{Y}, Y(s)\}$ with two consecutive peaks $\{Y(r), \hat{Y}, Y_{r'})$ and $(Y_{r'}, \hat{Y}, Y(s)\}$. We call this procedure inserting the reducing operation r' at $\hat{Y}$, and it essentially amounts to replacing one peak with two peaks of the same height. In this way, we continue on until the intersection numbers of all adjacent pairs is at most 1.

Now suppose the arcs α_r, α_s associated to a given peak $\{Y(r), \hat{Y}, Y(s)\}$ have intersection number 1. If there exists a reducing arc α_t such that $\alpha_t \cap \alpha_r = \alpha_t \cap \alpha_s = \emptyset$, then we can insert the reducing operation t at $\hat{Y}$ to produce two peaks, each with a disjoint pair of associated reducing arcs. Suppose that the defect region which supports α_r and α_s contains no third reducing arc which is disjoint from both α_r and α_s. There is only one possible arrangement for such a defect region, shown in Figure 10.

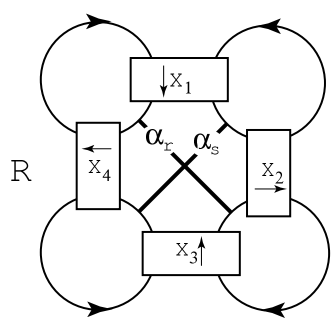


Figure 10: The case where two reducing arcs at a peak intersect once.

With this arrangement, the reducing arcs α_r and α_s must act on four distinct Seifert circles, possibly joined by signed arcs. We will now examine the region labelled R which lies 'outside' of the circles involved in our defect region and the signed arcs which join them. If R contains a Seifert circle, then it is easy to check that there must be a region somewhere in the diagram with three exposed circle. Such a region, as observed in the previous section, is necessarily a defect region which would contain a reducing arc disjoint from both α_r and α_s. If, on the other hand, R contains no Seifert circles, then it contains no signed arcs either, since all possible signed arcs between the four exposed circles of region R already appear in Figure 10. Thus we can join either pair of diagonally opposed circles by a reducing arc in R.

We conclude that if a peak $\{Y(r), \hat{Y}, Y(s)\}$ has associated arcs of intersection 1, we can always find a third reducing arc α_t so that we can insert the reducing operation t at $\hat{Y}$, thereby replacing the original peak with two peaks, each having a disjoint pair of reducing arcs. Since the operation of inserting a reducing operation at a peak preserves the height of the sequence, this finishes the proof of the lemma. $\|$

Thanks to the previous lemma, we can assume from now on that each peak in the graph of our height function corresponds to a disjoint pair of reducing arcs. Before we can state the next lemma, which concerns the peaks in the height function, we need to introduce a few some concepts. Note that when the reducing arcs involved in a peak $\{Y(r), \hat{Y}, Y(s)\}$ are disjoint, the reducing moves commute, that is, they can be performed in either order, starting with the diagram $\hat{Y}$ and resulting in the same diagram Y'. Further, as long as the reducing arcs α_r, α_s act on 3 or 4 distinct Seifert circles, we can perform the two reducing moves in either order with the same result. Observe that, since the reducing arcs α_r, α_s are disjoint in the diagram $\hat{Y}$, it makes sense to talk about the arc α_s (resp. α_r) in the context of the diagram $Y(r)$ (resp. $Y(s)$) obtained by reducing $\hat{Y}$ along α_r (resp. α_s). In this case we say we have a 'commuting pair' of reducing moves associated to the peak. If a peak $\{Y(r), \hat{Y}, Y(s)\}$ has a commuting pair, then we may replace it by a 'valley', that is, a subsequence $\{Y(r), Y', Y(s)\}$ where $h(Y') = h(Y) - 2$ and $Y' = Y(s \circ r) = Y(r \circ s)$ is the result of reducing $Y(r)$ along α_s, or equivalently reducing $Y(s)$ along α_r. Thus we can eliminate any peak corresponding to a commuting pair (such a peak necessarily has height at least 2).

In the case where two reducing arcs at a peak act on the same 2 circles, then after one move is performed, the second Reidemeister move will no longer be a reducing move; we call this a 'non-commuting pair' of reducing moves. Let $\{Y(r), \hat{Y}, Y(s)\}$ be a peak corresponding to a non-commuting pair of reducing arcs, and let C_1, C_2 be the two Seifert circles involved. Suppose there is a reducing arc α_t such that $\alpha_t \cap \alpha_r = \alpha_t \cap \alpha_s = \emptyset$ and such that t involves a circle other than C_1 or C_2, then we can insert t at $\hat{Y}$ to replace our peak $\{Y(r), \hat{Y}, Y(s)\}$ with two new peaks with commuting pairs of reducing moves:$\{Y(r), \hat{Y}, Y(t)\}$ and $\{Y(t), \hat{Y}, Y(s)\}$. As above, we now replace each peak with a 'valley': $\{Y(r), Y', Y(t)\}$ and $\{Y(t), Y'', Y(s)\}$, respectively, where $Y' = Y(t \circ r) = Y(r \circ t)$ is the diagram resulting from reducing $Y(r)$ by t (or equivalently, from reducing $Y(t)$ by r, and Y'' is the diagram resulting from reducing $Y(s)$ by t (or equivalently, from reducing $Y(t)$ by s). Again, this implies that the height of the original non-commuting peak $\{Y(r), \hat{Y}, Y(s)\}$ was at least 2. Thus, we can replace such a peak with peaks of strictly smaller height and repeat this process until all peaks either have height 1 or do not admit such a reducing arc α_t as above; we call a peak of the latter type <u>irreducible</u>.

We can now state the next lemma:

Lemma 2.5 *We may assume that any peak in the height function of our sequence either has height 1 or is irreducible.*

Proof: The proof is clear. In the discussion that preceded the statement of the Lemma we defined a peak to be irreducible in such a way that it subsumed all possibilities which did not allow us to reduce to height 1. $\|$

We note that each condition of Lemma 2.5 necessarily implies the non-commuting condition.

Lemma 2.6 *We may assume that no peaks in the height function of our sequence have height 1.*

Proof: Let $\{Y(r), \hat{Y}, Y(s)\}$ be a peak of height 1. We recall that height 1 implies that the two reducing arcs are non-commutative and hence involve precisely two circles. it is an easy exercise to show that these two circles must live either on the 'inside' or on the 'outside' of a band of circles, and that α_r and α_s are in fact equivalent as reducing moves. Therefore the diagram $Y(r)$ is equivalent to $Y(s)$ and we can simply eliminate this peak from our sequence. $\|$

It remains to deal with irreducible peaks in the height function of our sequence. Fortunately, it turns out that they only occur in a very particular way. To describe the particular way, define a weighted Seifert circle in a manner which is similar to the weights that we attach to closed braid diagrams, e.g. as in Figure 7. That is, A Seifert circle with weight w attached means a collection of w coherently oriented, nested, parallel Seifert circles. We use the term band for a Seifert circle with an attached weight.

Lemma 2.7 *If $\{Y(r), \hat{Y}, Y(s)\}$ is an irreducible peak in the height function of our sequence, then the diagram $\hat{Y}$ contains at most four bands, arranged as in Figure 11.*

Proof: Let $\{Y(r), \hat{Y}, Y(s)\}$ be an irreducible peak, and let C_1, C_2 be the two circles involved in the reducing arcs α_r and α_s. See Figure 11(i). For $i = 1, 2$, let D_i denote

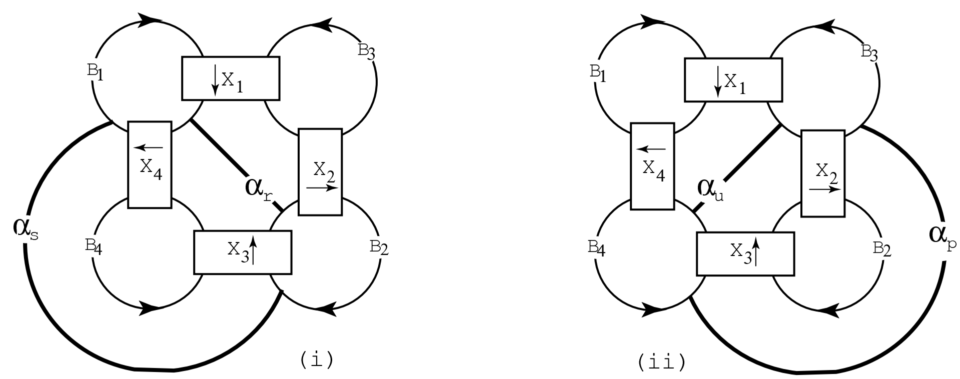


Figure 11: (i) The diagram corresponding to to the irreducible peak $\{Y(r), \hat{Y}, Y(s)\}$. (ii) Two additional reducing arcs used to eliminate irreducible peaks. Each pair α_r, α_s and α_p, α_u is a possible non-commuting pair associated to the peak. The blocks X_1, X_2, X_3, X_4 indicate (possibly empty) collections of signed arcs.

the disk bounded by C_i in S^2 which does not contain the reducing arcs α_r and α_s. Then $S^2 \setminus (D_1 \cup D_2 \cup \alpha_r \cup \alpha_s)$ has two components. By assumption, neither component can contain a defect region (or else we could find a reducing arc α_t as above, contradicting the irreducibility of our peak). Thus if either component contains any Seifert circles, the circles must form a band, oriented oppositely to C_1 and C_2. We allow the possibility that the weight of these bands is zero. The same reasoning shows that D_i cannot contain a defect region for $i = 1, 2$ and hence that C_i must be the outer circle of a band. It is of course

possible that some braiding takes place between adjacent coherent circles, as indicated in Figure 11(i). The braids joining the various bands are labelled X_i. Thus we have a diagram of the given form, in which each circle pictured represents a band, and the lemma is proved.$\|$
The following lemma will allows us to replace irreducible peaks in the height function of

our sequence with peaks of strictly smaller height.

Lemma 2.8 *Let $\{Y(r), \hat{Y}, Y(s)\}$ be an irreducible peak of height $n+1$ in the height function of our sequence. Then there exist sequences of diagrams $Y(r) = Y_1^r, \ldots, Y_n^r = Y(p \circ r)$ and $Y(s) = Y_1^s, \ldots, Y_n^s = Y(u \circ s)$ such that Y_{i+1}^r (resp. Y_{i+1}^s) is obtained from Y_i^r (resp. Y_i^s) by a reducing move and such that $h(Y(p \circ r)) = h(Y(u \circ s) = 0$ and $Y(p \circ r)$ and $Y(u \circ s))$ are Markov equivalent.*

Before discussing the proof of this lemma, we show how to use it to prove the Markov theorem. Using Lemmas 2.3-2.7, we have reduced the proof of Markov's Theorem to the situation of two closed braid diagrams X and X' related by a sequence of diagrams related by reducing moves (and their inverses) such that the height of each intermediate diagram is strictly positive and such that any peak in the height function of the sequence is irreducible (of height at least 2). Let $\{Y(r), \hat{Y}, Y(s)\}$ be an irreducible peak of height n in the height function of our sequence. By Lemma 2.8, we can replace this subsequence with a subsequence of strictly smaller height (possibly including a subsequence entirely at height 0 related by Markov moves, which we can remove from consideration as before). In doing so, we create new peaks whose height is *strictly lower* than the height of the peak being replaced. If we perform this operation at every irreducible peak, then we obtain a new sequence relating our closed braid diagrams Y, Y'.

The new peaks may or may not be irreducible; in fact, their corresponding arcs may not even be disjoint, but now we are back in the same situation as we were before Lemma 2.4 except that we are starting with a sequence of lower height. Thus by induction on the height of the sequence (with base case provided by Lemma 2.6), we can replace any sequence as described in Lemma 2.3 with a sequence consisting entirely of diagrams of height zero. This completes the proof of Theorem 4, modulo the proof of Lemma 2.8.$\|$

Sketch of the Proof of Lemma 2.8. By Lemma 2.7, the diagram $\hat{Y}$ has at most four bands which we label B_1, B_2, B_3, and B_4, joined by four (possibly trivial) braids X_1, X_2, X_3, X_4, as in Figure 11(i) and (ii). Let w_i denote the weight of the band B_i. Note that, e.g., the braid X_1 has $w_1 + w_3$ strands, and similarly for the other X_i. If w_i or w_j is greater than 1, then a reducing arc joining B_i, B_j is understood to indicate a sequence of $w_i w_j$ reducing moves. There are several ways to construct such a sequence; we adopt the convention that we choose reducing moves in such a way that the strands of one band all slide under or else all slide over the strands of the other band.

Referring again to Figure 11, we first perform the reducing move r along the arc α_r and then reduce again via the arc α_p involving B_3 and B_4, giving us some number of reducing moves depending on the weights of the bands involved. The resulting diagram $Y(p \circ r)$ is the closure of the first braid shown in Figure 12, up to a choice of arcs passing over or under in the various Reidemeister moves of type II. This gives us the first sequence of the lemma.

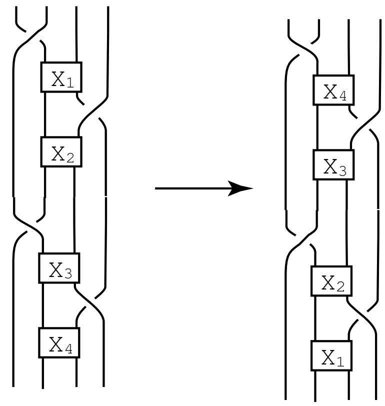


Figure 12: Two braids whose respective closures are the result of performing the two pairs of reducing moves r, p and s, u indicated in Figure 11.

To get the second sequence of the lemma, we begin instead with the reducing move s along the arc α_s and then reduce again via the arc α_u. The resulting diagram $Y(u \circ s)$ is the closure of the second braid shown in Figure 12.

We have found our two sequences of reducing moves, and we have reduced the proof of Markov's Theorem to just one specific calculation, namely, showing that the closed braids $Y(p \circ r)$ and $Y(u \circ s)$ corresponding to the two braids of Figure 12 are M-equivalent. The two braids can be related by a sequence consisting of several braid isotopies as well as two stabilizations and two destabilizations; the reader is referred to [129] for the details of this calculation. This concludes the proof of Lemma 2.8, and so also of Theorem 4. $\|$

Remark 2.2 Our choice of 'over' or 'under' in the reducing moves p and u leads to possible ambiguity, but the various diagrams which would result from different choices are all related by 'exchange moves', which are defined and discussed at the beginning of the next section. We will show in the next section that an exchange move replaces a sequence of 4 Markov moves: a braid isotopy, a stabilization, a second braid isotopy and a destabilization.

The essential groundwork has been laid regarding the connection between braids and knots. From this point on, all knot diagrams will be assumed to be in closed braid form, i.e. in the form indicated in Figure 3. In the sections that follow we will examine many consequences.

3 Braid foliations

We begin our study of new results which relate to the study of knots via closed braids by presenting some results that use the theory of braid foliations of a Seifert surface bounded by a knot which is represented as a closed braid. We will develop three applications of braid foliations: The first is Theorem 5, in §3.1. We give an essentially complete proof of the 'Markov theorem without stabilization' (MTWS) in the special case of the unknot, based upon the presentation in [21]. In §3.2 we state the MTWS, Theorem 6, in the general case. A full proof of that theorem can be found in [30]. In the final section, §3.3 we give an application of the MTWS to contact topology.

3.1 The Markov Theorem Without Stabilization (special case: the unknot)

After describing the basic ideas about braid foliations, we will apply them to the study of a classical problem in topology, the unknot recognition problem. Alexander's Theorem (Theorem 2) tells us that every link K may be represented as a closed n-braid, for some n. Markov's Theorem (Theorem 4) tells us how any two closed braid representatives of the same knot or link are related. Looking for a way to simplify a given closed braid representative of a knot or link systematically, the first author and Menasco were lead to the study of the unknot as a key example. There is an obvious choice of a simplest representative, namely a 1-braid representative. The Markov Theorem Without Stabilization for the unknot, which is stated below as Theorem 5, asserts that, in the special case of the unknot, the stabilization move of Markov's Theorem can be eliminated, at the expense of adding the exchange move. Therefore we begin with a discussion of the exchange move, which is defined in Figure 14, and the reason why it is so important.

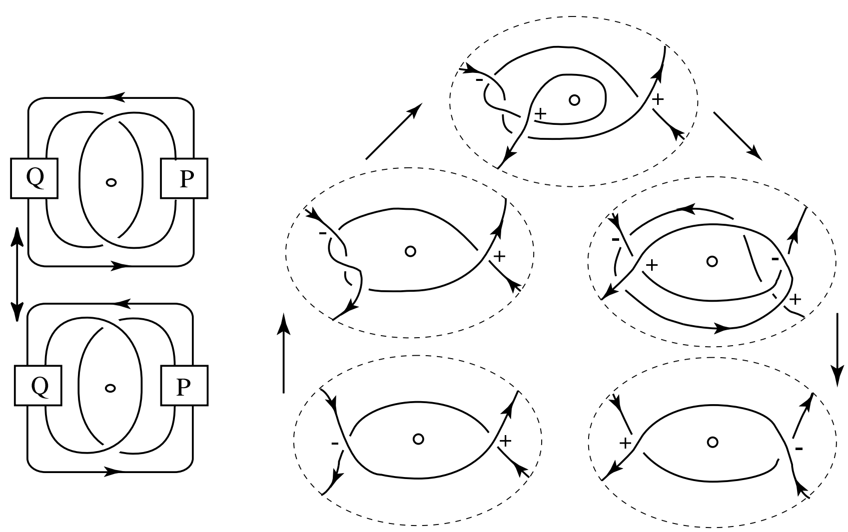


Figure 13: The left top and bottom sketches define the exchange move. The right sequence of 5 sketches shows how it replaces a sequence of Markov moves which include braid isotopy, a single stabilization, additional braid isotopy and a single destabilization.

A natural question to ask is what is accomplished by stabilization, braid isotopy and destabilization? Figure 14 shows that exchange moves are the obstruction to sliding a

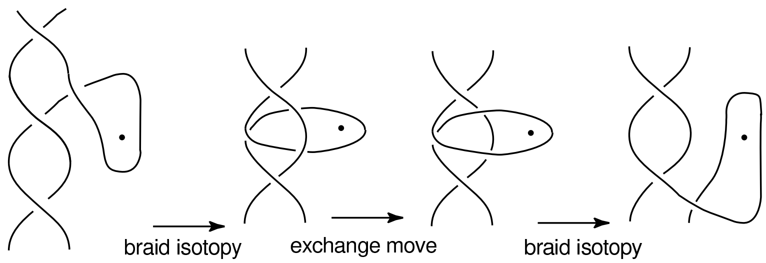


Figure 14: Exchange moves are the obstruction to sliding a trivial loop around a closed braid.

trivial loop around a braid [34]. However there is more at issue than just sliding trivial loops around a braid. As is illustrated in Figure 15 a sequence of exchange moves, together with braid isotopy, can create infinitely many closed braid representatives of a single knot type, all of the same braid index. This bypasses a key question: do exchange moves actually change conjugacy class? The answer is "yes", and in fact the phenomenon first appears in the study of closed 4-braid representatives of the unknot. In §5 we show that there is a definitive test for proving it.

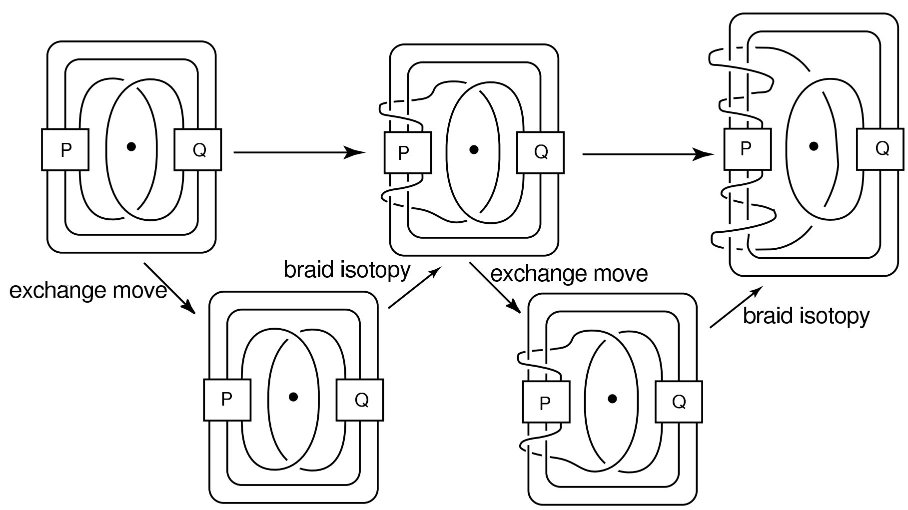


Figure 15: A sequence of exchange moves and braid isotopies with non-trivial consequences.

Keeping Figures 13, 14 and 15 in mind, we are ready to state the MTWS for the unknot:

Theorem 5 [29] *Every closed braid representative K of the unknot $\mathcal{U}$ may be reduced to the standard 1-braid representative U_1, by a finite sequence of braid isotopies, destabilizations and exchange moves. Moreover there is a complexity function associated to closed braid representative in the sequence, such that each destabilization and exchange move is strictly complexity-reducing.*

The first proof of Theorem 5 was the one in [29]. A somewhat different and slicker proof can be found in [21], but it requires more machinery than was necessary for present purposes. We follow the proof in [29]. However, parts of our presentation and most of our figures were essentially lifted (with the permission of both authors) from [21], a review article on braid foliation techniques. Our initial goal is to set up the machinery needed for the proof.

Recalling the definition of a closed braid from §2.1, we adopt some additional structure and say that K is in closed n-braid form if there is an unknot A in $S^3 \setminus K$, and a choice of fibration H of the solid torus $S^3 \setminus A$ by meridian disks, such that K intersects each fiber of H transversely. Sometimes it is convenient to replace S^3 by R^3 and to think of the fibration H as being by half-planes $\{H_\theta; \theta \in [0, 2\pi]\}$ of constant polar angle θ through the z-axis. Note that K intersects each fiber H_θ in the same number of points, that number being the index n of the closed braid K. We may always assume that K and A can be oriented so that K travels around A in the positive direction, using the right hand rule. Now let K be a closed braid representative of the unknot, and let $\mathcal{D}$ denote a disk spanned by K, oriented so that the positive normal bundle to each component has the orientation induced by that on $K = \partial\mathcal{D}$. In this section we will describe a set of ideas which shows that there is a very simple method that changes the pair $(K, \mathcal{D})$ to a planar circle that bounds a planar disc, via closed braids, moreover there is an associated complexity function that is strictly reducing. The ideas that we describe come from [29], however our main reference will be to the review article [21].

The braid axis A and the fibers of H will serve as a coordinate system in 3-space in which to study $\mathcal{D}$. A singular foliation of $\mathcal{D}$ is induced by its intersection with fibers of H. A singular leaf in the foliation is one which contains a point of tangency with a fiber of H. All other leaves are non-singular.

It follows from standard general position arguments that the disk $\mathcal{D}$ can be chosen to be 'nice' with respect to our fibration. More precisely, we can assume the following:

(i) The intersections of A and $\mathcal{D}$ are finite in number and transverse.

(ii) There is a neighborhood N_A of A in $R^3 \setminus K$ such that each component of $\mathcal{D} \cap N_A$ is a disk, and each disk is radially foliated by its arcs of intersection with fibers of H. There is also a neighborhood N_K of K in R^3 such that $N_K \cap \mathcal{D}$ is foliated by arcs of intersection with fibers of H which are transverse to K.

(iii) All but finitely many fibers H_θ of H meet $\mathcal{D}$ transversely, and those which do not (the singular fibers) are each tangent to $\mathcal{D}$ at exactly one point in the interior of both $\mathcal{D}$ and H_θ. Moreover, each point of tangency is a saddle point (with respect to the parameter θ). Finally, each singular fiber contains exactly one singularity of the foliation, each of which is a saddle point.

It can also be assumed (see [21] for details) that

(iv) Each non-singular leaf is either an *a*-arc, which has one endpoint on A and one on $K = \partial\mathcal{D}$ or a *b*-arc, which has both endpoints on A.

(vii) Each b-arc in a fiber H_θ separates that fiber into two components. Call the b-arc essential if each of these components is pierced at least once by K, and inessential otherwise. Then we also have that all b-arcs in the foliation of $\mathcal{D}$ may be assumed to be essential.

Note that $\mathcal{D}$ cannot be foliated entirely by b-arcs, since such a surface is necessarily a 2-sphere. Thus, if the foliation of $\mathcal{D}$ contains no singularities, $\mathcal{D}$ is foliated entirely by a-arcs. Otherwise, let U be the union of all the singular leaves in the foliation of $\mathcal{D}$. Moving forward through the fibration, we see that any singular leaf in the foliation is formed by non-singular leaves moving together to touch at a saddle singularity. The three types of singular leaves which can occur are labelled aa, ab or bb, corresponding to the non-singular leaves associated to them. Now each singular leaf λ in U has a foliated neighborhood N_λ in $\mathcal{D}$ such that $N_\lambda \cap U = \lambda$. According to whether λ has type aa, ab, or bb, N_λ is one of the foliated open 2-cells shown in Figure 16(i), with the arrows indicating the direction of increasing θ.

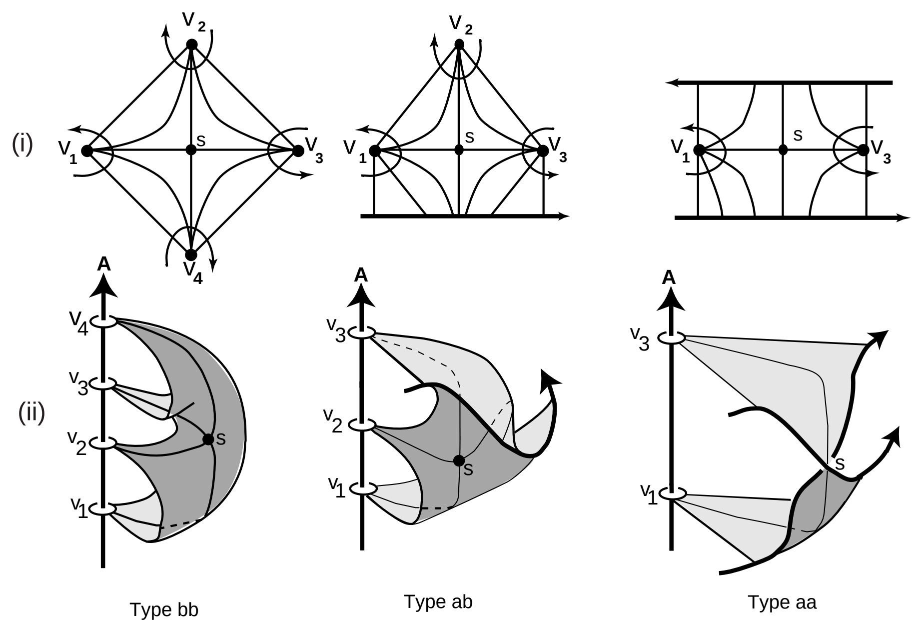


Figure 16: (i) The three types of tiles in the decomposition of $\mathcal{D}$; (ii)The canonical embedding of each type of tile.

