The Abdus Salam International Centre for Theoretical Physics

United Nations Educational, Scientific and Cultural Organization

IAEA International Atomic Energy Agency

2034-6

Advanced School and Conference on Knot Theory and its Applications to Physics and Biology

11 - 29 May 2009

Invariants of Links from Diagram Colouring (Quandles and Biquandles)

Roger Fenn

Sussex University
Falmer Brighton BN1 9QH
England

Invariants of Links from Diagram Colouring

(Quandles and Biquandles)

ROGER FENN

Mathematics Department
Sussex University
Falmer Brighton BN1 9QH England

Trieste: May 2009

This is based on the work of several people. The main contributions come from joint papers by myself and A. Bartholomew, S. Budden, M. Jordan, L. Kauffman, R. Rimanyi, C. Rourke, B. Sanderson and V. Turaev.

ABSTRACT

The notion of developing a knot invariant by colouring or labelling the arcs of a diagram is an old one. The very first usable knot invariant was "3-colouring" which is a representation of the fundamental group. The fundamental group itself can be thought of, via the Wirtinger presentation, as a knot invariant obtained by labelling the arcs of a diagram by group generators. The universal application of this principal leads to the notion of a rack, which is the algebraic distillation of two of the Reidemeister moves. It is a set with a binary operation such that right multiplication is an automorphism. If in addition the first Reidemeister move is shadowed then the result is a quandle. Since a quandle is a rack but not necessarily conversely we shall refer to racks rather than quandles except when this notion is explicitly needed.

Any codimension two link has a fundamental rack which contains more information than the fundamental group. Racks provide an elegant and complete algebraic framework in which to study links and knots in 3–manifolds, and also for the 3–manifolds themselves. Racks have been studied by several previous authors and have been called a variety of names. In the first two chapters of these notes we consolidate the algebra of racks and show that the fundamental rack is a complete invariant for irreducible framed links in a 3–manifold. We give some examples of computable link invariants derived from the fundamental rack and explain the connection of the theory of racks with that of braids.

In later chapters we extend the notion of a rack to a birack (and biquandle). This leads to important invariants of virtual links, that is, links in a surface cross an interval modulo surgery.

1. Previous work

The earliest