The complement of U in $\mathcal{D}$ is a union $B_1 \cup B_2 \cup ... \cup B_k$, where each B_i is foliated entirely by a-arcs or entirely by b-arcs. Choose one non-singular leaf in each B_i and declare it to be a boundary arc of type a or type b according to whether it is an a-arc or b-arc, respectively. Then the union of all boundary arcs determines a tiling of $\mathcal{D}$, that is, a decomposition into regions called tiles, each of which is a foliated neighborhood of one singular leaf. Each tile has type aa, ab and bb, according to the type of its unique singularity. Note that a tiling of $\mathcal{D}$ is a foliated cell-decomposition. We further define the sign of a singular point s in the foliation of $\mathcal{D}$ to be positive if the positive normal to $\mathcal{D}$ points in the direction of increasing θ in the fibration, negative otherwise. The sign of a tile is then defined to be the sign of its singularity.

The axis A intersects the surface in a finite number of points, called vertices of the tiles. Each vertex v is an endpoint of finitely many boundary arcs in the surface decomposition. Let the type of v be the cyclic sequence $(x_1, ..., x_r)$, where each x_i is either a or b, and the sequence lists the types of boundary arcs meeting at v in the cyclic order in which they occur in the fibration. See Figure 17. The valence of a vertex v is the number of distinct

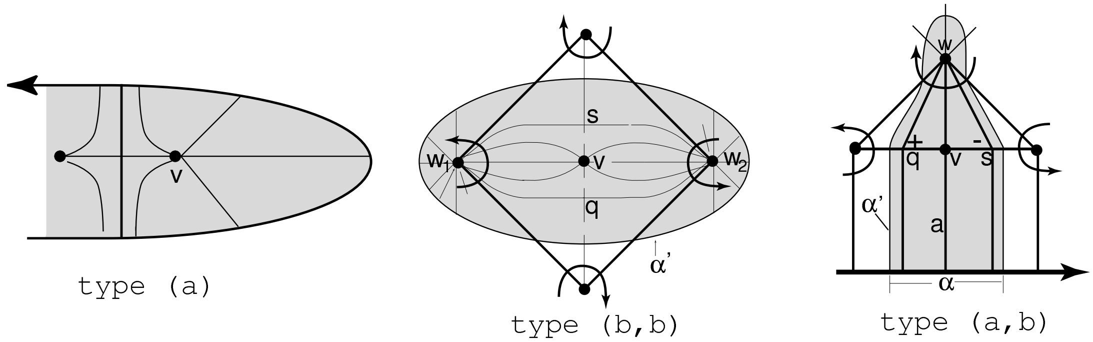


Figure 17: A vertex v of type $(a), (ab)$ and (bb).

tiles intersecting at v. The sign of a vertex v is the cyclic array of signs of the tiles meeting at v. See Figure 17.

The parity of v is said to be positive (resp. negative) if the outward-drawn normal to the surface has the same (resp. opposite) orientation as the braid axis at the vertex. Thus when we view the positive side of the surface, the sense of increasing θ around a vertex will be counterclockwise (resp. clockwise) when the vertex is positive (resp. negative), as illustrated in Figure 16(i).

Finally, we note that (again, see [21] for a proof) that, up to a choice of their sign, tiles of type aa, ab, bb each have a canonical embedding in 3-space, which is determined up to an isotopy of 3-space which preserves the axis A and each fiber of H setwise. The canonical embedding for each tile is shown in Figure 16(ii).

The decompositions of the disk $\mathcal{D}$ which we have just described are not unique. We shall now describe three ways in which they can be changed. In each of the three cases the possibility of making the change is indicated by examining the combinatorics of the tiling. The change is realized by an isotopy of the disk which is supported in a neighborhood N of a specified small number of tiles of type aa, ab or bb, leaving the decomposition of $\mathcal{D}$ unchanged outside N.

For a spanning disk $\mathcal{D}$ and a tiling $\mathcal{T}$, we denote the complexity of the tiling by $c(\mathcal{D}, \mathcal{T})$, where we define $c(\mathcal{D}, \mathcal{T}) = (n, S)$, where n is the braid index of the boundary and S is the number of singularities in the tiling. Setting V_+ and V_- equal to the number of positive and negative vertices, one has $n = V_+ - V_-$, so n is also determined by the tiling. Tiled discs $(\mathcal{D}, \mathcal{T})$ can then be ordered, using lexicographical ordering of the associated pair $c(\mathcal{D}, \mathcal{T})$. Each of our three moves will replace the given disk and tiling $\mathcal{D}, \mathcal{T}$ with some $\mathcal{D}', \mathcal{T}'$, with the following effects on complexity:

- **Change in foliation:** $(n', S') = (n, S)$.

- **Destabilization:** $(n', S') = (n-1, S-1)$.
- **Exchange moves (two types):** $(n', S') = (n, S-2)$.

We focus here only on the combinatorics of tilings which admit one of the above moves, as well as the effect of each move on the embedding of $\mathcal{D}$ and the new tiling of $\mathcal{D}$ which results. For proof and further details, see [21].

•Changes in foliation. The choice of a foliation of $\mathcal{D}$ is not unique, and our first move involves ways in which the surface decomposition can be changed by an isotopy of $\mathcal{D}$ or, equivalently, by an isotopy of the fibers of H keeping $\mathcal{D}$ fixed. This particular change was introduced in [28] for 2-spheres and was modified in [29] for certain spanning surfaces. In what follows, we say that two tiles are adjacent if they have a common boundary arc in the given tiling.

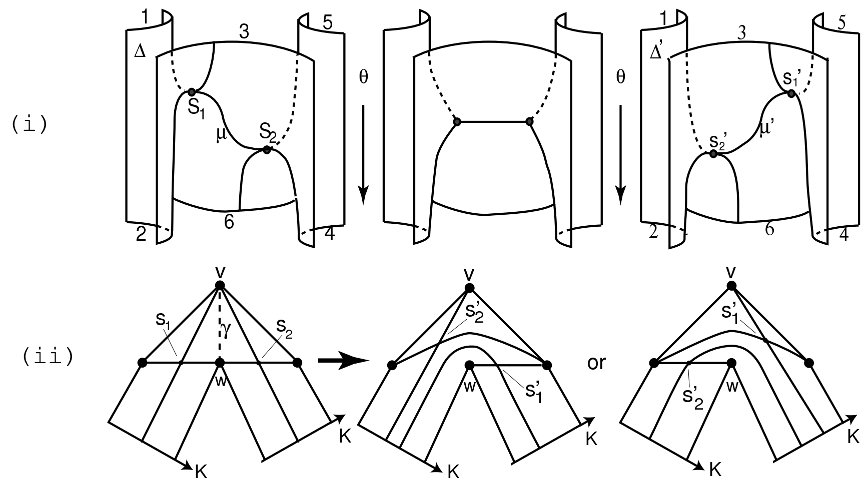


Figure 18: Sketch (i) shows an isotopy of $\mathcal{D}$ that induces a change in foliation. Sketch (ii) shows the effect of the change on two ab-tiles

A tiling admits a change in foliation whenever there are two tiles T_1, T_2 of the same sign adjacent at a b-arc. Roughly speaking, a change in foliation is a local isotopy of the surface which pushes two saddle points past each other. Locally, the disk $\mathcal{D}$ is embedded as in the left sketch in Figure 18(i). A change in foliation is defined as the passage from the embedding left to the right embedding. Figure 18(ii) the effect of this move on the local foliation, while Figure 18(iii) illustrates the effect of a change in foliation on the tiling, in the case of two adjacent (ab)-tiles.

•Destabilization via a type (a) **vertex.** A vertex v of valence 1 in a tiling of $\mathcal{D}$ occurs when two of the edges in a single tile T are identified in $\mathcal{D}$. Such a vertex must have type (a), and therefore T is an aa-tile. See Figure 17. Since there is a canonical embedding for an aa-tile in 3-space (Figure 16), identifying two edges of an aa-tile with endpoints on a common vertex yields the canonical embedding for T shown in Figure 19(i). Notice that there is a radially foliated disk D in T cut off from $\mathcal{D}$ by the arc of the singular leaf of T

with both endpoints on K. Hence there must be a trivial loop in the braid representation of K, and we can modify the braid and the disk $\mathcal{D}$ in the manner illustrated. We call this modification destabilization via a type (a) vertex. The effect on the tiling of $\mathcal{D}$ is that the aa-tile T and its type (a) vertex v are deleted, while the tiling outside T is unaltered. The braid index is decreased by one. See the right sketch in Figure 19(i).

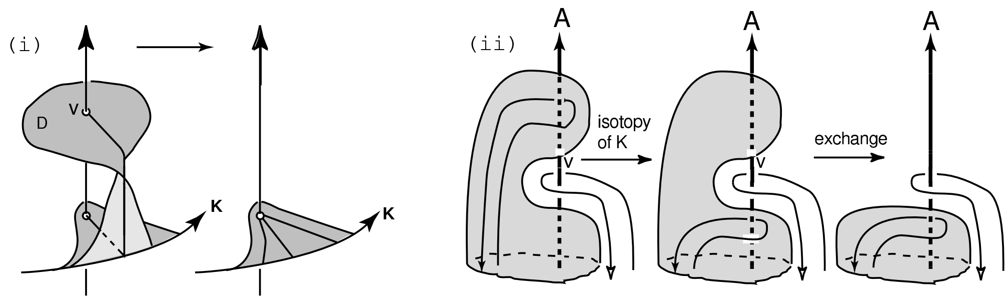


Figure 19: The effect of (i) a destabilization, and (ii) an exchange move of type (bb), on the embedded (and foliated) foliated disc $\mathcal{D}$ and on the braid K that is its boundary.

•**Exchange moves.** The destabilization move that we just described can be accomplished if a vertex of valence 1 exists in our tiling. Exchange moves are based upon the existence of a vertex of valence 2. An example of the type of move we would like to achieve on the braid itself was shown in sketch (ii) of Figure 19(ii), and also in in the second row from the top in Figure 20. There are two different types of tiling patterns, and hence two different types of local embeddings of $\mathcal{D}$, which support exchange moves:

Exchange Move Type (bb): We now consider a vertex v with sign $(\pm, \mp)$ and type (b, b). In this situation, two bb-tiles are adjacent along consecutive b-arcs. The canonical embeddings of the tiles in this case are shown in Figure 16, and the overall modifications to $\mathcal{D}$ are suggested by Figure 19(ii). The end result is a new surface $\mathcal{D}'$ whose decomposition has at least two fewer vertices (in particular, the vertex v is deleted) and at least two fewer regions than the decomposition of $\mathcal{D}$. $\|$

Exchange Move, Type (ab) An exchange move of type (ab) is possible whenever the tiling of $\mathcal{D}$ has a vertex v of valence 2, type (a,b), and sign $(\pm, \mp)$. Such a vertex occurs only when two ab-tiles are adjacent along corresponding a and b edges, as in Figure 17. The isotopy of $\partial\mathcal{D} = K$ is achieved by pushing a subarc α of K across a disk which is contained in the union of two (ab) tiles T_1, T_2, to a new arc α'. The effect on the tiling is that the ab-tiles T_1 and T_2 are deleted and any adjacent ab-tiles (resp. bb-tiles) become aa-tiles (resp. ab-tiles).

In order to prove the MTWS for the unknot, we must locate the places on the tiling of $\mathcal{D}$ where the three moves just described can be made. We begin with two lemmas. The first of the two is implicit in [9], but we refer the reader to [21] for an explicit proof.

Lemma 3.1 [9] *Let v be a vertex of type $(b, b, \ldots, b)$. Then the set of all singularities s which lie on a singular leaf ending at v contains both positive and negative singular points.*

Lemma 3.2 [29] *Suppose that the disk $\mathcal{D}$ is nontrivially tiled and has no vertices of valence 1. Then the tiling of $\mathcal{D}$ contains a vertex of type (ab), (bb), or (bbb).*

Proof. The proof is essentially an Euler characteristic argument. A tiling $\mathcal{T}$ of the disk $\mathcal{D}$ corresponds to a cell decomposition of the 2-sphere in the following way. Let V be the number of vertices in $\mathcal{T}$, let E be the number of boundary arcs in $\mathcal{T}$, which we will think of as edges, and let F be the number of tiles in $\mathcal{T}$. Now let Σ be the 2-sphere obtained by collapsing $\partial\mathcal{D} = K$ to a point. This gives a cell decomposition of Σ with $V + 1$ 0-cells (vertices), E 1-cells (edges), and F 2-cells (faces). We know that the Euler characteristic $\chi(\Sigma) = 2$, and therefore we have that $V - E + F = 1$. We also know that every face in our cell decomposition of Σ has 4 edges, and that every edge has 2 adjacent faces, and so $E = 2F$. We therefore have the following equation.

$$2V - E = 2. \tag{10}$$

Now let $V(\alpha, \beta)$ denote the number of vertices in $\mathcal{T}$ with α adjacent a-arcs and β adjacent b-arcs. If v is the valence of such a vertex, then $v = \alpha + \beta$, and so $V(\alpha, v - \alpha)$ denotes the number of vertices in $\mathcal{T}$ with valence v and α adjacent a-arcs. Since by hypothesis $v \neq 1$, we can now count the number of vertices in $\mathcal{T}$ using the following summation:

$$V = \sum_{v=2}^{\infty} \sum_{\alpha=0}^{v} V(\alpha, v - \alpha)$$

Now write $E = E_a + E_b$, where E_a denotes the number of edges in $\mathcal{T}$ of type a, and similarly for E_b. Since an a-edge is incident at one vertex in $\mathcal{T}$, and a b-edge is incident at two vertices in $\mathcal{T}$, we can write:

$$E_a = \sum_{v=2}^{\infty} \sum_{\alpha=0}^{v} \alpha V(\alpha, v - \alpha) \tag{11}$$

$$E_b = \frac{1}{2} \sum_{v=2}^{\infty} \sum_{\alpha=0}^{v} (v - \alpha) V(\alpha, v - \alpha) \tag{12}$$

Substituting into Equation 10, we now have

$$\sum_{v=2}^{\infty} \sum_{\alpha=0}^{v} (4 - v - \alpha) V(\alpha, v - \alpha) = 4.$$

Observe now that if $v \geq 4$, the coefficient $(4 - v - \alpha) \leq 0$. Writing out the terms corresponding to $v = 2$ and $v = 3$ gives, respectively, $2V(0,2) + V(1,1)$ and $V(0,3) - V(2,1) - 2V(3,0)$. Each $V(\alpha, \beta) \geq 0$ by definition. Hence we can move terms around in the above equation so that each term is nonnegative, as follows:

$$2V(0,2) + V(1,1) + V(0,3) = 4 + V(2,1) + 2V(3,0) + \sum_{v=4}^{\infty} \sum_{\alpha=0}^{v} (4 - v - \alpha) V(\alpha, v - \alpha).$$

The right-hand side of the equation is clearly ≥ 4, and therefore at least one term on the left-hand side is nonzero. Since $V(0,2), V(1,1)$, and $V(0,3)$ record the number of type (bb), type (ab), and type (bbb) vertices, respectively, the lemma is proved. $\|$

We are now ready to prove Theorem 5, the MTWS in the case of the unknot.

Proof of Theorem 4. We start with an arbitrary closed braid representative K of the unknot. Our closed braid K is the boundary of a disk $\mathcal{D}$ which admits a tiling $\mathcal{T}$. Note that if the complexity of the tiling $c(\mathcal{D},\mathcal{T})$ is equal to $(1,0)$, then our disk is radially foliated by a-arcs and K is the standard embedding of the unknot. Therefore we assume $c(\mathcal{D},\mathcal{T}) > (1,0)$. We will use induction on $c(\mathcal{D},\mathcal{T})$ to show that after a finite sequence of our three moves, $\mathcal{D}$ is radially foliated by a-arcs.

We first observe that if there exists a vertex of type (a) in $\mathcal{T}$, then we can destabilize along this vertex, thereby reducing $c(\mathcal{D},\mathcal{T})$. Hence we can asume that all vertices in $\mathcal{T}$ have valence at least 2. By Lemma 3.2, $\mathcal{D}$ must have a vertex of type $(ab), (bb)$, or (bbb). If there is a type (bbb) vertex, then at least two of the adjacent tiles must have the same sign. In this case we can do a change in foliation, replacing our (bbb) vertex with a (bb) vertex. Therefore we may assume that $\mathcal{D}$ contains a vertex of type (ab) or type (bb).

If $\mathcal{D}$ has a type (ab) vertex with sign (+,+) or (-,-), then we can apply a change in foliation which replaces this vertex with a type (a) vertex. We now destabilize along this type (a) vertex in order to reduce the complexity. If $\mathcal{D}$ has a type (ab) vertex with sign (+, -), then we can do an exchange move of type (ab), thereby reducing complexity. If $\mathcal{D}$ has a type (bb) vertex, then it is an interior vertex, and by Lemma 3.1, it has sign (+,-). Then we can do an exchange move of type (bb), which reduces complexity.

Thus if $c(\mathcal{D},\mathcal{T}) > (1,0)$, we can always reduce until $c(\mathcal{D},\mathcal{T}) = (1,0)$, and the theorem is proved. $\|$

Example 3.1 [Morton's irreducible 4-braid] We now give an example illustrating the necessity of exchange moves. In [109], Morton gave the following example of an <u>irreducible</u> braid:

$$X = \sigma_3^{-2}\sigma_2\sigma_3^{-1}\sigma_2\sigma_1^3\sigma_2^{-1}\sigma_1\sigma_2^{-1},$$

i.e., a braid whose closure represents the unknot but cannot be isotoped to the unknot in the complement of the braid axis. Now, it is clear that if a braid in $\mathbf{B}_n$ admits a factorization of the form $w_1\sigma_{n-1}w_2\sigma_{n-1}^{-1}$, where w_1 and w_2 are words in $\sigma_1, \ldots, \sigma_{n-2}$, then its closure admits an exchange move. As observed in [29], the conjugate braid $(\sigma_3\sigma_2\sigma_1)X(\sigma_1^{-1}\sigma_2^{-1}\sigma_3^{-1})$ is isotopic to the braid

$$X' = (\sigma_2^{-2}\sigma_1\sigma_2^{-1})\sigma_3(\sigma_2^3\sigma_1^{-1}\sigma_2)\sigma_3^{-1}.$$

Thus X' admits an exchange move to obtain the new braid word $(\sigma_2^{-2}\sigma_1\sigma_2^{-1})\sigma_3^{-1}(\sigma_2^3\sigma_1^{-1}\sigma_2)\sigma_3$, which can be cyclically rewritten as

$$X'' = (\sigma_2\sigma_3\sigma_2^{-2})\sigma_1(\sigma_2^{-1}\sigma_3^{-1}\sigma_2^3)\sigma_1^{-1}.$$

If we now interchange the axis of our braid with the point at infinity, we can perform another exchange move to obtain

$$X''' = (\sigma_2\sigma_3\sigma_2^{-2})\sigma_1^{-1}(\sigma_2^{-1}\sigma_3^{-1}\sigma_2^3)\sigma_1$$

which can in fact be isotoped to the unknot in the complement of its braid axis [29].

Wright has given a foliated disk corresponding to this braid in [133]. The reader is cautioned that Wright uses a different notation convention for braid words than that used in this paper (see p. 98 of [133]). We also refer the reader to [32] for further examples. ♠

We note that Theorem 5 has been used as the basis for an algorithm for recognizing the unknot. See [22]. This algorithm has been put on a computer [32], however it needs more work before it can become a practical tool for recognizing the unknot.

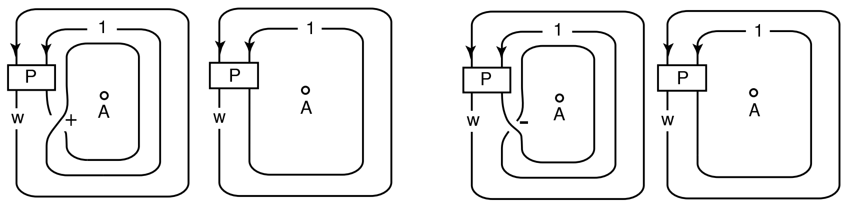
Open Problem 2 *At this writing the development of a sound and practical algorithm for unknot recognition remains one of the major open problems in low dimensional topology.* ♣

With regard to Problem 2, we note that Theorem 5 proves the existence of a monotonic simplification process which begins with an arbitrary closed n-braid representative of the unknot and ends with a 1-braid representative, but the complexity function $c(\mathcal{D}, \mathcal{T})$ which gives instructions for the process is 'hidden' in the tiling of the surface $\mathcal{D}$. We need it in order to know when complexity-reducing destabilizations and exchange moves are possible. The reader who is interested in this problem might wish to consult [22], where a somewhat different approach suggests itself. Instead of working with the tiled surface, one may work with the 'extended boundary word' of [22]. The latter is a closed braid which is obtained from an arbitrary closed braid representative of the given knot by threading in additional 1-braids, and it seems likely that it will give an alternative monotonic reduction process. Thus we suggest:

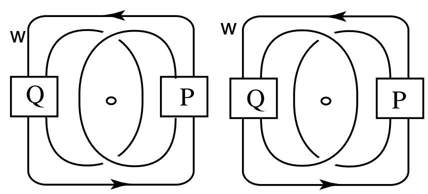
Open Problem 3 *Investigate the monotonic reduction process of Theorem 5 from the point of view of the extended boundary word of* [22]. ♣

3.2 The Markov Theorem Without Stabilization, general case

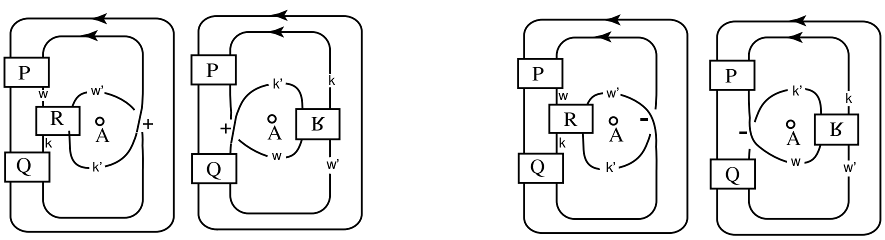
In this section we state the generalized version of Theorem 5 which was established in [30] for arbitrary closed braid representatives of arbitrary knots and links. The moves which are needed are of course much more complicated than in the case of the unknot. They are described in terms of 'block strand diagrams' and 'templates'. The concept of a block-strand diagram is fairly easy to understand from the 8 examples in Figure 20. The first important feature of a block-strand diagram is that after an assignment of a braided tangle to each block, it becomes a closed braid that represents a specific knot or link. Thus each block strand diagram determines infinitely many closed braid representatives of (presumably) infinitely many knots and links. A <u>template</u> is a pair of block-strand diagrams, both of which represent the same knot or link, when we make the same braiding assignments to corresponding blocks. Note that there are two destabilization templates (they differ in the sign of the 'trivial loop' that is removed). There are also two flype templates, which differ in the sign of the single crossing which is outside all the braid boxes. Templates always have an associated braid index, namely the braid index of any knot or link of minimum braid index that they carry. Destabilization templates occur for every braid index ≥ 2. Since the exchange move can be realized by braid isotopy when the braid index is ≤ 3, it does not play a role until braid index ≥ 4. The flype templates occur for braid index ≥ 3.



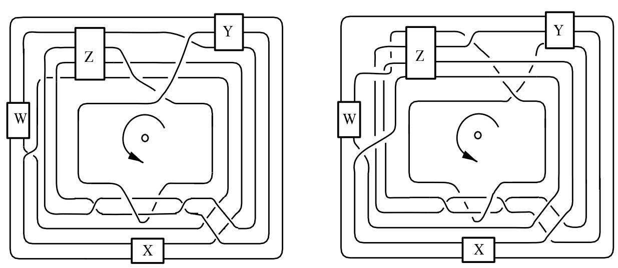
The two destabilization templates



The exchange move template



The two flype templates



A 6-braid template

Figure 20: Examples of templates

The method by which the 6-braid template in Figure 20 was constructed is described in the manuscript [30].

The two block strand diagrams in a template are always related by a sequence of Markov moves, however the sequence may be quite complicated, and so the isotopy that takes the left diagram to the right diagram is in general not obvious. This fact is illustrated by the 6-braid template of Figure 20. (See [30] for the isotopy, as an explicit sequence of Markov moves.)

With this brief introduction, we are able to state the Markov Theorem Without Stabilization in the general case:

Theorem 6 [30] *Let $\mathcal{B}$ be the collection of all braid isotopy classes of closed braid representatives of oriented knot and link types in oriented 3-space. Among these, consider the subcollection $\mathcal{B}(K)$ of representatives of a fixed link type K. Among these, let $\mathcal{B}_{min}(K)$ be the subcollection of representatives whose braid index is equal to the braid index of K. Choose any $X_+ \in \mathcal{B}(K)$ and any $X_- \in \mathcal{B}_{min}(K)$. Then there is a complexity function which is associated to X_+, X_-, and for each braid index m a finite set $\mathcal{T}(m)$ of templates is introduced, each template determining a move which is non-increasing on braid index, such that the following hold: First, there is are initial sequences which modify $X_- \to X'_-$ and $X_+ \to X'_+$:*

$$X_- = X_-^1 \to \cdots \to X_-^p = X'_-, \qquad X_+ = X_+^1 \to \ldots \to X_+^q = X'_+$$

Each passage $X_-^j \to X_-^{j+1}$ is strictly complexity reducing and is realized by an exchange move, so that $b(X_-^{j+1}) = b(X_-^j)$. These moves 'unwind' X_-, if it is wound up as in the top right sketch in Figure 15. Each passage $X_+^j \to X_+^{j+1}$ is strictly complexity-reducing and is realized by either an exchange move or a destabilization, so that $b(X_+^{j+1}) \le b(X_+^j)$. Replacing X_+ with X'_+ and X_- with X'_-, there is an additional sequence which modifies X'_+, keeping X'_- fixed:

$$X'_+ = X^q \to \cdots \to X^r = X'_-$$

Each passage $X^j \to X^{j+1}$ in this sequence is also strictly complexity-reducing. It is realized by one of the moves defined by a template $\mathcal{T}$ in the finite set $\mathcal{T}(m)$, where $m = b(X_+)$. The inequality $b(X^{j+1}) \le b(X^j)$ holds for each $j = q, \ldots, r-1$ and so also for each $j = 1, \ldots, r-1$.

The proof of Theorem 6 uses the braid foliation techniques that were used in the proof of Theorem 5, but in a more complicated setting. Instead of looking at a foliated embedded disc which is bounded by a given unknotted closed braid, we are given two closed braids, X_+ and X_-, and an isotopy that takes X_+ to X_-. The trace of the isotopy sweeps out an annulus, but in general it is not embedded. The proof begins by showing that the given isotopy can be split into two parts, over which we have some control. An intermediate link X_0 which represents the same link type K as X_+ and X_- is constructed, such that the trace of the isotopy from X_+ to X_0 is an embedded annulus $\mathcal{A}_+$. Also the trace of the isotopy from X_0 to X_- is a second embedded annulus $\mathcal{A}_-$. The union of these two embedded annuli $\mathcal{TA} = \mathcal{A}_+ \cup \mathcal{A}_-$ is an immersed annulus, but its self-intersection set is controlled, and is a finite number of clasp arcs. The main tool in the proof of Theorem 6 is the study of the braid foliation of the immersed annulus $\mathcal{TA}$. In the next section, i.e. §3.3, we will see how Theorem 6 was used to settle a long-standing problem about contact structures.

3.3 Braids and contact structures

In this section we describe how Theorem 6 was used in [31] to settle a problem about contact structures on $\mathbb{R}^3$ and S^3.

Let $\mathbf{A}$ be the z-axis in $\mathbb{R}^3$, with standard cylindrical coordinates (ρ, θ, z) and let $\mathbf{H}$ be the collection of all half-planes H_θ through $\mathbf{A}$. The pair $(\mathbf{A}, \mathbf{H})$ defines the standard braid structure on $\mathbb{R}^3$. Using the same cylindrical coordinates, let α be the 1-form $\alpha = \rho^2 d\theta + dz$. The kernel ξ of the 1-form α defines a contact structure on $\mathbb{R}^3$. We can visualize ξ by imagining that there is a 2-plane (spanned by $\partial/\partial x$ and $\partial/\partial \rho$) attached to every point in $\mathbb{R}^3$. Figure 21 shows both the braid structure and the polar contact structure, for comparison. The family of 2-planes that define ξ twist (to the left) as one moves along the x-axis from 0 to ∞. The family is invariant under rotation of 3-space about the z-axis and under translation of 3-space along rays parallel to the z-axis. Its salient feature is that it is totally non-integrable, that is there is no surface in $\mathbb{R}^3$ which is everywhere tangent to the 2-planes of (ξ) in any neighborhood of any point in $\mathbb{R}^3$. (Of course this makes it hard to visualize). The twisting is generic in the sense that, if p is a point in a contact 3-manifold M^3, then in every neighborhood of p in M^3 the contact structure is locally like the one we depicted in Figure 21.

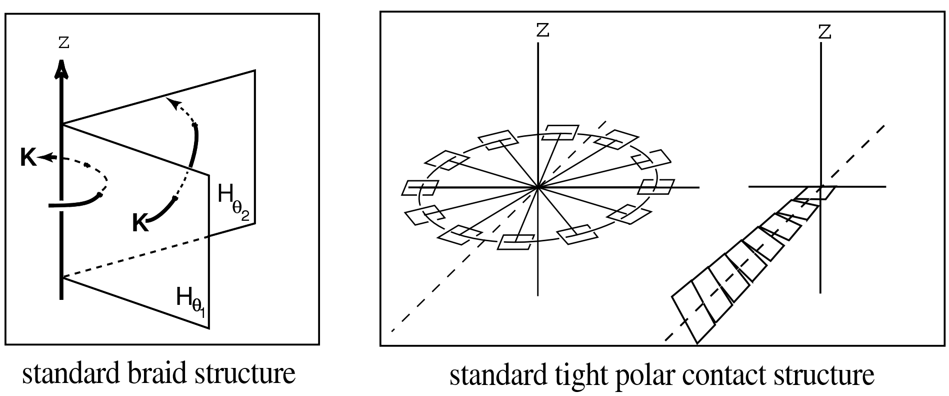


Figure 21: The standard braid structure and tight contact structure on $\mathbb{R}^3$

Let K be a knot (for simplicity we restrict to knots here, but everything works equally well for links) which is parametrized by cylindrical coordinates $(\rho(t), \theta(t), z(t))$, where $t \in [0, 2\pi]$. Then, as defined in §2.1, K is a closed braid if $\rho(t) > 0$ and $d\theta/dt > 0$ for all t. On the other hand, K is a Legendrian (resp. transversal) knot if it is everywhere (resp. nowhere) tangent to the 2-planes of ξ. In the Legendrian case this means that on K we have $d\theta/dt = (-1/\rho^2)(dz/dt)$. In the transversal case we require that $d\theta/dt > (-1/\rho^2)(dz/dt)$ at every point of $K(t)$. We are interested here primarily in the transversal case.

The total twist of the contact structure is the number of multiples of $\pi/2$ as one traverses the positive real axis from the origin to ∞. The case when the total twist angle is π is known as the standard (polar) contact structure. We call it ξ_π. While the braid structure is very different from the standard polar contact structure near $\mathbf{A}$, for large values of ρ the 2-planes in ξ_π are very close to the half-planes H_θ of the braid structure. In fact, assume that X is a closed braid which represents a knot K, and that X bounds a Seifert surface of minimum genus which supports a foliation as in §3.1. Assume further that the closed braid X intersects every 2-plane in the contact structure transversally. Then, in the complement of a tubular neighborhood of the braid axis, from the point of view of a topologist, the braid foliation and the foliation induced by the contact structure will be 'the same'.

Recall that Theorem 2 (Alexander's Theorem) was first proved in 1925. Sixty years later, Bennequin adapted Alexander's original proof (which is different from the proof given in this article) to the setting of transversal knots in [9], where he showed that every transversal knot is isotopic, through transversal knots, to a closed braid. In 2002 Orevkov and Shevchishin extended Bennequin's ideas and proved a version of Theorem 4 (Markov's Theorem) which holds in the transversal setting:

Theorem 7 [116]: *Let TX_+, TX_- be closed braid representatives of the same oriented link type K in oriented 3-space. Then there exists a sequence of closed braid representatives of $\mathcal{TK}$:*

$$TX_+ = TX_1 \to TX_2 \to \cdots \to TX_r = TX_1$$

such that, up to braid isotopy, each TX_{i+1} is obtained from TX_i by a single positive stabilization or destabilization.

Is the Transverse Markov Theorem really different from the Markov Theorem? Are there transversal knots which are isotopic as topological knots but are not transversally isotopic? To answer this question we take a small detour and review the contributions of Bennequin in [9].

Why did topologists begin to think about contact structures, and analysts begin to think about knots? While we might wish that analysts suddenly became overwhelmed with the beauty of knots, there was a more specific and focused reason. At the time that Bennequin did his foundational work in [9] it was not known whether a 3-manifold could support more than one isotopy class of contact structures. Bennequin answered this question in the affirmative, in the case of contact structures on $\mathbb{R}^3$ or S^3 which were known to be homotopic to the standard one. His tool for answering it was highly original, and it had to do with braids and knots.

Let TK be a transversal knot. By the above, we can assume without loss of generality that it is a closed braid. Let $\mathcal{TK}$ be its transversal knot type, i.e., its knot type under transversal isotopies, and let $[\mathcal{TK}]_{top}$ be its topological knot type. Choose a representative TX of $\mathcal{TK}$, which (by Bennequin's transversal version of Alexander's theorem) is always possible. Choose a Seifert surface F of minimal genus, with $TX = \partial F$. Bennequin studied the foliation of F which is induced by the intersections of F with the plane field determined by ξ_π. Let $n(TX)$ be the braid index and let $e(TX)$ be the algebraic crossing number of a projection of TX onto the plane $z = 0$. Both can be determined from the foliation. Bennequin found an invariant of $\mathcal{TK}$, given by the formula $\beta(\mathcal{TK}) = e(TX) - n(TX)$. Of course if he had known Theorem 7 the proof that $\beta(\mathcal{TK})$ is an invariant of $\mathcal{TK}$ would have been trivial, but he did not have that tool. He then showed a little bit more: he showed that $\beta(\mathcal{TK})$ is bounded above by the negative of the Euler characteristic of F in ξ_π. He then showed that this bound fails in one of the contact structures $\xi_{>\pi}$. In this way he proved that the contact structures $\xi_{>\pi}$ cannot be isotopic to ξ_{π}.

To knot theorists, Bennequin's proof should seem intuitively natural, because the invariant $\beta(\mathcal{TK})$ is a self-linking number of a representative $TX \in \mathcal{TK}$ (the sense of push-off being determined by ξ), and the more twisting there is the higher this number can be. For an explanation of the self-linking, and lots more about Legendrian and transversal knots we refer the reader to John Etnyre's excellent review article [63]. The basic idea is that TX bounds a Seifert surface, and this Seifert surface is foliated by the plane field associated to ξ. Call this foliation the <u>characteristic foliation</u>. Near the boundary, the characteristic foliation is transverse to the boundary. The Bennequin invariant is the linking number of TX, TX', where TX' is is a copy of TX, obtained by pushing TX off itself onto F, using the direction determined by the characteristic foliation of F.

Bennequin's paper was truly important. Shortly after it was written Eliashberg showed in [59] that the phenomenon of an infinite sequence of contact structures related to a single one of minimal twist angle occured generically in every 3-manifold, and introduced the term 'tight' and 'overtwisted' to distinguish the two cases. Here too, there is a reason that will seem natural to topologists. In 2003 Giroux proved [79] that every contact structure on every closed, orientable 3-manifold M^3 can be obtained in the following way: represent M^3 as a branched covering space of S^3, branched over a knot or link, and lift the standard and overtwisted contact structures on S^3 to M^3.

Returning to knot theory, the invariant $\beta(\mathcal{TK})$ allows us to answer a fundamental question: is the equivalence relation on knots that is defined by transversal isotopy really different from the equivalence relation defined by topological isotopy? The Bennequin invariant will be used to answer this question in the affirmative:

Theorem 8 [9] *There are infinitely many distinct transversal knot types associated to each topological knot type.*

Proof: Choose a transversal knot type $\mathcal{TK}$ and a closed braid representative TX_0. Stabilizing the closed braid TX_0 once negatively (recall the definition of positive/negative stabilizations given in §2.3, Figure 7), we obtain the transverse closed braid TX_1, with $e(TX_1) = e(TX_0) - 1$ and $n(TX_1) = n(TX_0) + 1$, so that $\beta(TX_1) = \beta(TX_0) - 2$. Iterating, we obtain transverse closed braids $TX_2, TX_3, \ldots$, defining transverse knot types $\mathcal{TK}_1, \mathcal{TK}_2, \mathcal{TK}_3, \cdots$, and no two have the same Bennequin invariant. Since stabilization does not change the topological knot type, the assertion follows. $\|$

At this writing, it is an open problem to find computable invariants of $\mathcal{TK}$ which are not determined by $[\mathcal{TK}]_{top}$ and $\beta(\mathcal{TK})$. A hint that the problem might turn out to be quite subtle was in the paper [72] by Fuchs and Tabachnikov, who proved that while ragbags filled with polynomial and finite type invariants of transversal knot types $\mathcal{TK}$ exist, based upon the work of Arnold in [5], they are all determined by $[\mathcal{TK}]_{top}$ and $\beta(\mathcal{TK})$. Thus, the seemingly new invariants that many people had discovered by using Arnold's ideas were just a fancy way of encoding $[\mathcal{TK}]_{top}$ and $\beta(\mathcal{TK})$.

This leads naturally to a question: Are there computable invariants of transversal knots which are *not* determined by $[\mathcal{TK}]_{top}$ and $\beta(\mathcal{TK})$? A similar question arises in the setting of Legendrian knots. Each Legendrian knot $\mathcal{LK}$ determines a topological knot type $[\mathcal{LK}]_{top}$, and just as in the transverse case it is an invariant of the Legendrian knot type. There are also two numerical invariants of $\mathcal{LK}$: the Thurston-Bennequin invariant $tb(\mathcal{LK})$ (a self-linking number) and the Maslov index $M(\mathcal{LK})$ (a rotation number). So until a few years ago the same question existed in the Legendrian setting, but the Legendrian case has recently been settled by Yuri Chekanov:

Theorem 9 [45] *There exist distinct Legendian knot types which have the same topological knot type*$[\mathcal{LK}]_{top}$*, and also the same Thurston-Bennequin invariant* $tb(\mathcal{LK})$ *and Maslov index* $M(\mathcal{LK})$*.*

The analogous result for transversal knots proved to be quite difficult, so to begin to understand whether something could be done via braid theory, the first author and Nancy Wrinkle asked an easier question, which they answered in part in [34]: are there transversal knot types which *are determined* by their topological knot type and Bennequin number? This question lead to a definition: a transversal knot type $\mathcal{TK}$ is transversally simple if it is determined by $[\mathcal{TK}]_{top}$ and $\beta(\mathcal{TK})$. So the question is: are there transversally simple knots? The manuscript [34] gives a purely topological (in fact braid-theoretic) criterion which enables one to answer the question affirmatively, adding one more piece of evidence that topology and analysis walk hand in hand. To explain it, recall the destabilization and exchange moves, and the two flypes, depicted in Figure 20. A topological knot or link type K is said to be exchange reducible if an arbitrary closed braid representative X of K can be changed to an arbitrary representative of minimum braid index by braid isotopy, positive and negative destabilizations and exchange moves. We have:

Theorem 10 [34] *If a knot type K is exchange-reducible, then any transversal knot type $\mathcal{TK}$ which has $[\mathcal{TK}]_{top} = K$ is transversally simple.*

This theorem was used to give a new proof of a theorem of Eliashberg [60], which asserts that the unlink is transversally simple, and also (with the help of [105]) to prove the then-new result that most iterated torus knots are transversally simple.

The rest of this section will be directed at explaining the main result of [31]:

Theorem 11 [31] *There exist transversal knot types which are not transversally simple. Explicitly, the transverse closed 3-braids $TX_+ = \sigma_1^5\sigma_2^4\sigma_1^6\sigma_2^{-1}$ and $TX_- = \sigma_1^5\sigma_2^{-1}\sigma_1^6\sigma_2^4$ determine transverse knot types $\mathcal{TK}_+, \mathcal{TK}_-$ with $(\mathcal{TK}_+)_{top} = (\mathcal{TK}_-)_{top}$ and $\beta(\mathcal{TK}_+) = \beta(\mathcal{TK}_-)$, but $\mathcal{TK}_+ \neq \mathcal{TK}_-$.*

Sketch of the proof of Theorem 11. See [31] for all details. The examples in Theorem 11 were obtained by choosing all the weights in the negative flype template to be 1, and assigning explicit 2-braids to the blocks P, Q, R of the negative flype template of Figure 20. If the weights are all chosen to be 1, the blocks P, Q, R are 2-braids and (except in very special cases) the flype templates have braid index 3. First one must show that the examples satisfy the conditions of the theorem. The topological knot types defined by the closed 3-braids $\sigma_1^5\sigma_2^4\sigma_1^6\sigma_2^{-1}$ and $\sigma_1^5\sigma_2^{-1}\sigma_1^6\sigma_2^4$ coincide because they are carried by the block strand diagrams for the negative flype template of Figure 20. The Bennequin invariant can be computed as the exponent sum of the braid word (14 in both cases) minus the braid index (3 in both cases). So the examples have the required properties.

The hard part is the establishment of a special version of Theorem 6 which is applicable to the situation. Its special features are as follows:

1. Both X_+ and X_- have braid index 3.

2. Since it is well known that exchange moves can be replaced by braid isotopy for 3-braids, the first two sequences in Theorem 6 are vacuous, i.e. $X_\pm = X'_\pm$.

3. Because of the special assumption just noted, the templates that are needed, in the topological setting, can be enumerated explicitly: they are the positive and negative destabilization and the positive and negative flype templates. No others are needed.

It is proved in [31] that if X_- and X_+ are transversal closed braids TX_+ and TX_-, then the isotopy that takes TX_+ to TX_- may be assumed to be transversal. So, suppose that a transversal isotopy exists from the transverse closed braid TX_+ to the transverse closed braid TX_-. Then there is a 3-braid template that carries the braids $\sigma_1^5\sigma_2^4\sigma_1^6\sigma_2^{-1}$ and $\sigma_1^5\sigma_2^{-1}\sigma_1^6\sigma_2^4$. This is the first key fact that we use from Theorem 6. Instead of having to consider *all* possible transversal isotopies from TX_+ to TX_-, we only need to consider those that relate the left and right block-strand diagrams in one of the four 3-braid templates. So the braids in question are carried either by one of the two destabilization templates or by one of the two flype templates. If it was one of the destabilization templates, then the knots in question could be represented by 2 or 1-braids, i.e. they would be type $(2,n)$ torus knots or the unknot, however an easy argument shows that the knots in Theorem 11 are neither type $(2,n)$ torus knots or the unknot. The positive flype templates are ruled out in different way: topologically, our closed braids admit a negative flype, so if they are also carried by the positive flype template they admit flypes of both signs. However, the manuscript [93] gives conditions under which a closed 3-braid admits flypes of both signs, and the examples were chosen explicitly to rule out that possibility.

We are reduced to isotopies that are supported by the negative flype template. It is straightforward to show that the obvious isotopy is not transversal, but maybe there is some *other* isotopy which is transversal. Here a key fact about the definition of a template is used (and this is a second very strong aspect of the MTWS). If such a transversal isotopy exists, then it exists for every knot or link defined by a fixed choice of braiding assignments to the blocks. Choose the braiding assignments $\sigma_1^3, \sigma_2^4, \sigma_1^{-5}$ to the blocks P, R, Q. This braiding assignment gives a 2-component link $L_1 \sqcup L_2$ which has two distinct isotopy classes of closed 3-braid representatives. If L_1 is the component associated to the left strand entering the block P, then $\beta(L_1) = -1$ and $\beta(L_2) = -3$ before the flype, but after the flype the representative will be $\sigma_1^3\sigma_2^{-1}\sigma_1^{-5}\sigma_2^4$, with $\beta(L_1) = -3$ and $\beta(L_2) = -1$. However, by Eliashberg's isotopy extension theorem (Proposition 2.1.2 of [60]) a transversal isotopy of a knot/link extends to an ambient transversal isotopy of the 3-sphere. Any transversal isotopy of $L_1 \sqcup L_2$ must preserve the β-invariants of the components. It follows that no such transversal isotopy exists, a contradiction of our assumption that TX_+ and TX_- are transversally isotopic. $\|$

Other examples of a similar nature were discovered by Etnyre and Honda [64] after the proof of Theorem 11 was posted on the arXiv. Their methods are very different from the proof that we just described (being based on contact theory techniques rather than topological techniques), but are equally indirect. They do not produce explicit examples, rather they present a bag of pairs of transverse knots and prove that at least one pair in the bag exists with the properties given by Theorem 11. Therefore we pose, as an important open problem:

Open Problem 4 Find new computable invariants of transversal knot types. Here 'new' means an invariant which is not determined by the topological knot type $\mathcal{TX}$ and and the Bennequin invariant $\beta(\mathcal{TX})$. 'Computable' means that it should be computable from either a closed braid diagram or some other representation of the transversal knot. Braid groups seem to be a natural setting for investigating this problem. ♣

Remark 3.1 The connections between closed braids and contact structures does not end with transversal knots. There are fundamental relationships between open book structures on 3-manifolds and contact structures on those manifolds, both untwisted and twisted. See [79] for an introduction to this interesting new area, and see [63] for a review of the mathematics and a discussion of many open problems waiting to be investigated.

4 Representations of the braid groups

Before the discovery of Hecke algbra representations of the braid group (discussed in this section) very little was known about finite dimensional but infinite representations of $\mathbf{B}_n$, except for the ubiquitous Burau representation. That matter changed dramatically in 1987 with the publication of [89]. Suddenly, we had more knot invariants and with them more braid group representations than anyone could deal with, and the issue became one of organizing them. However, we shall not attempt to give a comprehensive overview of the rich theory of representations of braid groups in this section. Instead, we focus here on the representations of $\mathbf{B}_n$ which have played the greatest roles in the development of that theory: the Burau representation, the Hecke algebra representations, and, more recently, the Lawrence-Krammer representation.

4.1 A brief look at representations of Σ_n

The fact that the representation theory of $\mathbf{B}_n$ is rooted in the representation theory of the symmetric group Σ_n is a consequence of the surjection $\mathbf{B}_n \to \Sigma_n$, given by mapping the elementary braid σ_i to the transposition $s_i = (i, i+1)$. While the kernel of this homomorphism is very big (it's the entire pure braid group $\mathbf{P}_n$), it nevertheless turns out that a great deal can be learned about representations of $\mathbf{B}_n$ by studying the collection of irreducible representations of Σ_n, and attempting to lift them to reprentations of $\mathbf{B}_n$ by 'deforming' them, and hoping that you get something new. For the record, we now note that the group algebra $\mathbb{C}\Sigma_n$ has generators $1, s_1, \ldots, s_{n-1}$ and defining relations:

$$s_i s_k = s_k s_i \text{ if } |i-k| \geq 2, \quad s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1}, \; s_i^2 = 1, \tag{13}$$

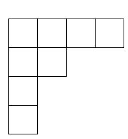
where $1 \leq i \neq k \leq n-1$. As a vector space, $\mathbb{C}\Sigma_n$ is spanned by $n!$ reduced words [35] in the transpositions s_i:

$$\{(s_{i_1} s_{i_1-1} \ldots s_{i_1-k_1})(s_{i_2} s_{i_2-1} \ldots s_{i_2-k_2}) \cdots (s_{i_r} x_{i_r-1} \ldots s_{i_r-k_r})\} \tag{14}$$

where $1 \leq i_1 < i_2 < \cdots < i_r \leq n-1$ and $i_j - k_j \geq 1$.

Irreducible representations of Σ_n are parametrized by Young diagrams, which in turn are parametrized by partitions of n. (See [73] for one of the many good discussions of this subject in the literature.) A Young diagram consists of stacked rows of boxes, aligned on the left, with the number of boxes in each row strictly nonincreasing as you go from top to bottom. The number of boxes in a row corresponds to a term in a given partition of n. For example, the Young diagram shown in Figure 22 corresponds to the partition $8 = 4+2+1+1$.

Figure 22: The Young diagram corresponding to the partition $4+2+1+1$.



As it turns out, the irreducible representations of the symmetric group are in 1-1 correspondence with Young diagrams, or equivalently with partitions of n. There is a row-column symmetry, and we adopt the convention that the Young diagram which has one row of n boxes, which comes from the partition $n = n$, corresponds to the trivial representation. Then the Young diagram which has n rows consisting of 1 box per row, which comes from the partition $n = 1 + 1 + \cdots + 1$, corresponds to the parity representation, mapping each element of Σ_n to its sign. The first interesting representation of a braid group occurs when $n = 3$, corresponding to the partition $3 = 2 + 1$. More generally, we consider the partition $n = (n-1) + 1$, or the Young diagram with two rows, the top row having $(n-1)$ boxes and the bottom having a single box. This Young diagram corresponds to the standard representation $\Sigma_n \to GL_n(\mathbb{C})$, given by the usual action of Σ_n on $\mathbb{C}^n$ by permuting basis vectors. Thus if s_i denotes the transposition $(i \;\; i+1) \in \Sigma_n$, the standard representation sends s_i to the following:

$$I_{i-1} \oplus \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \oplus I_{n-i-1}$$

where I_k denotes the $k \times k$ identity matrix.

The standard representation is reducible; it is easy to see that it fixes a 1-dimensional subspace, namely, the span of the sum of the basis vectors. The complementary $(n-1)$-dimensional subspace is the set of all points $(z_1, \ldots, z_n) \in \mathbb{C}^n$ such that $z_1 + \ldots + z_n = 0$. This subspace corresponds to the irreducible representation given by the Young diagram in question. One can use the hook length formula to calculate directly that the dimension of the representation corresponding to this Young diagram is indeed $n-1$ (we refer the reader again to [73] for an explanation of this formula).

4.2 The Burau representation and polynomial invariants of knots.

Burau first introduced his representation of the braid group in 1936 [42]. Much later, it was realized that it could be thought of as a deformation of the standard representation of Σ_n corresponding to the partition $n = (n-1) + 1$. For many years it was the focus of the representation theory of braid groups. We define the Burau representation $\rho : \mathbf{B}_n \to GL_n(\mathbb{Z}[t, t^{-1}])$ as follows:

$$\sigma_i \mapsto I_{i-1} \oplus \begin{pmatrix} 1-t & t \\ 1 & 0 \end{pmatrix} \oplus I_{n-i-1}$$

Note that substituting $t = 1$ gives back the representation (of $\mathbf{B}_n$ factoring through Σ_n), and this is why we say that it is a deformation of the standard representation of Σ_n. Like the representation of Σ_n, the Burau representation splits into a 1-dimensional representation and an $(n-1)$-dimensional irreducible representation known as the reduced Burau representation which we denote by $\bar{\rho} : \mathbf{B}_n \to GL_n(\mathbb{Z}[t, t^{-1}])$ as follows:

$$\sigma_i \mapsto I_{i-2} \oplus \begin{pmatrix} 1 & -t & 0 \\ 0 & -t & 0 \\ 0 & -1 & 1 \end{pmatrix} \oplus I_{n-i-2}$$

where the $-t$ in the middle of the 3×3 matrix is always in the $(i,i)^{th}$ spot.

It has been known for a long time that the Burau representation is faithful for $n \leq 3$ (see [18], for example), and for many years the representation was held to be a reasonable candidate for a faithful linear representation of $\mathbf{B}_n$ for all n. However, in 1991, Moody showed that for $n \geq 9$, the Burau representation is *not* faithful [107]. Long and Paton later improved Moody's result to $n \geq 6$ [103]. Bigelow further improved this to $n \geq 5$ [13]. At the time of this writing, the case $n = 4$ remains open.

Despite such results, the Burau representation continues to play an important role in the study of representations of $\mathbf{B}_n$. It was known classically that the Alexander polynomial $\Delta_K(t)$ of a knot or link K can be calculated directly from the image under the reduced Burau representation of a braid X such that $b(X)$ represents K, as follows (see [18], e.g., for a proof based upon Theorem 4, the Markov Theorem):

$$\Delta_{b(X)}(t) = \frac{\det(\bar{\rho}(X) - I_{n-1})}{1 + t + \cdots + t^{n-1}}. \tag{15}$$

Thus the Alexander polynomial of the closed braid associated to the open braid X, i.e. $\Delta_{b(X)}(t)$, is a rescaling of the characteristic polynomial of the the image of X in the reduced representation. In what follows, we shall see how a property of the Burau representation motivated the definition of a variant on the two-variable HOMFLY polynomial [75], a knot invariant of which both the Alexander polynomial and the Jones polynomial are specializations. Further, in §4.4 we shall give a topological interpretation of the Burau representation which naturally leads to the definition of a faithful linear representation of $\mathbf{B}_n$ known as the Lawrence-Krammer representation, which will be our focus in §4.5.

4.3 Hecke algebras representations of braid groups and polynomial invariants of knots

A simple calculation, together with the Cayley-Hamilton theorem, shows that the image of each of our braid group generators under the Burau representation, $\rho(\sigma_i)$, satisfies the characteristic equation $x^2 = (1-t)x + t$ and thus has two distinct eigenvalues. This prompted Jones to study all representations $\rho : \mathbf{B}_n \to GL_n(\mathbb{C})$ which have at most two distinct eigenvalues [89]. Let $x_i = \rho(\sigma_i)$. Then for all i, x_i must satisfy a quadratic equation of the form $x_i^2 + ax_i + b = 0$. By rescaling, we may assume that one of the eigenvalues is 1 and eliminate one of the variables, e.g., $a = -(1+b)$. Note that by rewriting our quadratic equation and making the substitution $b = -t$ we regain the characteristic equation from the Burau representation. However, the convention in the literature seems to be to rescale our representation by (-1) so that the equation takes the form $x_i^2 = (t-1)x_i + t$. With this motivation, we define the <u>Hecke algebra</u> $H_n(t)$ to be the algebra with generators $1, x_1, \ldots, x_{n-1}$ and defining relations as follows:

$$x_i x_k = x_k x_i \quad \text{if} \quad |i-k| \geq 2, \quad x_i x_{i+1} x_i = x_{i+1} x_i x_{i+1}, \quad x_i^2 = (t-1)x_i + t, \tag{16}$$

where $1 \leq i \neq k \leq n-1$. Comparing the relations in (13) and (16), we see that $H_n(1) \cong \mathbb{C}\Sigma_n$, the group algebra of the symmetric group. Hence we can think of $H_n(t)$ as a 'deformation' of $\mathbb{C}\Sigma_n$.

The connection between $H_n(t)$ and $\mathbb{C}\Sigma_n$ is made even more transparent by noting that as a vector space, $H_n(t)$ is spanned by $n!$ lifts of a system of reduced words in the transpositions

$s_i \in \Sigma_n$. For example, we can take as a spanning set

$$\{(x_{i_1}x_{i_1-1}\dots x_{i_1-k_1})(x_{i_2}x_{i_2-1}\dots x_{i_2-k_2})\cdots(x_{i_r}x_{i_r-1}\dots x_{i_r-k_r})\} \tag{17}$$

where $1 \le i_1 < i_2 < \cdots < i_r \le n-1$ and $i_j - k_j \ge 1$ [35], [89].

Remark 4.1 In §1.4.2 we defined an algebra $J_n(t)$ with generators $1, g_1, \dots, g_{n-1}$ and defining relations (7), the Jones algebra. As it turns out, its irreducible summands are in 1-1 correspondence with the irreducible representations of the Hecke algebra that are parametrized by Young diagrams with exactly 2 rows. The Hecke algebra, as defined above, has generators maps $1, x_1, \dots, x_{n-1}$ with defining relations (16), so that the map $\xi : H_n(t) \to J_n(t)$ that is defined by $\xi(x_i) = g_i$ is a homorphism of algebras.

Our main purpose in this section is to outline Jones' development in [89] of a two-variable polynomial knot invariant arising from representations of the Hecke algebras $H_n(t)$. This polynomial is essentially the well-known HOMFLY polynomial, and includes the Jones polynomial as a specialization. We have just seen that the Jones algebra is the quotient of the Hecke algebra $H_n(t)$ by one extra relation (as an aside, we note that this extra relation is satisfied by the image of the transpositions generators $s_i = (i, i+1)$ in Σ_n under all representations arising from Young diagrams with at most two rows.) To pick up a second variable, we introduce an additional parameter by allowing a 1-parameter family of traces on Hecke algebras. We now pursue this point of view, and we also refer the reader to [85] for another exposition which follows the same point of view.

We begin by defining a function $f : \mathbf{B}_n \to H_n(t)$ by $f(\sigma_i) = x_i$. The function f is well-defined on reduced words in the generators σ_i and commutes with the natural inclusions $\mathbf{B}_{n-1} \subset \mathbf{B}_n$ and $H_{n-1}(t) \subset H_n(t)$, although in general f fails to be a homomorphism. We can then apply the following result due to Adrian Ocneanu which appeared in [75] and was proved inductively in [89] using the $n!$-element basis given above.

Theorem 12 [75], [89] *For each $z \in \mathbb{C}^*$ (and each $t \in \mathbb{C}^*$), there exists a unique trace function* $\mathrm{tr} : \cup_{n=1}^{\infty} H_n(t) \to \mathbb{C}$ *such that*

1. $\mathrm{tr}(1) = 1$
2. $\mathrm{tr}(ab) = \mathrm{tr}(ba)$
3. tr *is $\mathbb{C}$-linear*
4. $\mathrm{tr}(ux_{n-1}v) = z\ \mathrm{tr}(uv)$ *for all $u, v \in H_{n-1}(t)$.*

Theorem 12 gives us a one-parameter family of trace functions on a one-parameter family of algebras. In fact, using the properties of the trace function given in theorem it is possible to compute $\mathrm{tr}(f(X))$ for all $X \in \mathbf{B}_n$. (We note the fact that for any $w \in H_n(t)$ such that $w \notin H_{n-1}(t)$, there is a unique reduced word $w = x_{i_1} \cdots x_{i_r}$ in which x_{n-1} appears exactly once [89].) In practice, the third relation of $H_n(t)$ is quite useful for computing the trace function tr, both in its original form and in the following:

$$x_i^{-1} = t^{-1}x_i + (t^{-1} - 1).$$

Example 4.1 Let $X_1 = \sigma_1^3 \in \mathbf{B}_2$, and let $X_2 = \sigma_1\sigma_2^{-1}\sigma_1\sigma_2^{-1} \in \mathbf{B}_3$. Note that $\hat{X_1}$ is the right-handed trefoil knot and that $\hat{X_2}$ is the Figure-8 knot. We invite the reader to check that

$$\mathrm{tr}(f(X_1)) = (t^2 - t + 1)z + t(t-1)$$

and that

$$\mathrm{tr}(f(X_2)) = (3 - t^{-1} - t)t^{-1}z^2 + (3 - t^{-1} - t)(t^{-1} - 1)z - (2 - t^{-1} - t).\spadesuit$$

We also note that the second property of the trace function given in Theorem 12 implies that $\mathrm{tr} \circ f$ is invariant on conjugacy classes in $\mathbf{B}_n$. It remains to tweak the function a bit in order to obtain from a given braid a two-variable polynomial which is also invariant under stabilization and destabilization moves as defined in §2.3; such a polynomial will be Markov-invariant and hence an invariant of the knot type of the closed braid.

Algebraically, stabilization and destabilization each take the form $X \to X\sigma_n^{\pm 1}$, the only difference being appropriate conditions on the braid X. We would like to rescale our representation f in such a way that both versions of stabilization (resp. destabilization) have the same effect on the trace function. Suppose there exists a complex number k such that $\mathrm{tr}(kx_i) = \mathrm{tr}((kx_i)^{-1}))$. Then we can find a 'formula' for k as follows:

$$\begin{aligned} k^2\mathrm{tr}(x_i) &= \mathrm{tr}(x_i^{-1}) \\ k^2 z &= \mathrm{tr}(t^{-1}x_i + t^{-1} - 1) \\ k^2 &= \frac{t^{-1}z + t^{-1} - 1}{z} \\ k^2 &= \frac{1 + z - t}{tz} \end{aligned}$$

Solving this for z, we obtain

$$z = -\frac{1-t}{1-k^2 t}.$$

We set $\kappa = k^2$, and define $f_\kappa : \mathbf{B}_n \to H_n(t)$ by $f_\kappa(\sigma_i) = \sqrt{\kappa}\ \sigma_i$. Now we have

$$\begin{aligned} \mathrm{tr}(f_\kappa(\sigma_n)) &= \sqrt{\kappa}\ z \\ &= -\sqrt{\kappa}\frac{1-t}{1-\kappa t}. \end{aligned}$$

We would like to define a map $\mathbf{B}_n \to \mathbb{Z}[t^{\pm 1}, \kappa^{\pm 1}]$ which is Markov invariant. At the moment, we have that

$$\mathrm{tr}(f(w \cdot \sqrt{\kappa}\sigma_n)) = -\sqrt{\kappa}\frac{1-t}{1-\kappa t}\mathrm{tr}(f(w)) = \mathrm{tr}(f(w \cdot \frac{1}{\sqrt{\kappa}}\sigma_n^{-1}))$$

for any $w \in \mathbf{B}_n$. Now we simply define

$$\begin{aligned} F(X) = F_X(t,\kappa) &= (-\frac{1}{\sqrt{\kappa}} \cdot \frac{1-\kappa t}{1-t})^{n-1}\mathrm{tr}(f_\kappa(X)) \\ &= (-\frac{1}{\sqrt{\kappa}} \cdot \frac{1-\kappa t}{1-t})^{n-1}(\sqrt{\kappa})^E\mathrm{tr}(f(X)) \end{aligned}$$

for $X \in \mathbf{B}_n$, where E is the exponent sum of X as a word in $\sigma_1, \ldots, \sigma_{n-1}$. It is clear that $F(X)$ depends only on the knot type of $b(X)$. We now reparametrize one last time, setting

$$\begin{aligned} l &= \sqrt{\kappa}\sqrt{t} \\ m &= \sqrt{t} - \frac{1}{\sqrt{t}}. \end{aligned}$$

With this substitution, we obtain a Laurent polynomial in two variables l and m, which we denote $P_{b(X)}(l,m) = P_K(l,m)$, where K is the (oriented) knot or link type of $b(X)$. Furthermore, $P_K(l,m)$ satisfies the skein relation

$$mP_{K_0} = l^{-1}P_{K_+} - lP_{K_-} \tag{18}$$

where K_0, K_+, and K_- are oriented knots with identical diagrams except in a neighborhood of one crossing, where they have a diagram as given in Figure 23 (Proposition 6.2 of [89]). Thus by beginning with $P_U = 1$, where U denotes the unknot, it is possible to calculate P_K for any knot or link K using only the skein relation, which is often simpler than using the trace function.

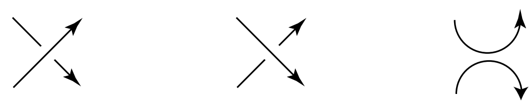


Figure 23: Crossings for K_+, K_-, and K_0, respectively, in the skein relation given in (18).

Remark 4.2 We note that the polynomial $P_K(l,m)$ obtained in this way is essentially the same as the two-variable polynomial known as the HOMFLY polynomial ([75]) which is usually reparametrized as $P_K(il^{-1}, im)$.

Example 4.2 Let X_1, X_2 be the braids defined in Example 4.1 whose closures are $K_1 =$ the right-handed trefoil knot and $K_2 =$ the Figure-8 knot, respectively. We leave it as an exercise for the reader to check that

$$\begin{aligned} F_{X_1}(t,\kappa) &= \kappa(1 + t^2 - \kappa t^2) \\ &= \kappa t(t + t^{-1} - \kappa t) \\ &= \kappa t(2 - \kappa t + t + t^{-1} - 2) \\ &= \kappa t\left(2 - \kappa t + (\sqrt{t} - \frac{1}{\sqrt{t}})^2\right) \end{aligned}$$

and hence we have

$$P_{K_1}(l,m) = 2l^2 - l^4 + l^2m^2.$$

Similarly, the reader can check that

$$\begin{aligned} F_{X_2}(t,\kappa) &= \frac{1 - \kappa(1 - t + t^2) + \kappa^2 t^2}{t\kappa} \\ P_{K_2}(l,m) &= l^{-2} - m^2 - 1 + l^2. \end{aligned}$$

For explicit calculations of F_{X_2} and P_{K_2} using the trace function, see p. 350 of [89].♠

Remark 4.3 There is also a 1-variable knot polynomial, the Jones polynomial, associated to the algebra $J_n(t)$ generated by $1, g_1, \dots, g_{n-1}$ with defining relations (7). In the situation of the Jones algebra the trace is unique, whereas in the situation of the Hecke algebra, as we presented it here, there is a 1-parameter family of traces. The 1-variable Jones polynomial was discovered before the 2-variable HOMFLY polynomial.

This two-variable knot polynomial has been much studied and reviewed in the literature. For the sake of completeness we list here a few of its noteworthy properties and applications.

1. **Connect sums**: $P_{K_1 \# K_2} = P_{K_1} \cdot P_{K_2}$

2. **Disjoint unions**: $P_{K_1 \amalg K_2} = (\frac{l^{-1}-l}{m}) P_{K_1} \cdot P_{K_2}$.

3. **Orientation**: $P_{\bar{K}} = P_K$, where $\bar{K}$ denotes the link obtained by reversing the orientation of every component of the link K.

4. **Chirality**: $P_{\tilde{K}}(l, m) = P_K(l^{-1}, -m)$, where $\tilde{K}$ denotes the mirror image of the link K.

5. **Alexander polynomial**: Note that $F_X(1, \kappa)$ is not defined. It comes as something of surprise, then, that the specialization $l = 1, m = \sqrt{t} - \frac{1}{\sqrt{t}}$ gives the Alexander polynomial $\Delta_K(t)$. Jones shows how to avoid the singularity by exploiting an alternate method of calculating the trace function using weighted sums of traces (see [132] and [90] as well as [75] and [89]). A by-product of this alternate method is another derivation of Equation 15 showing how to calculate $\Delta_K(t)$ from the Burau representation.

6. **Jones polynomial**: The famous Jones polynomial can be obtained from the two-variable polynomial by setting
$$V_K(t) = P(t, \sqrt{t} - \frac{1}{\sqrt{t}}).$$
Note that we are abusing notation by reusing the variable t here and above in $\Delta_K(t)$.

7. **A lower bound for braid index** At the end of §2.2 we remarked that it is an open problem to determine the braid index of a knot algorithmically. However, the HOMFLY polynomial does give a remarkably useful lower bound, via a famous inequality which is known as the Morton-Franks-Williams inequality. It was proved simultaneously and independently by Hugh Morton in [111] and by John Franks and Robert Williams in [74]. While it proved to be sharp on all but 5 of the knots in the standard tables of knots having at most 10 crossings, there are also infinitely many knots on which it fails to be sharp. The first author came to a new appreciation of the importance of this problem when she was faced with the problem of determining, precisely, the braid index of certain 6-braid knots on which it failed. Consulting many people, it became apparent that there was essentially no other useful result on this basic question.

4.4 A topological interpretation of the Burau representation

We have seen that the Burau representation is deeply connected with the topology of closed braids. We shall now give a fully topological definition of the Burau representation. We outline here ideas given in [13]; for full details see that or Turaev's excellent survey article [130], among others.

Let D_n be an n-punctured disk, which we will think of as a disk D with n distinguished points $q_1, \ldots, q_n$. Choose a point $d_0 \in \partial D_n$ to serve as the basepoint. Then $\pi_1(D_n, d_0)$ is free on n generators which can be represented by loops x_i based at d_0 travelling counterclockwise about the puncture q_i for $i = 1, \ldots, n$. We define a surjective map $\epsilon : \pi_1(D_n, d_0) \to \mathbb{Z}$ as follows. Let $\gamma = x_{i_1}^{n_1} \cdots x_{i_r}^{n_r} \in \pi_1(D_n, d_0)$. Then we define the exponent sum $\epsilon(\gamma) = \sum_{i=1}^{r} n_i$. The integer $\epsilon(\gamma)$ can be interpreted as the total algebraic winding number of γ about the punctures $\{q_i\}$, i.e., the sum over all $i = 1, \ldots, n$ of the winding number of γ about q_i. Now there is a regular covering space $\tilde{D_n}$ of D_n corresponding to the kernel of the map ϵ. Since $\epsilon : \pi_1(D_n, d_0) \to \mathbb{Z}$ is surjective, the group of covering transformations $\mathrm{Aut}(\tilde{D_n}) \cong \mathbb{Z}$. Let τ be a generator of $\mathbb{Z}$, and let $\Lambda = \mathbb{Z}[t, t^{-1}]$. Then $H_1(\tilde{D_n})$ inherits a Λ-module structure from the action of the covering transformations: we simply set $t \cdot \gamma = \tau_*(\gamma)$, where $\gamma \in H_1(\tilde{D_n})$ and τ_* denotes the induced action on homology. We note that as a Λ-module, $H_1(\tilde{D_n})$ is free of rank $n - 1$.

In what follows it will be convenient to think of $\mathbf{B}_n$ as in §1.3, i.e., as the mapping class group of D_n where ∂D_n is fixed pointwise, while the punctures may be permuted. We shall abuse terminology by not distinguishing between a mapping class and a diffeomorphism which represents it. Therefore we think of any $X \in \mathbf{B}_n$ as a map $X : D_n \to D_n$. Then X lifts uniquely to a map $\tilde{X} : \tilde{D_n} \to \tilde{D_n}$ which fixes the fiber over the basepoint d_0 pointwise. Furthermore, $\tilde{X}$ induces a Λ-module automorphism $\tilde{X}_*$ of $H_1(\tilde{D_n})$. Since $H_1(\tilde{D_n})$ is a free Λ-module of rank $n-1$, we can now define a map $\mathbf{B}_n \to \mathrm{GL}_{n-1}(\Lambda)$ by $X \mapsto \tilde{X}_*$. This map turns out to be equivalent to the reduced Burau representation defined previously (see [68] for a classification of linear representations of the braid group of degree at most $n - 1$).

The main idea of Stephen Bigelow's proof of the non-faithfulness of the Burau representation in the case $n = 5$ is contained in the following theorem:

Theorem 13 [13] *For $n \geq 3$, the Burau representation $\bar{\rho} : \mathbf{B}_n \to \mathrm{GL}_{n-1}(\Lambda)$ is not faithful if and only if there exist arcs α, β embedded in D_n satisfying:*

1. *$\partial\alpha = \{q_1, q_2\}$ and $\partial\beta = \{d_0, q_3\}$ or $\{q_3, q_4\}$.*
2. *α intersects β nontrivially (more precisely, there exists no isotopy rel endpoints which carries α off β)*
3. *For some choice of lifts $\tilde{\alpha}, \tilde{\beta}$, we have $\sum_{k\in\mathbb{Z}}(t^k\tilde{\alpha}, \tilde{\beta})t^k = 0$, where (x, y) denotes the algebraic intersection number of two (oriented) arcs in $\tilde{D_n}$.*

We note that the case $\partial\beta = \{d_0, q_3\}$ follows from Theorem 1.5 of [103]. Bigelow has produced an explicit example of arcs α and β satisfying the criteria of Theorem 13 in the case $n = 5$ (see p. 402 of [13]). It follows that the Burau representation of $\mathbf{B}_n$ is not faithful for $n \geq 5$. It has been known for many years that it *is* faithful for $n = 3$.

Open Problem 5 Is the Burau representation of $\mathbf{B}_4$ faithful? ♣

In the next section, we will explore another representation of $\mathbf{B}_n$, with a topological interpretation analogous to that of the Burau representation.

4.5 The Lawrence-Krammer representation

In 1990, Ruth Lawrence introduced a family of representations of $\mathbf{B}_n$ corresponding to Young diagrams with two rows which arise out of a topological construction of representations of the Hecke algebras $H_n(t)$ [100]. Later, Krammer gave an entirely algebraic definition of one of these representations:

$$\lambda : \mathbf{B}_n \to \mathrm{GL}_r(A)$$

where $A = \mathbb{Z}[t^{\pm 1}, q^{\pm 1}]$ and showed that it was faithful for $n = 4$ [97]. The representation λ has become known as the Lawrence-Krammer representation. Shortly after Krammer's result appeared, Bigelow was able to use topological methods to show that λ is faithful for all n [14]. Krammer later gave an algebraic proof of the same result [98]. Therefore we now have:

Theorem 14 [14], [98] *The map λ is a faithful representation of $\mathbf{B}_n$, and hence $\mathbf{B}_n$ is a linear group for all n.*

We observe that Krammer's result in [97] also implies that $\mathrm{Aut}(F_2)$, the group of automorphisms of a free group of rank 2, is linear.

We shall not present here a full proof of Theorem 14; our aim is to define the representation λ, to present the key ingredients in Bigelow's proof: 'forks', 'noodles', and the pairing between them, and to describe Bigelow's characterization of the kernel of λ in terms of these ingredients.

As mentioned previously, Bigelow's topological definition of λ is somewhat analogous to the topological definition of the Burau representation given in the previous section. This time, we begin the construction with a certain configuration space of a punctured disk rather than the punctured disk itself. In the notation of §1.1, we let $C = \mathcal{C}_{0,2}(D_n)$, where D_n is the n-punctured disk described in the previous section. Elements of C will be denoted $\{z_1, z_2\} = \{z_2, z_1\}$, where $z_1 \neq z_2 \in D_n$. Now we choose two distinct points $d_1, d_2 \in \partial D_n$, and let $c_0 = \{d_1, d_2\}$ be the basepoint of C. We are now in a position to define a map $\phi : \pi_1(C, c_0) \to \mathbb{Z} \times \mathbb{Z}$ as follows. Let $[\gamma] \in \pi_1(C, c_0)$. Then the loop $\gamma(s)$ in C can be expressed as $\gamma(s) = \{\gamma_1(s), \gamma_2(s)\}$, where γ_i is an arc in D_n for $i = 1, 2$.

We define two integers a, b as follows.

$$\begin{aligned} a &= \frac{1}{2\pi i} \sum_{j=1}^{n} \int_{\gamma_1} \frac{dz}{z - q_j} + \int_{\gamma_2} \frac{dz}{z - q_j} \\ b &= \frac{1}{\pi i} \int_{\gamma_1 - \gamma_2} \frac{dz}{z} \end{aligned}$$

Since γ is a loop in C, we either have that both γ_1 and γ_2 are both loops, i.e., $\gamma_1(0) = \gamma_1(1)$ and $\gamma_2(0) = \gamma_2(1)$, or else, as Bigelow puts it, the γ_i are arcs which 'switch places', i.e., $\gamma_1(1) = \gamma_2(0)$ and $\gamma_2(1) = \gamma_1(0)$. In the first case, Bigelow observes that we can interpret a as the sum of the total algebraic winding numbers of γ_1 and γ_2 about the punctures q_i

and b as twice the winding number of γ_1 and γ_2 about each other. In the second case, the product $\gamma_1\gamma_2$ is a loop, and so we still have a nice interpretation of a as the total algebraic winding number of $\gamma_1\gamma_2$ about the punctures q_i. Bigelow also points out that in this case the condition $\gamma_1 - \gamma_2)(1) = -(\gamma_1 - \gamma_2)(0)$ implies that b is an odd integer. It is worth noting that Turaev [130] interprets the integer b in a different way (in both cases). The composition of the map $\pi_1(C) \to S^1$ defined by

$$\gamma \mapsto \frac{\gamma_1(s) - \gamma_2(s)}{|\gamma_1(s) - \gamma_2(s)|}$$

with the usual projection from S^1 onto $\mathbb{R}P^1$ sends our loop γ to a loop in $\mathbb{R}P^1$. Let $\bar{\gamma}$ denote the homology class of this loop in $H_1(\mathbb{R}P^1)$. Up to a choice of generator u for $H_1(\mathbb{R}P^1) \cong \mathbb{Z}$, we have $\bar{\gamma} = u^b$. We define the map ϕ by setting $\phi(\gamma) = q^a t^b \in \mathbb{Z} \times \mathbb{Z}$, thought of as the free abelian group with basis $\{q, t\}$.

Now we proceed more or less as in the case of the Burau representation. Let $\tilde{C}$ be the regular cover of C which corresponds to the kernel of ϕ in $\pi_1(C, c_0)$. Let A denote the ring $\mathbb{Z}[t^{\pm 1}, q^{\pm 1}]$. The homology groups of $\tilde{C}$ naturally inherit an A-module structure via the action of $\mathrm{Aut}(\tilde{C})$.

It is clear that any homeomorphism $X : D_n \to D_n$ induces a homeomorphism $X' : C \to C$ defined by $X'(\{z_1, z_2\} = \{X(z_1), X(z_2)\}$. Thus X' necessarily fixes the basepoint c_0. If X'_* denotes the induced action on $\pi_1(C)$, then it is an easy exercise to check that X'_* preserves the values of the integers a and b defined above. Then X' lifts uniquely to a map $\tilde{X} : \tilde{C} \to \tilde{C}$ such that the fiber over the basepoint c_0 is fixed pointwise. Moreover, since $\tilde{X}$ commutes with the action of $\mathrm{Aut}(\tilde{C})$, we have that the induced action on $H_2(\tilde{C})$, denoted $\tilde{X}_*$, is an A-module automorphism. (Note that $\tilde{C}$ is a 4-manifold.) Now $H_2(\tilde{C})$ is a free A-module of rank $r = \begin{pmatrix} n \\ 2 \end{pmatrix}$ (Theorem 4.1 of [14]) and thus we can think of $\tilde{X}_* \in A$. We therefore define the <u>Lawrence-Krammer representation</u> as follows:

$$\begin{array}{rcl} \lambda : \mathbf{B}_n & \to & \mathrm{GL}_r(A) \\ X & \mapsto & \tilde{X}_* \end{array}$$

If we choose q and t to be algebraically independent in $\mathbb{C}$, we obtain a faithful representation $\lambda : \mathbf{B}_n \to GL_r(\mathbb{C})$.

Recall that $d_1, d_2 = c_0$ is the basepoint we have chosen for the configuration space C. A <u>fork</u> is a tree F embedded in the disk D such that

1. F has four vertices: d_1, z, q_i and q_j,

2. the three edges of F share z as a common vertex,

3. $F \cap \partial D_n = d_1$, and

4. $F \cap \{q_1, \ldots, q_n\} = \{q_i, q_j\}$.

We note that Krammer's algebraic definition of the representation λ depends on the induced action on an A-module generated forks.

The edge of F which contains the vertex d_1 is called the handle of F. The union of the other two edges is called the tine edge of F, denoted $T(F)$, and may contain punctures $q_k, k \neq i, j$. For any fork F we define a parallel copy F' to be a copy of the tree F embedded in D, with vertices $\{d_2, z', q_i$ and q_j such that $d_1 \neq d_2 \in \partial D_n$, $z \neq z'$, and F' is isotopic to F rel $\{q_i, q_j\}$. Given a fork F, we also define a noodle to be an arc N properly embedded in D_n with $\partial N = \{d_1, d_2\}$ and oriented so that its initial point is d_1 and its terminal point is d_2.

We next construct surfaces in $\tilde{C}$ associated to forks and noodles with which we will define a certain pairing on forks and noodles. For each fork F, we choose a parallel copy F' and first define a surface $S(F)$ as the set of all points of the form $\{x, y\} \in C$ with $x \in T(F) \setminus \{q_1, \ldots, q_n\}$ and $y \in T(F') \setminus \{q_1, \ldots, q_n\}$. Let ζ (resp. ζ') denote the handle of F (resp. F'), oriented from d_1 to z (resp. d_2 to z'). Define $\tilde{\zeta}$ to be the unique lift of $\{\zeta(s), \zeta'(s)\} \in \tilde{C}$ with initial point $\tilde{c_0}$, a fixed point in the fiber over c_0. Now let $\tilde{S}(F) \subset \tilde{C}$ be the unique lift of $S(F)$ which contains the terminal point of $\tilde{\zeta}$. Similarly, we can associate to each noodle N a surface $\tilde{S}(N) \subset \tilde{C}$. Define $S(N) \subset C$ as the set of all points of the form $\{x, y\} \in C$ such that $x, y \in N, x \neq y$, and let $\tilde{S}(N)$ be the unique lift of $S(N)$ which contains $\tilde{c_0}$.

Given a noodle N and a fork F, we define their pairing, denoted $\langle N, F\rangle$, to be an element of the ring A as follows. We can assume that (up to isotopy rel endpoints) $T(F)$ and N intersect transversely in a finite number of points $\{z_1, \ldots, z_r\}$ and similarly that $T(F')$ intersects N transversely at $\{z_1', \ldots, z_r'\}$. Then each pair z_i, z_i' cobounds an arc in N which lies between $T(F)$ and $T(F')$. Now for each $i, j = 1, \ldots, r$, there is a unique monomial $m_{i,j} = q^{a_{i,j}} t^{b_{i,j}}$ such that $m_{i,j}\tilde{S}(N)$ intersects the surface $\tilde{S}(F)$ at a point in the fiber over $z_i, z_j' \in C$. Letting $\epsilon_{i,j}$ denote the sign of that intersection, we define

$$\langle N, F\rangle = \sum_{i=1}^{r} \sum_{j=1}^{r} \epsilon_{i,j} q^{a_{i,j}} t^{b_{i,j}}.$$

Bigelow supplies a method for explicit calculation of the pairing for a given noodle and fork and shows that it is well defined on isotopy classes of forks rel q_i, q_j.

The proof that λ is faithful relies on two important lemmas (the 'Basic Lemma' and 'Key Lemma' of [14]), which we present here. As usual, we shall abuse notation and not distinguish between a mapping class and specific representatives.

Lemma 4.1 [14] *If $\lambda(X) = 1$, then $\langle N, F\rangle = \langle N, X(F)\rangle$ for every noodle N and every fork F.*

Lemma 4.2 [14] *For a noodle N and a fork F, the pairing $\langle N, F\rangle = 0$ if and only if $T(F)$ can be isotoped off N relative to q_i, q_j.*

Thus it is the pairing on noodles and forks which gives the essential characterization of the kernel of λ.

We end this section by noting two facts about the Lawrence-Krammer representation. First, Budney has shown that λ is unitary[1] for an appropriate choice of q and t (still

[1] Squier had previously shown that the Burau representation is unitary [125] .

algebraically independent) [41]. In other words, we have

$$\lambda : \mathbf{B}_n \to U_r(\mathbb{C})$$

where $\mathrm{U}_r = \{X \in \mathrm{GL}_r(\mathbb{C}) | \bar{X}^T = X^{-1}\}$. Thus the conjugacy class of $\lambda(X) \in U_r(\mathbb{C})$ is determined by its eigenvalues, and one could hope for an efficient solution to the conjugacy problem for $\mathbf{B}_n$ by passing to the whole of $U_r(\mathbb{C})$. However, Budney has given examples of non-conjugate braids $X_1, X_2 \in \mathbf{B}_n$ such that $\lambda(X_1)$ and $\lambda(X_2)$ are conjugate in $U_r(\mathbb{C})$ (see Section 4 of [41]). We will return to these examples in a bit more detail in §5.7.

Second, Matthew Zinno has shown a connection between the Lawrence-Krammer representation and the Birman-Murakami-Wenzl (BMW) algebra. The BMW algebra is related to Kauffman's knot polynomial and can be thought of as a deformation of the Brauer algebra in the same way that the Hecke algebra can be thought of as a deformation of the group algebra $\mathbb{C}\Sigma_n$ (see [33], [114]). Braid groups map homomorphically into the BMW algebra, giving rise to irreducible representations of $\mathbf{B}_n$. Zinno has identified a summand of the BMW algebra which corresponds exactly to the Lawrence-Krammer representation:

Theorem 15 [137] *The Lawrence-Krammer representation of* $\mathbf{B}_n$ *is equivalent to the* $(n-2) \times 1$ *irreducible representation of the BMW algebra.*

It follows immediately from Theorem 15 that the Lawrence-Krammer representation λ is irreducible and also that the regular representation of the BMW algebra is faithful.

4.6 Representations of other mapping class groups

The first theorem that we proved in this review article was Theorem 1, which asserted that the braid group $\mathbf{B}_{n-1}$ has a natural interpretation as the mapping class group $\mathcal{M}_{0,1,n-1}$ of the n-times punctured disc. One wonders, then, whether the linearity of $\mathbf{B}_n$ extends to a more general statement about the linearity of other mapping class groups $\mathcal{M}_{g,b,n}$?

The manuscripts of M. Korkmaz [96] and of S. Bigelow and R. Budney [16] exploit the very special connection between the braid group $\mathbf{B}_{n-1} = \mathcal{M}_{0,1,n-1}$ and the mapping class group $\mathcal{M}_{0,0,n}$, and between the mapping class groups $\mathcal{M}_{0,0,2g+2}$ and $\mathcal{M}_{g,0,0}$, to produce failthful finite dimensional representations of $\mathcal{M}_{0,0,n}$ for every n, also of $\mathcal{M}_{2,0,0}$, and finally of a particular subgroup of $\mathcal{M}_{g,0,0}$ for every $g \geq 3$. In the case of $\mathcal{M}_{2,0,0}$, the basic fact used by both Korkmaz and Bigelow-Budney is a theorem proved by Birman and Hilden in 1973 (see [18]), which asserts that the mapping class group $\mathcal{M}_{2,0,0}$ is a $\mathbb{Z}_2$ central extension of $\mathcal{M}_{0,0,6}$. This theorem is special to genus 2, which is the only group among the mapping class groups $\mathcal{M}_{g,0,0}$ which has a center. For $g \geq 3$ the so-called 'hyperelliptic involution' which generates the center of $\mathcal{M}_{2,0,0}$ generalizes to an involution whose centralizer is a subgroup of infinite index in $\mathcal{M}_{g,0,0}$. The same circle of ideas yield faithful matrix representations of those subgroups.

We now explain how Bigelow and Budney used the Lawrence-Krammer representation of the braid groups to obtain faithful matrix representations of $\mathcal{M}_{0,0,n.}$. If one considers any surface $S_{g,b,n}$ and caps one of the boundary components by a disc, one obtains a geometrically induced <u>disc-filling</u> homomorphism $d_\star : \mathcal{M}_{g,b,n-1} \to \mathcal{M}_{g,b-1,n}$. Its kernel is the Dehn twist about the distinguished boundary component (see §2.8 of [86]); a distinguished point in the interior of the new disc becomes the new fixed point. Applying these ideas to the braid group, one then sees that there is a natural homomorphism $d_\star : \mathbf{B}_{n-1} \to \mathcal{M}_{0,0,n}$. It turns out that image($d_\star$) is not the full group $\mathcal{M}_{0,1,n}$, but the stabilizer of the new fixed point, and kernel($d_\star$) is the infinite cyclic subgroup of $\mathbf{B}_{n-1}$ that is generated by the braid $C = (\sigma_{n-1}\sigma_{n-2}\cdots\sigma_2\sigma_1)^n$, a full twist of all of the n braid strands. Since the image $\lambda(C)$ in the Lawrence-Krammer representation is a scalar matrix (the diagonal entries are $q^{2(n-1)}t^2$ in the representation as it is defined in [14]), one obtains a faithful representation of image($d_\star$) by rescaling the Lawrence-Krammer matrices, setting $t^{-2} = q^{2(n-1)}$. The group image($d_\star$) is of finite index in $\mathcal{M}_{0,0,n}$, which therefore is a linear group. The dimension of the explicit representation of $\mathcal{M}_{0,0,n}$ constructed in [16] is $n\begin{pmatrix} n-1 \\ 2 \end{pmatrix}^2$. It leads to a related representation of dimension 64 of $\mathcal{M}_{2,0,0}$. In this regard we note that, while Korkmaz uses the identical geometry, he uses less care with regard to dimension, and his representation of $\mathcal{M}_{2,0,0}$ has dimension $2^{10}3^5 5^3$, which is very much bigger than 64.

Open Problem 6 Is there a faithful finite dimensional matrix representation of the mapping class group $\mathcal{M}_{g,b,n}$ for any values of the triplet (g, b, n) other than $(0, 1, n), (0, 0, n)$, $(2, 0, 0), (1, 0, 0)$ and $(1, 1, 0)$? A folklore conjecture is that most of the mapping class groups are, in fact linear. See [39] for evidence in this regard. New ideas seem to be needed to construct candidates. ♣

In view of the very large dimension of the representation of $\mathcal{M}_{2,0,0}$ which we just discussed, we note that there is an interesting 5-dimensional representation of the same group which occurs as one of the summands in the Hecke algebra representation of $\mathbf{B}_n$, namely the one belonging to the Young diagram with 2 rows and 3 columns. It is discussed in [89]. At this writing its kernel does not seem to be known, although it is known to be infinite.

4.7 Additional representations of $\mathbf{B}_n$.

We end our discussion of representations of the braid groups by describing a construction which yields infinitely many finite dimensional representations of $\mathbf{B}_n$, each one over a ring $\mathbb{C}[t_1, t_1^{-1}, t_2, t_2^{-1}, \ldots, t_k, t_k^{-1}]$, where $t_1, \ldots, t_k$ are parameter, for some $k \geq 1$. The construction includes all the summands in the Temperley-Lieb algebra and the Lawrence-Krammer representation too, and in addition infinitely many presumably new faithful representations of $\mathbf{B}_n$. It was first described in [25], generalizing ideas in [100]. It is due to Moody, with details first worked out by Long in [102].

We will be interested in the braid group $\mathbf{B}_n$ on n-strands, but to describe our construction it will be convenient to regard $\mathbf{B}_n$ as a subgroup of $\mathbf{B}_{n+1}$. Number the strands in the latter group as $0, 1, \ldots, n$. Let $\mathbf{B}_{1,n} \subset \mathbf{B}_{n+1}$ be the subgroup of braids in $\mathbf{B}_{n+1}$ whose associated permutation fixes the letter 0. Its relationship to $\mathbf{B}_n$ is given by the group extension

$$1 \to \mathbf{F}_n \to \mathbf{B}_{1,n} \to \mathbf{B}_n \to 1, \tag{19}$$

where the homomorphism $\mathbf{B}_{1,n} \to \mathbf{B}_n$ is defined by pulling out the zeroth braid strand. There is a cross section which is defined by mapping $\mathbf{B}_n$ to the subgroup of braids on strands $1, \ldots, n$ in $\mathbf{B}_{1,n}$. Therefore we may identify $\mathbf{B}_{1,n}$ with $\mathbf{F}_n \rtimes \mathbf{B}_n$. The semi-direct product structure arises when we regard $\mathbf{B}_n$ as a subgroup of the automorphism group of a free group. The action of $\mathbf{B}_n$ on F_n is well known and is given in [6], also [65], and also in [18]. Thinking of F_n as the fundamental group $\pi_1(S_{0,1,n})$ of the n-times punctured disc, the action of the elementary braid σ_i is given explicitly by:

$$\sigma_i \mathbf{x_j} \sigma_i^{-1} = \begin{cases} \mathbf{x_{i+1}} & \text{if } j = i; \\ \mathbf{x_{i+1}}^{-1}\mathbf{x_i}\mathbf{x_{i+1}} & \text{if j=i+1} \\ \mathbf{x_j} & \text{otherwise} \end{cases} \tag{20}$$

Since $\mathbf{B}_n$ is a subgroup of $\mathbf{B}_{n+1}$ we see In this way that the groups $\mathbf{B}_n,\ \mathbf{F}_n$ and also $\mathbf{F}_n \rtimes \mathbf{B}_n$ are all subgroups of $\mathbf{B}_{n+1}$.

In order to describe the idea behind the construction we recall the notion of homology or cohomology of a space with coefficients in a flat vector bundle. Suppose that X is a manifold and that we are given a representation $\rho : \pi_1(X) \to GL(V)$. This enables us to define a flat vector bundle E_ρ: Let $\tilde{X}$ be the universal covering of X. The group $\pi_1(X)$ acts on $\tilde{X} \times V$ by $g.(\tilde{x}, \mathbf{v}) = (g.\tilde{x}, \rho(g).\mathbf{v})$. Then E_ρ is the quotient of $\tilde{X} \times V$ by this action. We now form the cohomology groups of 1-forms with coefficients in E_ρ, denoting these by $H^1(X;\rho)$ or $H^1_c(X;\rho)$ for compactly supported cochains. In order to get an action of the braid groups, we use (20). The action gives a canonical way of forming a split extension $F_n \rtimes \mathbf{B}_n$. It turns out that in order to get an action on the twisted cohomology group what is required is exactly a representation of this split extension. Since $\mathbf{B}_{1,n}$ is a subgroup of $\mathbf{B}_{n+1}$, any representation of the latter will of course do the job, and that is why we think of $\mathbf{B}_n$ as a subgroup of $\mathbf{B}_{n+1}$.

Theorem 16 [25] *Given a representation $\rho : F_n \rtimes \mathbf{B}_n \to GL(V)$ we may construct another representation $\rho_t^+ : \mathbf{B}_n \to H^1_c(S_{0,1,n};\rho)$ where t is a new parameter. In particular, given any representation $\rho : B_{n+1} \to GL(V)$, we may construct a representation $\rho_t^+ : \mathbf{B}_n \to H^1_c(S_{0,1,n};\rho)$.*

This works in exactly the way one might expect. The representation restricted to the free factor gives rise to the local system on the punctured disc and thus the twisted cohomology group and the compatibility condition provided by the split extension structure gives the braid group action.

A comment is in order concerning Theorem 16. Although the theorem is stated abstractly, (20) gives a concrete recipe which enables one to write down the description of ρ_t^+ given ρ. Moreover, as we have noted, any time that we have a representation of $\mathbf{B}_{n+1}$ we also have one of the semi-direct product, which we identify with the subgroup $\mathbf{B}_{1,n}$.

The theorem shows that given a k parameter representation of the braid group, the construction yields a $k+1$ parameter representation, apparently in a nontrivial way. For example, if one starts with the (zero parameter) trivial representation of $F_n \rtimes \mathbf{B}_n$, the theorem produces the Burau representation, and starting with the Burau representation, the construction produces the Lawrence-Krammer representation [102]. However the role of this extra parameter is not purely to add extra complication, it also adds extra structure. For there is a natural notion of what it should mean for a representation of a braid group to be unitary (See [125], for example) and the results of Deligne-Mostow and Kohno imply:

Theorem 17 [102] *In the above notation, if ρ is unitary, then for generic values of s, so is ρ_s^+.*

Open Problem 7 This problem is somewhat vague. It begins with a suggestion that Long's construction be studied in greater detail, and goes on to ask whether (a wild guess) all finite dimensional unitary matrix representations of $\mathbf{B}_n$ arise in a manner which is related to the construction of Theorem 16? ♣

5 The word and conjugacy problems in the braid groups

Let G be any finitely generated group. Fix a set of generators for G. A word ω is a word in the given set of generators and their inverses. The element of G that it represents will be denoted $[\omega]$, and its conjugacy class will be denoted $\{\omega\}$. We consider two problems. The word problem begins with $\omega, \omega' \in G$, and asks for an algorithm that will decide whether $[\omega] = [\omega']$? The conjugacy problem asks for an algorithm to decide whether $\{\omega\} = \{\omega'\}$, i.e. whether there exists α such that $\{\omega'\} = \{\alpha^{-1}\omega\alpha\}$? A sharper version asks for a procedure for finding α, if it exists. In both cases we are interested in the complexity of the algorithm, and ask whether it is polynomial in either braid index n or word length $|\omega|$ or both? The word problem and the conjugacy are two of the three classical 'decision problems' first posed by Max Dehn [51]. In the case we consider here of $G = \mathbf{B}_n$, the set of generators we choose will be those of either the classical or the new presentation. Artin gave the first solution to the word problem for the braid group in 1925 [6], and a number of fundamentally different solutions to the word problem in $\mathbf{B}_n$ exist today, with at least two of them being polynomial in both n and $|\omega|$. We focus here on an approach due to Frank Garside [76] and improved on by a number of others, and briefly describe other methods at the end of this section.

On the other hand, we know of only one definitive solution to the conjugacy problem, namely the combinatorial solution that was discovered by Garside. He found a finite set of conjugates of an arbitrary element $[\omega] \in \mathbf{B}_n$, the 'Summit Set' $\mathbf{S}_\omega$ of ω, with the properties that if $\{\omega\} = \{\beta\}$, then $\mathbf{S}_\omega = \mathbf{S}_\beta$, whereas if $\{\omega\} \neq \{\beta\}$, then $\mathbf{S}_\omega \cap \mathbf{S}_\beta = \emptyset$. If $\mathbf{S}_\omega = \mathbf{S}_\beta$, his methods also find α such that $\beta = \alpha\omega\alpha^{-1}$. While his algorithm has been improved in major ways over the years, at this writing it is exponential in both n and $|\omega|$. The principle difficulties may be explained in the following way:

($\star$) The entire set $\mathbf{S}_\omega$ and a single element in $\mathbf{S}_\beta$ must be computed to decide whether ω and β are or are not conjugate. While the calculation of a single element can now be done rapidly, in the general case $\mathbf{S}_\omega$ has unpredictable size. The combinatorics that determine the size of $\mathbf{S}_\omega$ are particularly subtle and difficult to understand, and at this writing only partial progress has been made, in spite of much effort by experts.

The improvements that have been made over the years have included the replacement of $\mathbf{S}_\omega$ by a proper subset S_ω, the 'Super Summit Set', via the work of ElRifai and Morton in [61]. However, while S_ω is very much smaller than $\mathbf{S}_\omega$, the principle difficulty ($\star$) is unchanged. Very recently Gebhardt found a still smaller subset, the 'Ultra Summit Set U_ω', to replace S_ω [77]. The difficulty is unfortunately the same as it was for the summit set and the super summit set, but can be made more specific: the thing that one needs to understand is the number of 'closed cycling orbits' in U_ω and the length of each orbit, and how distinct orbits are related. Sang Jin Lee [99] has given examples to show that the number of orbits and their size can be arbitrarily large, and that it is in no way clear how distinct orbits are related, except in special cases. More work remains to be done.

It was shown in Theorem 1, proved earlier in this article, that there is a faithful action of the braid group $\mathbf{B}_n$ on the n-times punctured disc. Investigating this action (in the more general setting of the action of the mapping class group of a 2-manifold on the 2-manifold), Thurston proposed in the early 1980's a very different approach to the conjugacy problem which is based upon the dynamical properties of that action. These ideas were investigated in Lee Mosher's PhD thesis [112]. The curious fact is that while the dynamical picture quickly yields several very interesting class invariants not easily seen in the combinatorial picture, namely a graph (known as a 'train track') which is embedded in the surface and is invariant under the action of properly chosen representatives of $\{\omega\}$, and a real number λ, however they are not enough to determine conjugacy. Generic elements in the mapping class group of a surface have an invariant train track, but it is not unique. It is known that there are finitely many possible train tracks associated to a given element, but their enumeration remains an unsolved problem, and that is what gets in the way of a nice solution to the conjugacy problem. In fact, when one begins to understand the details, the entire picture suggests difficulties very much like those in ($\star$) again.

We have just one more remark of a general nature, before we proceed to review all these matters in more detail. One of the reasons that the Garside approach to the conjugacy problem is interesting is because the techniques that were developed for the braid groups revealed unexpected structures in $\mathbf{B}_n$ that generalize to related unexpected structures in Artin groups (see §5.4), and in the less well-known class of 'Garside groups'. See § 5.2. There has been major activity in recent years regarding these structures, and there are also many open problems. The same can also be said for the Thurston approach, as it too generalizes from an action of $\mathbf{B}_n$ on the n-times punctured disc to actions of mapping class groups on curves on surfaces, discussed in §5.6. All of these matters, and other related ones, will be discussed below.

5.1 The Garside approach, as improved over the years

We assume, initially, that the group $\mathbf{B}_n$ is defined by the classical presentation (2), with generators $\sigma_1, \ldots, \sigma_{n-1}$. In this subsection we describe the solution to the word and conjugacy problem which was discovered by Garside in 1968 (see [76]) and subsequently sharpened and expanded, in many ways, through the contributions of others, in particular Thurston [62], ElRifai and Morton [61], Birman, Ko and Lee ([23] and [24]), Franco and Gonzales-Meneses [70] and most recently Gebhardt [77]. Other relevant papers are [82] and [83].

In (G1)-(G6) below we describe a constructive method for finding a normal form for words. In (G7)-(G10) we describe how to find a unique finite set of words in normal form that characterizes $\{\omega\}$.

(G1) Elements of $\mathbf{B}_n$ which can be represented by braid words which only involve positive powers of the σ_i are called <u>positive braids</u>. A key fact which was observed by Garside in [76] is that the presentation (2) not only defines the group $\mathbf{B}_n$, it also defines a monoid $\mathbf{B}_n^+$. He then went on to prove that this monoid embeds in the obvious natural way in $\mathbf{B}_n$, in the strong sense that two positive words in $\mathbf{B}_n$ define the same element of $\mathbf{B}_n$ if and only if they also define the same element in the monoid $\mathbf{B}_n^+$. The monoid of positive braids is particularly useful in studying $\mathbf{B}_n$ because the defining relations all preserve word length, so that the number of candidates for a positive word which represents a positive braid is finite. However, all this is of little use unless we can show that the monoid of positive braids is more than just a small and very special subset of $\mathbf{B}_n$. In (G3) below we show that this is indeed the case.

(G2) A special role is played by the Garside braid Δ in $\mathbf{B}_n$. It is a positive half-twist of all the braid strands, and is defined by the braid word

$$\Delta_n = (\sigma_{n-1}\sigma_{n-2}\cdots\sigma_1)(\sigma_{n-1}\sigma_{n-2}\cdots\sigma_2)\cdots(\sigma_{n-1}\sigma_{n-2})(\sigma_{n-1}).$$

It is illustrated in sketch (i) of Figure 24, for $n = 5$. The square of Δ_n (a full twist of all

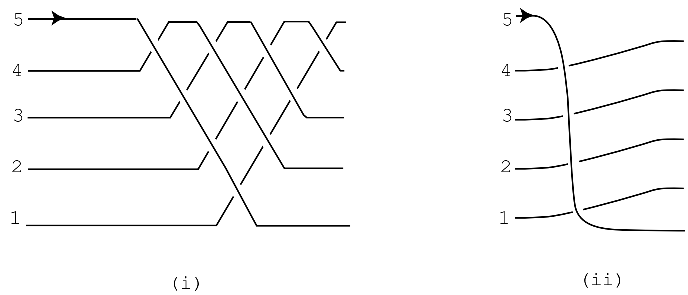


Figure 24: (i) The 5-braid Δ_5. (ii) The 5-braid δ_5.

the braid strands) generates the infinite cyclic center of $\mathbf{B}_n$. The inner automorphism $\tau : \mathbf{B}_n \to \mathbf{B}_n$ which is defined by $\tau(X) = \Delta^{-1}X\Delta$ is the symmetry which sends each σ_i to σ_{n-i}. Of course Δ itself is invariant under this symmetry. Later we will show that almost all of the combinatorics which we are describing hold equally well for the new presentation (3). The role of Δ in the classical presentation is replaced by that of the braid $\delta = \sigma_{n-1,n}\cdots\sigma_{2,3}\sigma_{1,2}$ which is depicted in Figure 24(ii). See §5.3 below.

(G3) The Garside braid Δ is very rich in elementary braid transformations. In particular, for each $i = 1, \ldots, n-1$ there exist (non-unique) positive braids L_i and R_i such that $\Delta = L_i\sigma_i = \sigma_i R_i$. This implies that $\sigma_i^{-1} = \Delta^{-1}L_i = R_i\Delta^{-1}$ for each $i = 1, \ldots, n-1$, so that an arbitrary word $\omega = \sigma_{\mu_1}^{\epsilon_1}\cdots\sigma_{\mu_r}^{\epsilon_r}, \quad \epsilon_q = \pm 1$, can be converted to a word which uses only positive braid generators, at the expense of inserting arbitrary powers of Δ^{-1}. Since, by (G2) we have $\beta\Delta^\epsilon = \Delta^\epsilon\tau(\beta)$ for every braid β, all powers of Δ^{-1} can be moved to the left (or right). It follows that $[\omega]$ is also defined by a word of the form $\Delta^r T$ where T is positive. This representation is non-unique.

(G4) To begin to find unique aspects, Garside observes that there is a maximum value, say i, of r with the property that the braid defined by ω is also represented by a word $\Delta^i Z$ where Z is positive and i is maximal for all braids of this form. Note that this means that Z cannot be written in the form $Z_1\Delta Z_2$ for any positive words Z_1, Z_2, because if it could, then it would be possible, by (G2), to rewrite $\Delta^i Z$ as $\Delta^{i+1}\tau(Z_1)Z_2$, with $\tau(Z_1)Z_2 \in \mathbf{B}_n^+$, contradicting the maximality of i. The integer i is known as the infimum of $[\omega]$, and written $inf(\omega)$. It is an invariant of $[\omega]$. Garside's complete invariant of $[\omega]$ is $\Delta^i Z_0$, where among all positive words $Z_0, Z_1, \ldots, Z_q$ that define the same element as Z, one chooses the unique word Z_0 whose subscript array (as a positive word in $\sigma_1, \ldots, \sigma_{n-1}$) is lexicographically minimal. This is Garside's solution to the word problem in $\mathbf{B}_n$. It is exponential in both n and $|\omega|$ because there exist braids which admit very many of the elementary braid transformations in (2). To make this very explicit, we note that by [126] the number of positive words of letter length $(n)(n-1)/2$ in the standard braid generators that represents Δ_n is the same as the number of standard Young tableaux of shape $(n, n-1, \ldots, 1)$. It is given by the hook length formula as $((n)(n-1)/2))!/(1^{n-1}3^{n-1}5^{n-2}\cdots(2n-3))$. Calculating, we learn that Δ_5 can be represented by 768 positive words, all of letter length 10, in the $\sigma_i's$, giving a good idea of the problem. Moreover, this formula shows that the number of words increases exponentially with n. Of course the elements that are of interest to us, that is the positive braids defined by words like Z above, do not contain Δ, however there is not reason why they cannot contain, for example, arbitrarily high powers of large subwords of Δ.

(G5) Let l and r be positive words such that there is a factorization of Δ in the form $\Delta = lr$. We call l a <u>left divisor</u> of Δ and r a <u>right divisor</u> of Δ. Let $\mathcal{P}$ (respectively $\mathcal{P}'$) be the collection of all left (resp. right) divisors of Δ. Note that we already used the fact that each of the generators σ_i belongs to both $\mathcal{P}$ and $\mathcal{P}'$. As it turns out, the sets $\mathcal{P}$ and $\mathcal{P}'$ coincide. The set $\mathcal{P}$ will be seen to play an essential role in Garside's solution to the conjugacy problem, but before we discuss that we describe how to use it to get a very fast solution to the word problem. The solution to the word problem was first discovered by Adyan [1] in 1984, but Adyan's work was not well-known in the West. It was not referenced by either Thurston (see Chapter 5 of [62]) or ElRifai and Morton [61], who rediscovered Adyan's work 8-10 years later, and added to it. Nevertheless, we use [61] and [62] as our main source rather than [1], because the point of view in those papers leads us more naturally to recent generalizations.

(G6) A key observation is that there are $n!$ braids in $\mathcal{P}$, and that these braids are in 1-1 correspondence with the permutations of their end-points, under the correspondence defined by sending each σ_i to the transposition $(i, i+1)$. For this reason $\mathcal{P}$ is known as the set of <u>permutation braids</u>. This, combined with the fact that the crossings are always positive, makes it possible to reconstruct any permutation braid from its permutation, and so to obtain a unique element, even though its representation as a word is highly non-unique. Permutation braids have a key property: after the braid is tightened, any two strands cross at most once, positively. To understand the importance of this fact, we note that any positive 5-braid in which 2 strands cross at most once, whose associated permutation is $(1,2,3,4,5) \to (5,4,3,2,1)$ must be Δ_5, a criterion which is an enormous improvement over searching through the 768 distinct positive words that the hook length formula showed us represent Δ_5, giving a good idea of the simplification that Adyan, Thurston and ElRifai and Morton discovered.

The final step in finding a very rapid solution to the word problem in $\mathbf{B}_n$ is a unique way to factorize the braid P in the partial normal form $\Delta^i P$ of (G4) above as a unique product of finitely many permutation braids. If the strands in P cross at most once, then $P \in \mathcal{P}$ and we may choose any positive braid word, say l_1, which has the same permutation as P as its representative. If not, set $P = l_1 l_1^\star$ where l_1 is a positive braid of maximal length in which two strands cross at most once and $l^\star$ is the rest of P. If no two strands in $l_1^\star$ cross more than once, set $l_2 = l_1^\star$, and stop. If not, repeat, setting $l_1^\star = l_2 l_2^\star$, where l_2 has maximal length among all positive braids whose strands cross at most once, and so forth to obtain $P = l_1 l_2 \cdots l_s$ where each $l_i \in \mathcal{P}$ and each l_i has maximal length for all factorizations of $l_{i-1}^{-1} \cdots l_1^{-1} L$ as $l_i l_i^\star$ with $l_i \in \mathcal{P}$ and $l_i^\star$

positive. This representation is unique, up to the choices of the words which represent $l_1, \ldots, l_s$. As we remarked earlier, each l_i is determined uniquely as an element in $\mathbf{B}_n^+$ by the permutation of its strands.

The factorization of an arbitrary braid word as $\Delta^i P = \Delta^i l_1 l_2 \cdots l_s$ is the left greedy normal form, where the term 'left greedy' suggests the fact that each l_i is a maximal permutation subbraid relative to the positive braid to its right. The left greedy normal form solves the word problem. The integer i was already defined to be $inf(\omega)$. The integer $i + s$ is known as the supremum of ω, written $sup(\omega)$. The integer s is the canonical length $L(\omega)$.

Remark 5.1 The Adyan-Thurston-ElRifai-Morton solution to the word problem is shown in [23] to be $\mathcal{O}(|\omega|^2 n\text{logn})$, where $|\omega|$ is the word length of the initial representative of the braid ω, as a word in the standard generators of $\mathbf{B}_n$.

Example 5.1 We illustrate how to find the left greedy normal form for a positive braid via an example. Assume that we are given the positive braid

$$P = \sigma_1\sigma_3\sigma_2^2\sigma_1\sigma_3^2\sigma_2\sigma_3\sigma_2.$$

The first step is to factor P as a product of permutation braids by working along the braid word from left to right and inserting the next partition whenever two strands are about to cross a second time since the beginning of the current partition. The reader may wish to draw a pictures to go with this example. This gives the factorization:

$$(\sigma_1\sigma_3\sigma_2)(\sigma_2\sigma_1\sigma_3)(\sigma_3\sigma_2\sigma_3)(\sigma_2)$$

Next, start at the right end of the braid word, and ask whether the elementary braid relations can be used to move crossings from one partition to the partition on its immediate left, without forcing an adjacent pair of strands in a factor to cross. This process is repeated three times. The first time one applies the braid relations to factor 3, and then to adjacent letters in factors 2 and 3, to increase the length of factor 2 at the expense of decreasing the length of factor 3:

$$(\sigma_1\sigma_3\sigma_2)(\sigma_2\sigma_1\sigma_3\sigma_2)(\sigma_3\sigma_2)(\sigma_2.)$$

In fact one can push one more crossing from factor 3 to factor 2:

$$(\sigma_1\sigma_3\sigma_2)(\sigma_2\sigma_1\sigma_2\sigma_3\sigma_2)(\sigma_2)(\sigma_2),$$

Now it is possible to move a crossing from factor 2 to factor 1, giving:

$$(\sigma_1\sigma_3\sigma_2\sigma_1)(\sigma_2\sigma_1\sigma_3\sigma_2)(\sigma_2)(\sigma_2),$$

No further changes are possible, we have achieved left greedy normal form. If it had happened that P contained Δ_4, our factorization would reveal it because $\Delta_4 \in \mathcal{P}$. Each canonical factor represents a braid in $\mathcal{P} \subset \mathbf{B}_4^+$ and so has length strictly less than $|\Delta_4| = 6$. The words which represent the canonical factors are non-unique. Their associated permutations determine the canonical factors uniquely. ♠

(G7) We pass to the conjugacy problem, which builds on the solution just given to the word problem. From now on a word ω be always be assumed to be in left greedy normal form. Notice that the abelianizing map $e : \mathbf{B}_n \to \mathbb{Z}$ has infinite cyclic image, and that $e(\omega)$ is the exponent sum of a representing word in the σ_i. Clearly $e(\omega)$ is an invariant of both word and conjugacy class. It follows that there are only finitely many braids ω' in the conjugacy class $\{\omega\}$ which have left greedy normal form with $inf(\omega') > inf(\omega)$, because $e(\Delta) = (n)(n-1)/2$ and any increase in $inf(\omega)$ must come at the expense of a corresponding decrease in word length of the positive word that remains after all powers of Δ have been pushed to the left. Let $Inf(\omega)$ be the maximum value of $inf(\omega')$ for all braids ω' in the conjugacy class $\{\omega\}$. Assume that a representative ω_1 of $\{\omega\}$ is given and that it has left greedy normal form $\Delta^I L_1 L_2 \ldots L_S$, where $I = Inf(\omega)$ and S is minimal for all braids in normal form which are conjugate to ω. Let $Sup(\omega)$ be the integer $I + S$. Then $Inf(\omega)$ and $Sup(\omega)$ are class invariants of ω. The ElRifai-Morton super summit set S_ω is the collection of all elements in left-greedy normal form which realize $Inf(\omega)$ and $Sup(\omega)$. It is a major improvement over Garside's summit set $\mathbf{S}(\omega)$, which is the larger set of all elements which realize $Inf(\omega)$ but not $Sup(\omega)$, with left greedy form replaced by the subscript ordering described in (G4) above.

One might wonder how the normal forms which we just described for words and conjugacy classes relate to length functions on $\mathbf{B}_n$. In this regard, R. Charney has introduced a concept of 'geodesic length' in [44]. As above, let $\mathcal{P}$ be the set of all permutation braids. Charney defines the geodesic length of a braid ω to be the smallest integer $K = K(\omega)$ such that there is a word $L_1^{\epsilon_1} L_2^{\epsilon_2} \cdots L_K^{\epsilon_K}$, where each $\epsilon_i = \pm 1$, which represents the conjugacy class $\{\omega\}$, with each $L_i \in \mathcal{P}$ and each L_i^{-1} the inverse of a word in $\mathcal{P}$. If $\Delta^u L_1 L_2 \cdots L_s$ is an arbitrary element in the super summit set of ω, it is not difficult to show that the geodesic length of $\{\omega\}$ is the maximum of the 3 integers $(s+u, -u, s)$. We remark that Charney's geodesic length is defined in [44] for all Artin groups of finite type. This and other ways in which the Garside machinery generalizes to other classes of groups, including all Artin groups of finite type, will be discussed (much too briefly) in §5.2.

Open Problem 8 In the manuscript [98] Krammer gives several proofs of the faithfulness of the Lawrence-Krammer representation of $\mathbf{B}_n$. One of his very interesting proofs shows that if $[\omega] \in \mathbf{B}_n$, then it is a very simple matter to read Charney's geodesic length from the matrix representation of ω. Since the unique element of geodesic length zero is the identity, it follows that the representation is faithful. Our suggestion for future work is to investigate the Garside solution to the conjugacy problem and its improvements (to be described below) via the Lawrence-Krammer representation of $\mathbf{B}_n$. We mention this because we feel that this aspect of Krammer's work has received very little attention. ♣

(G8) Here is a fast constructive procedure for finding $Inf(\omega)$ and $Sup(\omega)$, due to ElRifai and Morton [61] and to Birman, Ko and Lee in [24]. Starting with any element in $\{\omega\}$, define the cycling of $\omega = \Delta^i l_1 l_2 \cdots l_k$ to be the braid $c(\omega) = \tau^{-i}(l_1)\omega\tau(l_1)$ and the decycling of ω to be the braid $d(\omega) = l_k \omega (l_k)^{-1}$. Note that $c(\omega)$ and $d(\omega) \in \{\omega\}$.

Putting $c(\omega)$ and $d(\omega)$ into left greedy normal form, one obtains braids which have at least i powers of Δ, and possibly more because it can happen that after cycling the braid $\Delta^i l_1 l_2 \cdots l_k$ will change to one whose left greedy normal form is $\Delta^{i'} l'_1 l_2 \cdots l'_k s'$ with $i' \geq i$. This would, of course, reduce s if it increases i. In [61] ElRifai and Morton proved that if $inf(\omega)$ is not maximal for the conjugacy class, then it can be increased by repeated cycling. Also, if $sup(\omega)$ is not minimal, then it can be decreased by decycling.

They could not say, however, how many times one might have to cycle or decycle before being sure that no further improvement was possible. The solution to that problem was found by Birman, Ko and Lee in [24]. It was shown that if $inf(\omega)$ is not maximal for the conjugacy class if it will be increased after fewer than $(n)(n-1)/2$ cyclings, and similarly if $sup(\omega)$ is not minimal it will be decreased after fewer than $(n-1)(n)/2$ decyclings. One then has a tool for increasing $inf(\omega)$ and decreasing $sup(\omega)$, and also a test which tells, definitively, when no further increase or decrease is possible.

Remark 5.2 The fact that everything we do to compute Inf(ω) also applies to the computation of Sup(ω) is not surprising because ElRifai and Morton showed that $Sup(\omega)$ = -Inf(ω^{-1}).

(G9) We now come to Gebhardt's very new work. Following the steps given in (G8) above, one will have on hand the summit set $S(\omega)$, i.e. the set of all braids in the conjugacy class $\{\omega\}$ which have left greedy normal form $\Delta^I L_1 L_2 \cdots L_S$, where $I = Inf(\omega)$ is maximal for the class and $I + S = Sup(\omega)$ is minimal for the class. This set is still very big. Gebhardt's improvement is to show that it suffices to consider only the subset of braids which in a closed orbit under cycling. This finite set of words is called the <u>ultra summit set</u> U_ω. It has been proved by Volker Gebhardt [77] that $\{\omega\} = \{\omega'\}$ if and only if $U_\omega = U_{\omega'}$. Note that Gebhardt does not need to use decycling, he proves that it suffices to consider the closed orbits under cycling.

(G10) To compute U_ω, Gebhardt also shows that if $\omega_i, \omega_j \in U_\omega$, then there is a finite chain $\omega_i = \omega_{i,1} \to \omega_{i,2} \to \cdots \to \omega_{i,q} = \omega_j$ of braids, with each $\omega_{i,j}$ in U_ω such that each $\omega_{i,t}$ is obtained from $\omega_{i,t-1}$ by conjugating by a single element in $\mathcal{P}$. Thus the following steps suffice to compute U_ω, after one knows a single element $\rho \in U_\omega$: One first computes the conjugates of ρ by the $n!$ elements in $\mathcal{P}$. One then puts each into left greedy normal form, and discards any braid that either does not (a) realize Inf and Sup, or (b) realizes Inf and Sup but is not in a closed orbit under cycling, or (c) realizes Inf and Sup and is in a closed orbit under cycling but is not a new element in U_ω. Ultimately, the list of elements so obtained closes to give U_ω.

Note that in doing this computation, one not only learns, for each $\omega_i, \omega_j \in U_\omega$, that $\{\omega_j\} = \{\omega_j\}$, but one also computes an explicit element α such that $\omega_j = \alpha^{-1}\omega_i\alpha$.

An inefficient part of this computation is the constant need to access the $n! - 2$ non-trivial elements in $\mathcal{P}$. A more efficient process is known (see (G11) below), but to describe it, we need new notions, which will be introduced after we discuss a wider class of groups, known as Garside groups.

5.2 Generalizations: from $\mathbf{B}_n$ to Garside groups

The first person to realize that the structure described in (G1)-(G10) is not restricted to braids was Garside himself [76], but his generalizations were limited to examples. Soon after his paper was published, the ideas were shown to go through, appropriately modified, in all finite type Artin groups, i.e., Artin groups whose associated Coxeter group is finite, by Breiskorn and Saito [40] and by Deligne [56]. See [106] for a survey (and reworking) of the results first proved in [40] and [56]. (See §5.4 for definitions of Artin and Coxeter groups in general.) They used explicit properties of finite reflection groups in their proof, but Paris and Dehornoy were thinking more generally and defined a broader class which they called 'Garside groups'. The class included all finite type Artin groups. Over the last few years, several tentative definitions of the term 'Garside group', referring to various classes of groups that generalize the braid group in this way, were proposed and appear in the literature (see [55], [121], for example) before the one that we give below was agreed upon by many, although the search for the most general class of such groups continues and the definition of 'Garside groups' is likely to continue to be in flux for some time. As we proceed through the definitions of 'Garside monoids', 'Garside structures', and 'Garside groups', the reader can look to the classical presentation of $\mathbf{B}_n$ for examples.

Given a finitely generated monoid G^+ with identity e, we can define a partial orders on its elements. Let $a, b \in G^+$. We say that $a \prec b$ if a is a left divisor of b, i.e. there exists $c \in M,\ c \neq e$, with $ac = b$. Also $b \succ a$ if there exists c such that $b = ca$. Caution: the two orderings are really different, that is, $a \prec b$ does not imply that $b \succ a$. An interesting example in the positive braid monoid $\mathbf{B}_n^+$ is the partial order induced on the elements in the sets $\mathcal{P}$ of left divisors of Δ_n, which gives it the structure of a lattice. The reader who wishes to get a feeling for the ordering might wish to construct the lattice in the cases Δ_3 and Δ_4. The lattice for Δ_n has $n!$ elements.

Now given $a, b \in G^+$ we can define in a natural way the (left) greatest common divisor of a and b, if it exists, written $d = a \wedge b$, as follows: $d \prec a, d \prec b$ and if, for any x, it happens that $x \prec a$ and $x \prec b$ then $x \prec d$. Similarly, we define the (left) least common multiple of a, b, denoted $m = a \vee b$, if $a \prec m, b \prec m$ and if for any x it happens that $a \prec x$ and $b \prec x$ then $m \prec x$.

It turns out that in the case of the braid monoid $\mathbf{B}_n^+$ the left partial ordering extends to a right-invariant ordering on the full braid group $\mathbf{B}_n$, a matter which we will discuss in §5.7 below.

We continue with our description of the features of the monoid $\mathbf{B}_n^+ \subset \mathbf{B}_n$ which will lead us to define more general monoids G^+ and their associated groups G. An element $x \in G^+$ is an atom if $x \neq e$ and if x has no proper left or right divisors. For example, the generators $\sigma_1, \ldots, \sigma_{n-1}$ are the atoms in the monoid G^+. Note that by (G3) above the atoms in $\mathbf{B}_n^+$ are left and right divisors of Δ, also they generate $\mathbf{B}_n^+$.

While we have not had occasion to introduce a key property of the Garside braid Δ before this, we do so now: As noted earlier, its sets $\mathcal{P}$ and $\mathcal{P}'$ of left and right divisors of Δ coincide. This property was used in the proofs of the facts that we described in (G1)-(G10) above, and is part of a long story about symmetries in the braid group.

A monoid G^+ is a Garside monoid if:

1. G^+ is generated by its atoms,

2. For every $x \in G^+$ there exists an integer $l(x) > 0$ such that x cannot be written as a product of more than $l(x)$ atoms,

3. G^+ is left and right cancellative, and every pair of elements in G^+ admits a left and also a right least common multiple and greatest common divisor,

4. There is an element $\Delta \in G^+$, the Garside element, whose left divisors and right divisors coincide, also they form a lattice, also each generates G^+.

All of the data just listed is called a Garside structure. It follows (via the work of Ore, described in Volume 1 of the book [46]) that every Garside monoid embeds in its group of fractions, which is defined to be a Garside group. It is not difficult to prove that the Garside element in a Garside monoid is the least common multiple of its set of atoms. An interesting property is that a Garside monoid admits a presentation $\langle S \mid R \rangle$, where for every pair of generators x, y there is a relation of the form $x \cdots = y \cdots$ that prescribes how to complete x and y on the right in order to obtain equal elements.

Examples abound. While we shall see that the braid groups are torsion-free, there are examples of Garside groups which have torsion. Whereas the relations in the braid monoid all preserve word length (which results in various finiteness aspects of the algorithms that we described) there are examples of Garside monoids in which this is not the case. As noted above, every Artin group of finite type is a Garside group. Torus knot groups are Garside groups, as are fundamental groups of complements of complex lines through the origin [55]. See [121] for additional examples.

Our reason for introducing Garside groups is that the solutions to the word problem and conjugacy problem in $\mathbf{B}_n$ described in §5.1, suitably modified, generalize to the class of Garside groups ([55], [121]), as do the simplifications of (G10) which we now describe. These improvements were first discovered by Franco and Gonzales-Meneses [70], and later improved by Gebhardt [77].

We return to the braid group, with the partial ordering of $\mathbf{B}_n$ on hand:

(G11) We already learned that if we begin with an arbitrary braid ω, then after a bounded number of cyclings and decyclings we will obtain a braid in $\{\omega\}$ which realizes $Inf(\omega)$ and $Sup(\omega)$. Continuing to cycle, we will arrive at an element, say ρ, in U_ω. The remaining task is the computation of the full set U_ω, and the method described in (G10) is inefficient. The difficulty is that, starting with $\rho \in U_\omega$ one is forced to compute its conjugates by *all* the $n!$ elements in $\mathcal{P}$, even though many of those will turn out to either not realize $Inf(\omega)$ and/or $Sup(\omega)$, and so will be discarded, whereas others will turn out to be duplicates of ones already computed. Moreover, this inefficient step is done repeatedly.

The good news is that, following ideas first introduced by N. Franco and J. Gonzalez-Meneses in [70], Gebhardt proves in Theorem 1.17 of [77], that if $\rho \in U_\omega$, and if $a, b \in \mathcal{P}$, with $a^{-1}\rho a$ and $b^{-1}\rho b \in U_\omega$, then $c^{-1}\rho c \in U_\omega$, where $c = a \wedge b$. See [70], and then [77], for a systematic procedure that allows one to use this fact to find all the orbits in U_ω efficiently. In this regard we remark that the work of Gebhardt is very new. The major open problem that remains is to improve it to an algorithm which will be polynomial in the word length of ω:

Open Problem 9 The bad news is that, like the Garside's summit set and ElRifai and Morton's super summit set, Gebhardt's ultra summit set also has the key difficulty ($\star$). More work remains to be done. ♣

5.3 The new presentation and multiple Garside structures

As it turns out, essentially all of the structure that we just described also exists with respect to the second presentation of $\mathbf{B}_n$ which was given in §1 of this paper. This is the main result of [24]. The fundamental braid Δ_n is replaced by $\delta_n = \sigma_{n-1,n}\sigma_{n-2,n-1}\cdots\sigma_{2,3}\sigma_{1,2}$. It is proved in [24] that the associated monoid $\mathbf{B}_n^+$ embeds in $\mathbf{B}_n$, and that all the results described in (G1)-(G11) above have curious (and very surprising) variations which hold in the new situation.

For example, the elements in the set of left divisors of δ_n are in 1-1 correspondence with a set of permutations, namely permutations which are products of non-interlacing descending cycles. The set of left divisors of δ_n turns out to have order equal to the n^{th} Catalan number, $(2n!)/(n!(n+1)!)$, whereas the left divisors of Δ have order $n!$. The fact that the order of the set of permutation braids is smaller in the new presentation than in the classical presentation had led to the hope by the authors of [24], when they first discovered the new presentation, that it would result in a polynomial algorithm for the conjugacy problem, but the Catalan numbers grow exponentially with index, and once again ($\star$) proved to be a fundamental obstacle. Thus, while the new presentation is extremely interesting in its own right, and does lead to faster word and conjugacy algorithms, the improvement in that regard does not address the fundamental underlying difficulties.

Open Problem 10 Curiously, it appears very likely that the classical and new presentations of $\mathbf{B}_n$ are the only positive presentations of $\mathbf{B}_n$ in which the Garside structure exists, although that has not been proved at this time, and is an interesting open problem. For partial results in this direction, see [94]. ♣

There is a different aspect of the dual presentations which we mention now, which involves a small detour. Before we can explain it, recall that one of our earliest definitions of the braid group $\mathbf{P}_n$, given in §1.1 was as the fundamental group $\pi_1(\mathcal{C}_{0,\hat{n}}, \vec{p})$, of the space formed from $\mathbb{C}^n$ by deleting the hyperplanes along which two or more complex coordinates coincide. Of course this gives a natural cell decomposition for $\mathcal{C}_{0,\hat{n}}$ as a union of (open) cells of real dimension $2n$. The braid group $\mathbf{B}_n$, as we defined it in §1.1, is the fundamental group of the quotient $\mathcal{C}_{0,n} = \mathcal{C}_{0,\hat{n}}/\Sigma_n$, where the symmetric group Σ_n acts on $\mathcal{C}_{0,\hat{n}}$ by permuting coordinates. In the interesting manuscript [69], Fox and Neuwirth used this natural cell decomposition of $\mathcal{C}_{0,\hat{n}}$ to find a presentation for $\mathbf{B}_n$, arriving at the classical presentation (2) for the braid group $\mathbf{B}_n$. See Section C of Chapter 10 of [43] for a succinct presentation of the results in [69]. Fox and Neuwirth also prove that $\mathcal{C}_{0,\hat{n}}$ is aspherical, and use this to give the first proof that $\mathbf{B}_n$ is torsion-free. See §6.4 below for a different and very easily understood proof of that same fact.

Around the same time that the first author, together with Ko and Lee, wrote the paper [23], which introduced the dual presentation of $\mathbf{B}_n$, Thomas Brady was thinking about other complexes which, like the one just described might serve as a $K(\pi, 1)$ for the group $\mathbf{B}_n$. See [36], which describes the construction of a finite CW-complex K_n of dimension

$n-1$ which is homotopy equivalent to $C_{0,n}$, and which is defined combinatorially, using the partial ordering on Σ_n, as described in §5.2 in the discussion of the lattice of simple words in $\mathbf{B}_n$. Brady then used his complex to determine a presentation for $\mathbf{B}_n$, arriving at the dual presentation (3).

As it turned out, the new presentation was important for other reasons too. The braid groups inherit a second Garside structure from the new presentation (3). The fact that the same was true for many other Garside groups played a major role in the discovery of the appropriate definitions. The Garside structure on the braid group arising from the new presentation ends up revealing even more structure: it is dual to that coming from the standard presentation, in the sense of an action on a complex and a dual complex. As it turned out, all finite type Artin groups of finite type also have dual 'Garside structures, as proved by Bessis [11] (this reference also contains the details of this dual structure). See also [37]. We also refer the reader to [121] for explicit presentations for the dual monoids associated to the Artin groups of finite type, and also for an interesting table that gives the number of simple elements defined by the left divisors of the Garside element in the classical and dual monoids for the finite type Artin groups.

Open Problem 11 This one is a very big set of problems. It is not known whether all Garside groups have dual presentations, in fact, it is also not known how many distinct Garside structures a given Garside group may have. ♣

5.4 Artin monoids and their groups

There is another class of groups which is intimately related to the braid group and its associated monoid, but it is much less well understood than the Garside groups. Let $S = \{u, \dots, \dots, t, \dots, v\}$ be a finite set. A <u>Coxeter graph</u> Γ over S is a graph whose vertices are in 1-1 correspondence with the elements of S. There are no edges joining a vertex s to itself. There may or may not be an edge joining a vertex s to a vertex $t \neq s$. Each pair of vertices $(s,t) = (t,s)$ is labeled by a non-negative integer $m(s,t)$. There are 2 types of labels: The label $m(s,t)$ is 2 if Γ does not have an edge that joins s and t, and it is $\in \{3, 4, \dots, \infty\}$ if there is an edge joining s and t.

The <u>Artin group</u> $A(\Gamma)$ associated to Γ has generators $\{\sigma_s \ \ s \in S\}$. There is a relation for each label $m(s,t) < \infty$, namely $\sigma_s\sigma_t\sigma_s\sigma_t \cdots = \sigma_t\sigma_s\sigma_t\sigma_s \cdots$, where there are $m(s,t)$ terms on each side, $2 \leq m(s,t) < \infty$. The <u>Coxeter group</u> $C(\Gamma)$ associated to the Artin group $A(\Gamma)$ is obtained by adding the relations $\sigma_s^2 = 1$ for every $s \in S$. As previously mentioned, we say that the group $A(\Gamma)$ has finite type if its associated Coxeter group is finite. The braid group $\mathbf{B}_n$ is an example of an Artin group of finite type; its associated Coxeter group being the symmetric group Σ_n. Note that by definition the defining relations in the presentations that we just described for Artin groups all preserve word length.

One of the many interesting properties of Artin groups is that for every Artin group there is an associated monoid A^+, and just as every Garside monoid embeds in its group, it was proved by Paris in [117] that every Artin monoid injects in its group. The injectivity property holds in the following strong sense too: If two positive words P, P' represent the same element of A, then they represent the same element of A^+. Paris's proof is completely different from Garside's proof of injectivity in the case of the braid group, which is the basis for the known proofs of the same fact for Garside groups. On the other hand, the other properties that are needed to obtain a Garside structure may or may not hold, for example it is definitely not true that every element in an Artin group can be written in the form NP, where N is negative and P is positive.

Remark 5.3 An interesting special case of Artin groups are the right-angled Artin groups. They are Artin groups in which the defining relations are all of the form $\sigma_s\sigma_t = \sigma_t\sigma_s$. For example, the right-angled Artin group associated to the braid group is defined by the presentation:

$$\langle \sigma_1, \ldots, \sigma_{n-1} \mid \sigma_i\sigma_k = \sigma_k\sigma_i \;\text{ if }\; |i-k| \geq 2 \,\rangle \tag{21}$$

Open Problem 12 In §4.5 we explained the fairly recent proof that the braid groups are linear. This leads one to ask, immediately, whether the same is true for its natural generalizations, e.g., Garside groups and Artin groups. It turns out that, like the braid groups, all right-angled Artin groups have faithful matrix representations [58], [50]. Of course any Artin group that injects as a subgroup of a related braid group is also linear, e.g. Artin groups of type B_n. It was proved recently by A. Cohen and D. Wales [47] and simultaneously by F. Digne [58] that all finite type Artin groups are linear. Going beyond that, the matter seems to be wide open and interesting. ♣

Remark 5.4 For reasons of space, we have not included any significant discussion of the vast literature on Artin groups and associated complexes on which they act. It is a pity to omit it, because it is a major area, and much of it had its origins in work on $\mathbf{B}_n$.

5.5 Braid groups and public key cryptography

The problem which is the focus of of 'public key cryptography' was mentioned, very briefly, in §1.4.5. The basic issue is how to send information, securely, over an insecure channel. The solution is always to use some sort of code whose main features are known to the sender and recipient, but which cannot be deduced by a viewer who lacks knowledge of the shared keys. To the best of our knowledge, all solutions to this problem rest on the same underlying idea: they make use of problems which have a precise answer, which is known to both the sender and recipient, but one which is deemed to be so difficult to compute that it is, in effect, unavailable to a viewer, even though the viewer has all the necessary data to deduce it. The earliest such schemes were based upon the difficulty of factorizing large integers into a product of primes. The individuals who wish to exchange data over a public system are denoted A and B. In a vastly oversimplified version, Both A and B have agreed, privately, on the choice of a prime number p. The sender chooses another prime q and transmits the product $n = pq$. A viewer may learn n, but because of the difficulty of factorizing n into primes cannot deduce p and q. The recipient, who knows both p and n, has no problem computing q.

A more recent approach is due to W. Diffie and M. Hellman [57]. As before, both A and B have agreed, privately, on the choice of a prime number p and a generator g of the finite cyclic group $\mathbb{Z}/p\mathbb{Z}$. A chooses a number a at random and computes $g^a(\text{mod } p)$, which she sends to B on the public channel. As for B he chooses a number b at random, computes $g^b((\text{mod } p)$ and sends in to A on the public channel. Since A now knows both g^b and a she can compute $(g^a)^b = g^{ab}(\text{mod})\ p$. Similarly, since B now knows both g^a and b, he can compute $(g^b)^a = g^{ba}(\text{mod } p) = g^{ab}(\text{mod } p)$. So both know $g^{ab}(\text{mod } p)$. As for the viewer, he knows both $g^a(\text{mod } p)$ and $g^b(\text{mod } p)$, however, because of the known difficulty of computing discrete logs, the crucial information $g^{ab}(\text{mod } p)$ is in effect unavailable to the viewer.

A rather different set of ideas was proposed in [3] and [92], and this is where the braid group comes in. The security of a system that is based upon braid groups relies upon the assumption that the word and conjugacy problems in the braid group have both been solved, but the conjugacy problem is computationally intractable whereas the word problem is not. However, as we have seen, that matter seems to be wide open at this moment. We have described the underlying mathematics behind the ElRifai-Morton solution to the word problem, and the best of the current solutions to the word and conjugacy problem, namely that of Gebhardt. The solution to the word problem is used in the conjugacy problem. The reason that the conjugacy problem is not polynomial is that we do not understand enough about the structure of the summit set, the super summit set and the ultra summit set. Our strong belief is that these matters will be settled. The assumption that our current lack of understanding of aspects of the mathematics of braids means that they cannot be understood seems unwarranted.

5.6 The Nielsen-Thurston approach to the conjugacy problem in $\mathbf{B}_n$

In this subsection we consider the conjugacy problem in $\mathbf{B}_n$ from a new point of view. We return again to the interpretation given in Theorem 1 of §1.3 of the braid group $\mathbf{B}_n$ as the mapping class group $\mathcal{M}_{0,1,n}$ of the punctured disc $S_{0,1,n} = D^2_n$. In this context, then, we emphasize that the term <u>braid</u> refers to a mapping class, that is, an isotopy class of diffeomorphisms.

The Nielsen-Thurston classification of mapping classes of a surface is probably the single most important advance in this theory in the last century, and we review it here. For simplicity, we shall focus for now on the mapping class group of a closed surface $\mathcal{M}_g = \mathcal{M}_{g,0,0}$ in the notation of §1.3. We note that the groups $\mathcal{M}_g$ have trivial center for all $g \geq 3$.

We make the tentative definition that an element of $\mathcal{M}_g$ is 'reducible' if it preserves up to isotopy a family of disjoint nontrivial curves on the punctured disc. Such a family of curves is known as a <u>reduction system</u>. (Throughout this discussion, we use 'curve' to refer to the isotopy class of a curve. By 'disjoint curves', we mean two distinct isotopy classes of curves with respective representatives which are disjoint.) It was proved in [26] that for a reducible map ϕ, there exists a <u>essential reduction system</u>, denoted $\mathrm{ERS}(\phi)$. A curve $c \in \mathrm{ERS}(\phi)$ if and only if

1. There exists an integer k such that $\phi^k(c) = c$.

2. If a curve x has nonzero geometric intersection with c, then $\phi^m(x) \neq x$ for all integers m. In particular, any curve x which has nonzero geometric intersection with a curve $c \in \mathrm{ERS}(\phi)$ is not in $\mathrm{ERS}(\phi)$.

It is proved in [26] that the curves in $\mathrm{ERS}(\phi)$ are contained in every reducing system for ϕ, and are a minimal reduction system for ϕ. Also, the system of curves $\mathrm{ERS}(\phi)$ is unique. Keeping all this in mind, we now define an element $\phi \in \mathcal{M}_g$ to be <u>reducible</u> if it fixes the curves in an essential reduction system, i.e. in $\mathrm{ERS}(\phi)$.

A mapping class $\phi \in \mathcal{M}_g$ is periodic if $\phi^k = 1$ in $\mathcal{M}_g$.

We note that a reducible mapping class always contains a representative diffeomorphism which fixes a given reduction system setwise. Likewise, Nielsen showed that a periodic mapping class always contains a representative which is periodic as a diffeomorphism. Finally, a mapping class which is neither reducible nor periodic is pseudo-Anosov (abbreviated as PA).

Braid groups, as the mapping class group of a punctured disc, admit a similar classification. The group $\mathbf{B}_n$ is torsion-free, but by analogy with the above case of mapping class groups of closed surfaces of genus at least two, which are centerless, we say that an element $\omega \in \mathbf{B}_n$ is periodic if for some integer k, ω^k is isotopic to a full Dehn twist on the boundary of the disc D^2_n. In terms of generators and relations this is equivalent to saying that the braid is a root of Δ^2, where Δ is the Garside braid. (Recall from §5.1 that the square of Δ_n generates the center of $\mathbf{B}_n$.)

The definitions of reducible braids and pseudo-Anosov (PA) braids require no alteration other than replacing the surface S_g with the punctured disc.

Remark 5.5 It is clear for $\mathbf{B}_n$ and $\mathcal{M}_g$ alike, we also have a classification of elements as reducible, periodic and pseudo-Anosov up to conjugacy.

Each of these three possibilities reveals new structure, so we consider them one at a time. A major reference for us is the paper [8], which gives an algorithm for finding a system of reducing curves if they exist, and for recognizing periodic braids. This paper did not receive much attention at the time that it was written because it had the misfortune to be written simultaneously and independently with the ground-breaking and much more general papers of Bestvina and Handel [12]. However, as is often the case, one learns very different things by examining a particular case of a phenomenon in detail, and by proving a broad generalization of the same phenomenon, and that is what happened here.

Periodic braids. The first author learned from [8] that the classification of periodic braids had been solved via the work of Kerekjarto (1919) and Eilenberg (1935), who proved that up to conjugacy a periodic braid is a power of either δ or α, where $\delta = \sigma_1\sigma_2\ldots\sigma_{n-1}$ and $\alpha = \sigma_1\sigma_2\ldots\sigma_{n-1}\sigma_1$. One verifies, using the elementary braid relations in the classical presentation, that $\delta^n = \alpha^{n-1} = \Delta^2$. To visualize the assertion for δ geometrically, place the punctures at polar angles $2\pi k/n$ around a circle of radius $r < 1$, and think of δ as a rotation of D^2_n of angle $2\pi/n$. To visualize α, do the same, only now place one of the punctures at the origin and arrange the remaining $n-1$ punctures symmetrically at polar angles $2\pi k/(n-1)$.

Reducible braids. Let $\omega = \Delta^I L_1 L_2 \cdots L_s$ be a reducible mapping class in left greedy normal form, as described in §5.1. Our initial model for D^2_n will be the unit disc, with the n punctures arranged along the real axis, and placed symmetrically so that they divide the interval $[-1,1] \subset \mathbb{R}$ into $n+1$ equal line segments. It is shown in [8] that one can choose a braid ω' conjugate to ω and in the super summit set of ω which fixes the simplest possible family of closed curves, namely a family C of geometric ellipses whose centers are on a horizonal 'axis' through the punctures, and which are chosen in such a way that the axis bisects the discs that the curves in CRS(ω') bound.

Open Problem 13 Is it true that, if a reducible braid is in its ultra summit set, then its invariant multicurves are can always be chosen to be geometric ellipses? ♣

The remaining braiding may then be thought of as going on inside tubes, also the tubes may braid with other tubes. With this model it should be intuitively clear that the choice of such a representative for the conjugacy class of a braid which permutes the tubes in the required fashion may require very different choices (again up to conjugacy) for the braids which are inside the tubes, and that the sensible way to approach the problem is to cut the initial disc open along the reducing curves and focus on the periodic or PA maps inside the tubes. This is, of course, a very different approach from the one that was considered in §5.1 above, where no such considerations entered the picture.

Pseudo-Anosov braids. In the PA case (which is the generic case), there is additional structure, and now our description becomes very incomplete. In this case [128] there exist two projective measured foliations $\mathcal{F}^u$ and $\mathcal{F}^s$, which are preserved by an appropriate representative w of $\{\omega\}$. Moreover the action of w on $\mathcal{F}^u$ (the unstable foliation) scales its measure by a real number $\lambda > 1$, whereas the action on $\mathcal{F}^s$ (the stable foliation) scales its measure by $1/\lambda$. These two foliations and the scaling factor λ are uniquely determined by the conjugacy class of $\{\omega\}$, however the triplet $(\mathcal{F}^u, \mathcal{F}^s, \lambda)$ does not determine $\{\omega\}$. To explain the missing pieces, we replace the invariant foliations by an invariant 'train track', with weights associated to the various branches.

In the case of pseudo-Anosov mapping classes acting on a once-punctured surface, a method for enumerating the train tracks is in Lee Mosher's unpublished PhD thesis [112]. But Mosher's work was incomplete and remained unpublished for many years, even as many of the ideas in it were developed and even expanded, leading to some confusions in the literature about exactly what is known and what remains open. At this writing a complete solution to the conjugacy problem for braid groups or more generally surface mapping class groups, based upon the Nielsen-Thurston machinery, does not exist in the literature. Mosher has a partially completed monograph in preparation, *Train track expansions of measured foliations*, which promises to give such a solution, but for more general surfaces, i.e. for $S_{g,b,n}$ very little seems to be known. We therefore pose it as an open problem:

Open Problem 14 Investigate the conjugacy problem in the mapping class groups $\mathcal{M}_{g,b,n}$ with the goals of (i) pinning down precisely what is known for various triplets (g, b, n); (ii) describing all cases in which there is a complete solution; and (iii) describing what remains to be done in the simplest cases, that is $\mathcal{M}_{g,0,0}$ and $\mathcal{M}_{g,0,1}$. ♣

Essentially nothing is known, at this writing, about the interface between the dynamic and combinatorial solutions to the conjugacy problem in the braid groups, and still less in the general case of the more general mapping class groups $\mathcal{M}_{g,b,n}$, to which the entire Thurston machinery applies. We know of nothing which even hints at related dynamic structures in Artin groups or Garside groups.

Open Problem 15 Does the Garside approach to the conjugacy problem in $\mathcal{M}_{0,1,n}$ generalize to a related approach to the problem in $\mathcal{M}_{g,1,0}$ for any $g > 0$? ♣

Open Problem 16 Does the Nielsen-Thurston approach to the conjugacy problem in $\mathcal{M}_{0,1,n}$ generalize to a related approach to the problem in any other Artin group? ♣

5.7 Other solutions to the word problem

We review, very briefly, other ways that the word and conjugacy problems have been solved in the braid groups. We restrict ourselves to results which revealed aspects of the structure of $\mathbf{B}_n$ that has had major implications for our understanding of braid groups, even when the implications for the word and/or conjugacy problems fall short of that criterion.

1. Artin's solution to the word problem. The earliest solution to the word problem was discovered in 1925 by Artin, in his first paper on braids [6]. It was based upon his analysis of the structure of the pure braid group, which we described in §1.2, and in particular in the defining relations 4 and the sequence given just below it:

$$\{1\} \to \mathbf{F}_{n-1} \to \mathbf{P}_n \xrightarrow{\pi_n^\star} \mathbf{P}_{n-1} \to \{1\}.$$

The resulting normal form, described in [6] as <u>combing</u> a braid, is well known. We refer the reader to Artin's original paper for a very beautiful example. In spite of much effort over the years nobody has managed, to this day, to use related techniques to solve the conjugacy problem in $\mathbf{B}_n$, except in very special cases.

2. The Lawrence-Krammer representation. See §4 above. When a group admits a faithful matrix representation, there exists a fast way to solve the word problem. It is interesting to note that Krammer's first proof of linearity, in the case of $\mathbf{B}_4$, used the solution to the word problem which came from the presentation (3) in §1 to this paper. See [97]. Also, when he proved linearity in the general case, in [98], he gives two proofs. The second proof, Theorem 6.1 of [98], shows that the Lawrence-Krammer matrices detect the infimum and supremum of a braid. It follows as an immediate corollary (noted in [98]) that the representation is faithful, because any braid which has non-zero infimum and supremum cannot be the identity braid.

As regards the conjugacy problem, there is a difficulty. The image of $\mathbf{B}_n$ under the isomorphism given by the Lawrence-Krammer representation yields a group $\mathcal{B}_n$ which is a subgroup of the general linear group $GL_m(\mathbb{Z}[q^{\pm 1}, t^{\pm 1}])$, where $m = (n)(n-1)/2$. To the best of our knowledge it is unknown at this time how to describe $\mathcal{B}_n$ in any way that allows one to test membership in $\mathcal{B}_n$. Lacking such a test, any class invariants which one finds in this way will be limited in usefulness, because if $\omega \in \mathbf{B}_n$, there will be no way to distinguish between class invariants which arise from conjugation by elements in $\mathcal{B}_n$ from those which arise from conjugation by more general elements of $GL_m(\mathbb{Z}[q^{\pm 1}, t^{\pm 1}])$ from invariants thus one cannot hope for a complete solution.

Open Problem 17 Investigate the eigenvalues and the trace of the Lawrence-Krammer matrices. ♣

3. The Dehornoy ordering. A group or a monoid has a <u>right-invariant</u> (resp. <u>left-invariant</u>) ordering if there exists a strict linear ordering of its elements, denoted $<$, with the property that if $f, g, h \in G$, then $f < g$ implies $fh < gh$ (resp. $hf < hg$). To the best of our knowledge nobody had considered the question of whether $\mathbf{B}_n$ had this property before 1982, when Patrick Dehornoy announced his discovery that the groups $\mathbf{B}_n$ admit such an ordering, and that it can be chosen to be either right-invariant or left-invariant, but not both. In the 5-author paper [67] the Dehornoy ordering was shown to have the following topological meaning. We restrict to the the right-invariant case, the two results being essentially identical. We regard $\mathbf{B}_n$ as $\mathcal{M}_{0,1,n}$, where the surface $S_{0,1,n}$ is the unit disc in the complex plane and the punctures lie on the real axis in the interval $(-1,1)$. The punctures divide $[-1,1] \subset \mathbb{R}$ into $n+1$ segments which we label $\xi_1, \xi_2, \ldots, \xi_{n+1}$ in order, where ξ_1 joins $\{-1\}$ to the first puncture. Choose $[\omega_1], [\omega_2] \in \mathbf{B}_n$, and representatives ω_1, ω_2 with the property that $\omega_1([-1,1])$ and $\omega([-1,1])$ intersect minimally. Note that $\omega_k([-1,1])$ divides $S_{0,1,n}$ into two halves, and it makes sense to talk about the upper and lower half of $\omega_2([-1,1])$ because $\{-1\}$ and $\{1\}$ are on $\partial S_{0,1,n}$ and $\omega_k \mid \partial S_{0,1,n}$ is the identity map. Then $[\omega_2] > [\omega_1]$ if $\omega_1(\xi_i) = \omega_2(\xi_i)$ for $i = 1, \ldots, j-1$ and an initial segment of $\omega_2(\xi_j)$ lies in the upper component of $S_{0,1,n} - \omega_1([-1,1])$.

It was proved in [131] that the resulting left-invariant ordering of $\mathbf{B}_n$ extends the ElRifai-Morton left partial ordering that we described in §5.2 above. It is proved in [54] that there are infinitely many other left orderings.

The subject blossomed after the paper [67] appeared. In fact so much work resulted that there is now a 4-author monograph on the subject [54], written by some of those who were the main contributors, containing an excellent review of what has been learned during the 10 years since the initial discovery. We do not wish to repeat what is readily available elsewhere, especially because we are not experts, so we point the reader to Chapter 3 of [54], where it is shown that the Dehornoy ordering leads to a solution to the word problem. Unfortunately, the solution so-obtained is exponential in $|\omega|$, whereas the solution described in (G1)-(G6) above is quadratic in $|\omega|$.

4. $\mathbf{B}_\mathrm{n}$ as a subgroup of Aut ($\mathbf{F}_\mathrm{n}$). It is well known that $\mathbf{B}_n$ has a faithful representation as a subgroup of $\mathrm{Aut}(\mathrm{F_n})$, for example see [18] for a proof. This of course gives a solution to the word problem, but (like the Dehornoy solution) it is exponential in $|\omega|$. On the other hand, it is extremely interesting that the Nielsen-Thurston machinery, described in §5.6 above, generalizes to the entire group $\mathrm{Aut}(\mathrm{F_n})$, giving yet another instance where the braid group is at the intersection of two rather different parts of mathematics, and could be said to have pointed the way to structure in the second based upon known structure in the first.

In this section we have described several solutions to the word problem, and noted that one of them (the modified Garside approach of (G1)-(G6)) is $\mathcal{O}(|\omega|^2(n)(log(n))$, where $|\omega|$ is the letter length of two arbitrary representatives of elements of $\mathbf{B}_n$, using the classical presentation (2) for $\mathbf{B}_n$. The word problem seems to be one short step away from the <u>non-minimal braid</u> problem: given a word ω in the generators $\sigma_1, \ldots, \sigma_{n-1}$ and their inverses, determine whether there is a shorter word ω' in the same generators which represents the same element of $\mathbf{B}_n$? For, it is clear that a decision process exists: list all the words that are shorter than the given one, thin the list by eliminating as many candidates as possible by simple criteria such as preserving length mod 2, an obvious invariant, and then test, one by one, whether the survivors represent the same element of $\mathbf{B}_n$ as the given word ω? It therefore seemed totally surprising to us that in 1991 M. S. Patterson and A. A. Razborov proved:

Theorem 18 [119] : *The non-minimal braid problem is NP complete.*

Thus, if one could find an algorithm to decide whether a given word ω is non-minimal, and if the algorithm was polynomial in $|\omega|$, one would have proved that $P = NP$! To the best of our knowledge, essentially nothing has been done on this problem. In this regard we suggest two research problems:

Open Problem 18 The proof that is given in [119] is very specific to the classical presentation (2) for $\mathbf{B}_n$. Can it be adapted to the new presentation? To other presentations? ♣

Open Problem 19 Investigate the shortest word problem in the braid group $\mathbf{B}_n$, using the classical presentation. ♣

In §3.1 we discussed and proved the Markov Theorem Without Stabilization in the special case of the unknot. See Theorem 5, which asserts that there is a complexity measure on closed braid representatives of the unknot, and using it a sequence of strictly complexity-reducing destabilizations and exchange moves that reduce a closed braid diagram for the unknot to a round planar circle. In [22], theorem 5 was used to develop an algorithm for unknot recognition. In [32] the first steps were taken to develop a computer program to realize the resulting algorithm. The algorithm is far from being practical in even simple cases, however we are now in a position to describe three open problems, all closely related to (G1)-(G-11) above, which would lead to an effective solution to the unknot recognition problem:

Open Problem 20 Develop an algorithm that will detect when the conjugacy class of a closed braid admits a destabilization. ♣

An n- braid W admits an exchange move if it is conjugate to a braid of the form $U\sigma_{n-1}V\sigma_{n-1}^{-1}$, where U and V depend only on $\sigma_1, \ldots, \sigma_{n-2}$. Thus, up to conjugation, W is a product of two reducible braids, one positively reducible and the other negatively reducible. By Theorem 5, it may be necessary to modify the given conjugacy class by exchange moves in order to jump from the given class to a new class (if it exists) which admits a destabilization. Thus there is an unknown complexity measure on conjugacy classes, with the ones which admit destabilizations being especially nice. This leads us to our second open problem.

Open Problem 21 Develop an algorithm that will detect when the conjugacy class of a closed braid admits an exchange move, and when a collection of classes are exchange-equivalent. ♣

This would still not be a complete tool, because exchange moves can be either be complexity-reducing or complexity-increasing, and unless we can tell the difference (that is the content of open problem 3 below) we are left with the option of trying a sequence of exchange moves of unpredictable length in our search for destabilizations.

Open Problem 22 The complexity measure that was introduced in §3.1 is 'hidden' in the braid foliation of an incompressible surface whose boundary is the given knot. This is a highly non-trivial matter, because if we had the foliated surface in hand, then we would be able to compute its Euler characteristic and would know whether we had the unknot. On the other hand, the braid foliation determines the embedding of its boundary (see Theorem 4.1 of [21]), so there is no essential obstacle to 'seeing' the complexity measure in the given braid. That is the essence of our third open problem, which asks that we recognize how to translate the complexity measure that was used in the proof of Theorem 5 into a complexity measure on closed braids which will be able to distinguish exchange moves that reduce complexity from those which do not. ♣

6 A potpourri of miscellaneous results

This section is for leftovers— topics on which there have been interesting new discoveries which did not seem to fit well anywhere else.

6.1 Centralizers of braids and roots of braids

The mixed braid groups are defined in [83] to be the braids which preserve a given partition of the puncture points on the disc. In [83] Gonzales-Meneses and Wiest gave full descriptions of the centralizer of a braid in terms of semi-direct and direct products of mixed braid groups, and found sharp bounds on the number of generators of the centralizer of a braid.

In a related paper [82] Gonzales-Meneses proved that, up to conjugacy, braids have unique roots. That is, if $\omega \in \mathbf{B}_n$ and if $\alpha, \beta \in \mathbf{B_n}$ have the property $\alpha^k = \beta^k = \omega$, then $\{\alpha\} = \{\beta\}$.

6.2 Singular braids, the singular braid monoid, and the desingularization map

The singular braid monoid $\mathbf{SB}_n$ is a monoid extension of the braid group $\mathbf{B}_n$. It was introduced in [7] and, simultaneously and independently, in [20]. Its definition was suggested by the mathematics which revolved about Vassiliev invariants of knots and links. To define it, we need to describe a presentation for $\mathbf{SB}_n$, taken from [20]. There are three types of generators, which we call σ_i, σ_i^{-1} and τ_i, $1 \leq i \leq n-1$. Here the $\sigma_i^{\pm 1}$ are to be thought of as the classical positive and negative elementary braids and the τ_i are to be thought of as elementary singular braids. The braid τ_i is obtained from σ_i (see Figure 24 (ii)) by identifying strands i and $i+1$ as they cross. Defining relations in $\mathbf{SB}_n$ are:

$$\sigma_i\sigma_i^{-1} = \sigma_i^{-1}\sigma_i = 1, \quad \sigma_i\tau_i = \tau_i\sigma_i,$$

$$\sigma_i\sigma_j = \sigma_j\sigma_i, \quad \sigma_i\tau_j = \tau_j\sigma_i, \quad \tau_i\tau_j = \tau_j\tau_i \qquad \text{if} \qquad |i-j| \geq 2$$

$$\sigma_i\sigma_j\sigma_i = \sigma_j\sigma_i\sigma_j, \qquad \sigma_i\sigma_j\tau_i = \tau_i\sigma_i\sigma_j \qquad \text{if} \qquad |i-j| = 1.$$

The desingularization map is a homomorphism from $\mathbf{SB}_n$ to the group ring $\mathbb{Z}\mathbf{B}_n$ of the braid group, defined by $\psi(\sigma_i^{\pm 1}) = \sigma_i^{\pm 1}, \qquad \psi(\tau_i) = \sigma_i - \sigma_i^{-1}$. Birman used this map to develop the relationship between Vassiliev invariants and quantum groups'. It was conjectured in [20] that the map ψ is an embedding. After several proofs of special cases of the conjecture, it was settled by Luis Paris in the affirmative in 2003 [118].

During the 10 year interval after the introduction of the singular braid monoid, and before the proof of the embedding theorem, it came as quite a surprise when it was discovered that there was a new group on the scene–the singular braid group of [71]. To this day we are unsure of its significance, although its existence is unquestioned! It is most easily defined via generators and relations, starting with the presentation that we just gave for the singular braid monoid, and then adding one new generator $\bar{\tau}$ (to suggest that it behaves the way that the inverse of τ ought to behave). Defining relations for $\mathbf{GB}_n$ are all the relations in $\mathbf{SB}_n$, plus ones satisfied by the new generator. The latter are 'monoid relations' which are the same as those in the singular braid monoid, but substituting $\bar{\tau}$ for τ, and two additional relations, namely $\tau\bar{\tau} = \bar{\tau}\tau = 1$. Pictorially, one has two types of singular crossings, and they annihilate one another. The main result in [71] is that $\mathbf{SB}_n$ embeds in $\mathbf{GB}_n$.

6.3 The Tits conjecture

The Tits conjecture is very easy to state, in its simplest form. Consider the classical presentation of the braid group $\mathbf{B}_n$, i.e. the presentation (2). The Tits conjecture, in the special case of the braid group, is that the subgroup G of $\mathbf{B}_n$ generated by the elements $T_i = \sigma_i^2$ has the presentation $\langle T_1, \ldots, T_{n-1} \mid T_iT_j = T_jT_i \text{ if } |i-j| \geq 2\rangle$. A generalized version of the conjecture (the generalization relates to the arbitrary choices of the powers) was proved by Crisp and Paris in 2001, settling a question which had plagued the experts for many years:

Theorem 19 [49] *Let S be a finite set, let Γ be a Coxeter graph over S, and let $A(\Gamma)$ be the associated Artin group, as defined in §5.4, with generating set $\{\sigma_s, \ \ s \in S\}$. Associate further to each $s \in S$ an integer $m_s \geq 2$. Consider the subgroup G of $A(\Gamma)$ generated by the elements $\{T_{s,m_s} = \sigma_s^{m_s}\}$. Then defining relations among the generators of the subgroup G are 'the obvious ones', namely that T_{s,m_s} and T_{t,m_t} commute in G if and only if they commute in $A(\Gamma)$. No other relations are needed.*

The proof is interesting, because it introduces a technique which is very closely related to the themes that we have explored in this article. The basic idea is that, for a key subclass of Artin groups there is a representation f of $A(\Gamma)$ into the mapping class group $\mathcal{M}(S)$ of a connected surface S with boundary that is associated to the graph Γ, which induces an action of $A(\Gamma)$ on a monoid determined by Γ. Since the group $H(\Gamma)$ which is presented in the theorem has an obvious homomorphism onto $A(\Gamma)$, and the proof shows that the restriction of the action to $H(\Gamma)$ gives the desired isomorphism.

6.4 Braid groups are torsion-free: a new proof

As we stated in § 1, it was necessary to make choices in the writing of this review, and our decision was to be guided by the principle of focusing on new results or new proofs of known results. During the years since [18] was written, many people have written to the first author with questions about braids, and a regular question has been "Isn't there a *simple* proof that the braid groups have no elements of finite order?" It is therefore fitting that we end this review with exactly that – a beautiful simple proof, based upon the discovery, due to Dehornoy, that the braid groups admit a left-invariant ordering:

Theorem 20 *The groups $\mathbf{B}_n, \ n = 1, 2, 3, \ldots$ are all torsion free.*

Proof: See §5.7 for a discussion of the Dehornoy left-invariant ordering of the braid groups. Choose any element $g \in \mathbf{B}_n, \ \ g \neq 1$. Replacing g if necessary with g^{-1} we may assume that $1 < g$. Since the ordering is left-invariant we then have that $g < g^2$ and $g^{-1} < 1$. Iterating, $\cdots < g^{-3} < g^{-2} < g^{-1} < 1 < g < g^2 < g^3 < \cdots$. ‖

Appendix: Computer programs

In 2004, it is almost as important to know about computer tools as it is to have a guide to the literature, so we supplement our bibliography with a guide to the computer tools that we know about and which have been useful to us and colleagues.

Changing knots and links to closed braids. Vogel's proof of his method for changing arbitrary knot diagrams to closed braid diagrams is ideal for computer programming. We refer the reader to the URL http://www.layer8.co.uk/maths/braids/, for a program, due to Andrew Bartholomew and Roger Fenn, which does this and much more.

Garside's algorithm for the word and conjugacy problems. Many people have programmed Garside's algorithm for the word and conjugacy problem, but the one we know best is the program of Juan Gonzalez-Meneses, which can be downloaded from http:www.personal.us.es/meneses. The very robust version that the reader will find there computes Garside's normal forms for braids, and the ultra-summit set of a braid, a complete invariant of conjugacy.

Nielsen-Thurston classification of mapping classes in $\mathcal{M}_{0,n+1}$. We know of two very useful computer programs, all based upon the Bestvina-Handel algorithm [12]. Both assume that the boundary of the n-times punctured disc has been capped with a disc, so that they compute in the mapping class group of $\mathbf{B}_n/$ modulo its center, rather than in $\mathbf{B}_n$. Equivalently, they work with the mapping class group of an $(n+1)$-times punctured sphere, where admissible maps fix the distinguished point.

The first, due to W. Menasco and J. Ringland, can be downloaded from http://orange.math.buffalo.edu/software.html, by following the link to "BH2.1 An Implementation of the Bestvina-Handel Algorithm". The second, due to T. Hall, can be downloaded from http://www.liv.ac.uk/Maths/pure and following the links to the research group on Dynamical Systems, and then to the home page of Toby Hall. Both determine whether an input map is pseudo-Anosov, reducible or finite order and both find an invariant train track and a train track map (which makes it possible to calculate essential dynamics in the class by Markov partition techniques). In the reducible case Hall's program provides a set of reducing curves. Recently Hall updated his program to adapt it to MacIntosh OS-X computers, versions 10.2 and above.

Other software. We mention M. Thistlethwaite's Knotscape, because the program has a reputation for being very versatile and user-friendly. It is available for download from http://www.math.utk.edu∼morwen. It accepts as input knots that are defined as closed braids (and also knots defined by the Dowker code or mouse-drawn diagrams), and locates it in the tables if it has at most 16 crossings. The program computes numerous invariants, including the Alexander, Jones, Homfly and Kauffman polynomials, and hyperbolic invariants (assuming that the knot is hyperbolic). The hyperbolic routines in Knotscape were taken with permission from Jeff Weeks's program SnapPea; the procedures for calculating polynomials were supplied by Bruce Ewing and Ken Millett, and the procedure for producing a knot picture from Dowker code is part of Ken Stephenson's Circlepack program.

References

[1] S.I. Adyan, *Fragments of the word Δ in a braid group*, Mat. Zametki **36**, No. 1 (1984), 25-34. English translation Math. Notes **36** (1984), No. 1-2, 505-510.

[2] J. Alexander, *A lemma on a system of knotted curves*, Proc. Nat. Acad. Sci. USA. **9** (1923), 93-95.

[3] I. Anshel, M. Anshel and D. Goldfeld, *An algebraic method for public key crypography*, Math research letters **6**, No. 3-4 (1999), 287-291.

[4] V. Arnold, *On braids of algebraic functions and cohomologies of swallowtails*, Uspekhi Mat. Nauk. **23** (1968), 247-248.

[5] ___________, *Topological invariants of plane curves and caustics*, AMS University Lecture Series **5** (1994).

[6] E. Artin, *Theorie der Zopfe*, Abh. Math. Sem. Hamburg, **4** (1925), 47-72.

[7] J. Baez, *Link invariants of finite type and perturbation theory*, Lett. Math. Phys. **26** (1992), 43-51.

[8] D. Benardete, M. Guitierrez and Z. Nitecki, *Braids and the Nielsen-Thurston classification*, J. Knot Theory and its Ramifications **4** (1995), 549-618.

[9] D. Bennequin, *Entrelacements et équations de Pfaff*, Asterisque **107-108** (1983), 87-161. English version: Russian Math.Surveys, **44**, No. 3 (1989), 1-65.

[10] A. Berrick, F. Cohen, Y. Wong and J.Wu, *Configurations, braids and homotopy theory*, preprint Jan. 2003.

[11] D. Bessis, *The dual braid monoid*, Ann. Sci. Ecole Normal Sup (4), **36**, No. 5 (2003), 647-683.

[12] M. Bestvina and M. Handel, *Train tracks for surface automorphisms*, Topology **34** (1995), 109-140.

[13] S. Bigelow, *The Burau representation of the braid group B_n is not faithful for $n = 5$*, Geometry and Topology **3** (1999), 397-404.

[14] ___________, *Braid groups are linear*, J. Amer. Math. Soc. **14** No. 2, (2001), 471-486.

[15] ___________, *Representations of braid groups*, Proceedings of the ICM, Vol. II, Higher Ed. Press, Beijing (2002), 37-45.

[16] S. Bigelow and R. Budney, *The mapping class group of a genus 2 surface is linear*, Algebraic and Geometric Topology, **1** (2001), 699-708.

[17] J. Birman, *Mapping class groups and their relationship to braid groups*, Comm. Pure and Appl. Math. **22** (1969), 213–238.

[18] ___________, *BRAIDS, LINKS AND MAPPING CLASS GROUPS*, Annals of Math. Studies **82** (1974).

[19] ___________, *The work of Vaughan F.R. Jones*, Proc. Int. Congress of Mathematicians, Kyoto, Japan, Mathematical Society of Japan and Springer-Verlag (1991), 9-18.

[20] ___________, *New points of view in knot theory*, Bull. AMS **28** (1993), 253-287.

[21] J. Birman and E. Finkelstein, *Studying surfaces via closed braids*, J. of Knot Theory and Its Ramifications **7**, No. 3 (1998), 267-334.

[22] J. Birman and M. Hirsch, *A new algorithm for recognizing the unknot*, Geometry and Topology **2** (1998), 175-220.

[23] J. Birman, K.Y.Ko and S.J.Lee, *A new approach to the word and conjugacy problems in the braid groups*, Adv. Math. **139**, No. 2, (1998), 322-353.

[24] ___________, *The infimum, supremum and geodesic length of a braid conjugacy class*, Adv. Math. (2001), **164**, No. 1, (2001), 41-56.

[25] J. Birman, D. Long, J. Moody, *Finite-dimensional representations of Artin's braid group*, in "THE MATHEMATCAL LEGACY OF WILHELM MAGNUS'", Contemporary Math **169** (1994), 123-132.

[26] J. Birman, A. Lubotzky and J. McCarthy, *Abelian and solvable subgroups of the mapping class groups*, Duke Math. J. **50** (1983), 1107-1120.

[27] J. Birman and W. Menasco, *Studying links via closed braids III: Classifying links which are closed 3-braids*, Pac. J. Math. **161**, No. 1 (1993), 25-113.

[28] ___________, *Studying links via closed braids IV: Split links and composite links*, Inventiones mathematicae **102** (1990), 115-139.

[29] ___________, *Studying links via closed braids V: The unlink*, Trans. AMS **329**, No. 2 (1992), 585-606.

[30] ___________, *Stabilization in the braid groups I: MTWS*, preprint, submitted for publication, revised July 2003, arXiv math.GT/0310279

[31] ___________, *Stabilization in the braid groups II: On transversal simplicity of transversal knots*, preprint, submitted for publication, revised October 2003, arXiv math.GT/0310280

[32] J. Birman, M. Rampichini, P. Boldi, and S. Vigna, *Towards an implementation of the B-H algorithm for recognizing the unknot*, J. of Knot Theory and Its Ramifications **11**, No. 4 (2002), 601-645.

[33] J. Birman and H. Wenzl, *Braids, link polynomials, and a new algebra*, Trans. AMS **313**, No. 1 (1989), 249-273.

[34] J. Birman and N. Wrinkle, *On transversally simple knots*, J. of Differential Geometry **55** (2000), 325-354.

[35] N. Bourbaki, *GROUPS ET ALGEBRES DE LIE*, Chapitres 4,5,6, Hermann, Paris 1968.

[36] T. Brady, *A partial order on the Symmetric Group and new $K(\pi, 1)'s$ for the braid groups*, Adv. Math **161** (2001), 20-40.

[37] T. Brady and C. Watt, *$K(\pi, 1)$'s for Artin groups of finite type*, Geom. Dedicata **94** (2002), 225–250.

[38] G. Bredon, *TOPOLOGY AND GEOMETRY*, Springer-Verlag Graduate Texts in Mathematics **139** (1991).

[39] T. Brendle and H. Hamidi-Tehrani, *On the linearity problem for mapping class groups*, Algebraic and Geometric Topology **1** (2001), 445-468.

[40] E. Brieskorn and K. Saito, *Artin-Gruppen und Coxeter-Gruppen*, Invent. Math. **17** (1972), 245-272.

[41] R. Budney, *On the image of the Lawrence-Krammer representation*, arXiv:math.GT/0202246.

[42] W. Burau, *Über Zopfgruppen und gleischsinning verdrillte Verkettungen*, Abh. Math. Sem. Ham. II (1936), 171-178.

[43] G. Burde and H. Zieschang, KNOTS, de Gruyter, Berlin and New York 1985

[44] R. Charney, *Geodesic automation and growth functions for Artin groups of finite type*, Math.Ann. **301** (1995), 307-324.

[45] Y. Chekanov, *Differential algebra of Legendrian knots*, Invent. Math. **150** (2002), NO. 3, 441-483.

[46] A. Clifford and G. Preston, *THE ALGEBRAIC THEORY OF SEMIGROUPS*, AMS Surveys **7** (1961). Volume 1.

[47] A. Cohen and D. Wales, *Linearity of Artin groups of finite type*, Israel J. Math. **131** (2002), 101–123.

[48] F. Cohen and J. Wu, *On braid groups, free groups, and the loop space of the* 2*-sphere*, Progress in Mathematics, **215**(2003), 93-105.

[49] J. Crisp and L. Paris, *The solution to a conjecture of Tits on the subgroup generated by the squares of the generators of an Artin group*, Invent. Math. **145** (2001), 19-36.

[50] M. Davis and T. Januszkiewicz, *Right- angled Artin groups are commensurable with right-angled Coxeter groups*, J. Pure Appl. Algebra **153** (2000), no. 3, 229–235.

[51] M. Dehn, *Über unendliche discontinuerliche Gruppen*, Math. Ann. **69** (1911), 116-144.

[52] P. Dehornoy, *A fast way of comparing braids*, Adv. in Math. **125** (1997), 200-235.

[53] ___________, *Groupes de Garside*, Ann. Scient. Ec. Norm. Sup. **35** (2002), 267-306.

[54] ______, I. Dynnikov, D. Rolfsen and B. Wiest, *WHY ARE BRAIDS ORDERABLE?*, Panoramas et Synthéses **14**, Sociètè Mathèmatique de France 2002.

[55] ______ and L. Paris, *Gaussian groups and Garside groups, two generalizations of Artin groups*, Proc. London Math. Soc. **79**(1999), No. 3, 569-604.

[56] P. Deligne, *Les immeubles des groupes de tresses generalises*, Invent. Math. **17** (1972), 273-302.

[57] W. Diffie and M. Hellman, *New directions in cryptography*, IEEE Trans. in Information Theory **IT-22**, No. 6 (June 1977), 74-84.

[58] F. Digne, *On the linearity of Artin braid groups*, J. Algebra **268** (2003), no. 1, 39-57.

[59] Y. Eliashberg, *Classification of overtwisted contact structures on 3-manifolds*, Invent. Math. **98** (1989), 623-637.

[60] Y. Eliashberg, *Legendrian and transversal knots in tight contact 3-manifolds*, Topological methods in Modern Mathematics, Proceedings of a Symposium at SUNY Stonybrook in 1991, Publish or Perish, Houston, Texas (1993)

[61] E. ElRifai and H. Morton, *Algorithms for positive braids*, Quart.J. Math. Oxford Ser (2), **45** (180) (1994), 479-497.

[62] D. Epstein, J. Cannon, F. Holt, S. Levy, M. Patterson and W. Thurston, *WORD PROCESSING IN GROUPS*, Jones and Bartlett, Boston, MA 1992, Chapter 9.

[63] J. Etnyre, *Legendrian and Transversal knots*, arXiv math.SG/0306256.

[64] J. Etnyre and K. Honda, *Cabling and transverse simplicity*, arXiv:math.SG(GT)/0306330.

[65] E. Fadell and J. Van Buskirk, *The braid groups of* E^2 *and* S^2, Duke Math J. **29** (1962), 243-257.

[66] E. Fadell and L. Neuwirth, *Configuration spaces*, Math Scand. **10** (1962), 111-118.

[67] R. Fenn, M. Greene, D. Rolfsen, C. Rourke and B. Wiest, *Ordering the braid groups*, Pac. J. Math. **191**, No. 1 (1999), 49-74.

[68] E. Formanek, *Braid group representations of low degree*, Proc. London Math Soc.(3), **73** (1996), No. 2, 279-322.

[69] R. Fox and L. Neuwirth, *The braid groups*, Math. Scand. **10** (1962), 119-126.

[70] N. Franco and J. Gonzalez-Meneses, *The conjugacy problem for braid groups and Garside groups*, Journal of Algebra, **266**, No. 1 (2003), 112-132.

[71] R. Fenn, E. Keyman and C. Rourke, *The singular braid monoid embeds in a group*, Journal of Knot Theory and its Ramifications, **7** (1998), No. 7, 881-892.

[72] D. Fuchs and S. Tabachinikov, *Invariants of Legendrian and transverse knots in the standard contact space*, Topology **36** (1997) No. 5, 1025-1053.

[73] W. Fulton and J. Harris, *REPRESENTATION THEORY*, Springer-Verlag (1991).

[74] J. Franks and B. Williams, *Braids and the Jones-Conway polynomial*, Trans. AMS **303** (1987), No. 1, 97-108.

[75] P. Freyd, D. Yetter, J. Hoste, W. Lickorish, K. Millett, and A. Ocneanu, *A new polynomial invariant of knots and links*, Bull. Amer. Math. Soc. **12** (1985), 183-312.

[76] F. Garside, *The braid group and other groups*, Quart. J. Math Oxford **20** (1969), 235-254.

[77] V. Gebhardt, *A new approach to the conjugacy problem in Garside groups*, preprint arXiv:math.GT/0306199 v2, 21 October 2003.

[78] R. Ghrist, *Configuration spaces and braid groups on graphs in Robotics*, in "Knots, braids and mapping class groups", AMS/IP Studies in Advanced Mathematics, Amer. Math. Soc. and Int. Press 2001, 29-40.

[79] E. Giroux, *Géométrie de Contact:de la Dimension Trois vers les Dimensions Supérieures*, Proc. ICM (Beijing 2002), **II**, 405-414.

[80] D. Goldsmith, *The theory of motion groups*, Mich. Math. J. **28**, No. 1 (1981), 3-17.

[81] ____________, 1982: *Motions of links in the 3-sphere*, Math Scand. **50**, 167-205.

[82] J. Gonzalez-Meneses, *The* n^{th} *root of a braid is unique up to conjugacy*, Algebraic and Geometric Topology **3** (2003), 1103-1118.

[83] J. Gonzalez-Meneses and B. Wiest, *On the structure of the centralizer of a braid*, preprint, arXiv math.GT/0305156.

[84] E. Gorin and V. Lin, *Algebraic equations with continuous coefficients and some problems in the algebraic theory of braids*, Mat. Sbornik, Volume 7 (1969), No.4.

[85] P. de la Harpe, M. Kervaire and C. Weber, *On the Jones polynomial*, L'Enseignement Mathematique **32** (1986), 271-335.

[86] N. Ivanov, *Mapping Class Groups*, pages 523-633 of Handbook of Geometric Topology, Editors R. Daverman and R. Sher, Elsevier Science 2002.

[87] V. Jones, *Index for subfactors*, Invent. Math. **72** (1983), 1-25.

[88] __________, *Braid groups, Hecke algebras and subfactors*, Geometric Methods in Operator Algebras, Pitman Research Notes in Mathematics **123** (1986), 242-273.

[89] __________, *Hecke algebra representations of braid groups and link polynomials*, Annals of Math. **126** (1987), 335-388.

[90] D. Kazdan and G. Lusztig, *Representations of Coxeter groups and Hecke algebras*, Invent. Math. **53** (1979), 165-184.

[91] S. Kerckhoff, *The Nielsen realization problem*, Annals of Mathematics (2) **117**, No. 2 (1983), 235-265.

[92] K. H. Ko, S. J. Lee, J. H. Cheon, J. W. Han, J. Kang and C. Park, *New publlic key cryptosystems using braid groups*, in Lecture Notes in Computer Science **1880**, Springer, Berlin 2000.

[93] K. Y. Ko and S. J. Lee, *Flypes of closed 3-braids in the standard contact space*, J. Korean Math Soc.**36** (1999), 51-71.

[94] K. Y. Ko and J. W. Han, *Positive presentations of the braid groups and the embedding problem*, Math. Zeit. **240**, No. 1 (2002), 211-232.

[95] R. Kleinberg and W. Menasco, *Train tracks and zipping sequences for pseudo-Anosov braids*, Chaos, Solitons and Fractals, **9**, No. 4/5 (1998), 793-809.

[96] M. Korkmaz, *On the linearity of certain mapping class groups*, Turkish J. Math. **24**, No. 4 (2000), 367-371.

[97] D. Krammer, *The braid group B_4 is linear*, Invent. Math. **142** No. 3, (2000), 451-486.

[98] __________, *Braid groups are linear*, Ann. of Math. (2) 155 No. 1, (2002), 131-156.

[99] S. J. Lee, data presented at a talk at the First East Asian Conference on Knots and Related Topics, Seoul, Korea, February 16-19, 2004.

[100] R. Lawrence, *Homological representations of the Hecke algebra*, Comm. Math. Phys. **135** No. 1, (1990), 141-191.

[101] W. B. R. Lickorish, *AN INTRODUCTION TO KNOT THEORY*, Springer-Verlag Graduate Texts in Mathematics **175** (1997).

[102] D. Long, *Constructing representations of braid groups*, Comm. Math. Anal. Geom. **2**(1994), 217-238.

[103] D. Long and M. Paton, *The Burau representation of the braid group B_n is not faithful for $n \geq 6$*, Topology **32** (1993), 439-447.

[104] A. Markov, *Über die freie Äquivqlenz geschlossener Zöpfe* , Recueil Math., Moscou **1** (1935).

[105] W. Menasco, *On iterated torus knots and transversal knots*, Geom. Top. **5** (2001), 651-682.

[106] J. Michel, *A note on words in braid monoids*, J. Algebra **215**, No. 1 (1999), 366-377.

[107] J. Moody, *The Burau representation of the braid group B_n is not faithful for large n*, Bull. Amer. Math. Soc. **25** (1991), 379-384.

[108] H. Morton, *Infinitely many fibered knots having the same Alexander polynomial*, Topology **17** (1978), 101-104.

[109] ___________, *An irreducible 4-string braid with unknotted closure*, Math. Proc. Camb. Phil. Soc. **93** (1983), 259-261.

[110] ___________, *Threading knot diagrams*, Math. Proc. Camb. Phil. Soc. **99** (1986), 247-260.

[111] ___________, *Seifert circles and knot polynomials*, Math. Proc. Camb. Phil. Soc. **99** (1986), 107-109.

[112] L. Mosher, *Pseudo-Anosov's on punctured surfaces*, PhD thesis, Princeton University, 1983.

[113] L. Mosher, *Train track expansions of measured foliations*, partially completed preprint, http://newark.rutgers.edu∼mosher/

[114] J. Murakami, *The Kauffman polynomial of links and representation theory*, Osaka J. Math., **24**, No. 4 (1987), 745-758.

[115] J. Nielsen, *COLLECTED MATHEMATICAL PAPERS OF JAKOB NIELSEN*, edited by V. Hansen, Birkhauser (1986), [N-18], [N-20] and [N-21].

[116] S. Orevkov and V. Shevchisin, *Markov Theorem for Transversal Links*, preprint Feb 2003, math.GT/0112207.

[117] L. Paris, *Artin monoids embed in their groups*, Cooment. Math. Helv. **77**, No. 3 (2002), 609-637.

[118] __________, *The proof of Birman's conjecture on singular braid monoids*, arXiv:math.GR/0306422.

[119] M. Patterson and A. Razborov, *The set of minimal braids is Co-NP-complete*, J. of Algorithms **12** (1991), 393-408.

[120] R. Penner (with J. Harer), *Combinatorics of train tracks*, Annals of Math Studies No. **125** (1992) Princeton University Press, Princeton NJ.

[121] M. Picantin, *The conjugacy problem in small Gaussian groups*, Comm. Algebra **29**, No. 3 (2001), 1021-1039.

[122] __________, *Explicit presentations for the dual braid monoids*, arXiv:math.GR/0111280

[123] D. Rolfsen, *KNOTS AND LINKS*, American Mathematical Society (2003).

[124] I. Smith, *Geometric monodromy and the hyperbolic disc*, Q. J. Math. **52** (2001), no. 2, 217–228.

[125] C. Squier, *The Burau representation is unitary*, Proc. Amer. Math. Soc. **90** No. 2, (1984), 199-202.

[126] R. Stanley, *On the number of reduced decompositions of elements of Coxeter groups*, European J. Comb. **5** (1984), 359-372.

[127] N. Steenrod, *The topology of Fibre Bundles*, Princeton Univ. Press 1974

[128] W. Thurston, *On the topology and geometry of diffeomorphisms of surfaces*, Bull. Amer. Math. Soc. **19** (1988), 109-140.

[129] P. Traczyk, *A new proof of Markov's braid theorem*, Knot Theory, Banach Center Publications **42** (1998), Warsaw.

[130] V. Turaev, *Faithful linear representations of the braid groups*, arXiv math.GT/0006202.

[131] B. Wiest, *Dehornoy's ordering of the braid groups extends the subword ordering*, Pacific J.Math. **191** (1999), 183-188.

[132] H. Wenzl, *Representations of Hecke algebras and subfactors*, PhD thesis, Univ. of Pennsylvania, 1985.

[133] G. Wright, *A foliated disk whose boundary is Morton's irreducible 4-braid*, Math. Proc. Camb. Phil. Soc. **128** (2000), 95-101.

[134] N. Wrinkle, PhD thesis, Columbia University, 2002.

[135] P. Vogel, *Representation of links by braids: A new algorithm*, Comment. Math. Helvetici, **65**, No. 1 (1990), 104-113.

[136] S. Yamada, *The minimal number of Seifert circles equals the braid index of a link*, Invent. Math., **89**, No. 2 (1987), 347-356.

[137] M. Zinno, *On Krammer's Representation of the Braid Group*, Math. Ann. **321**, No. 1 (2001), 192-211.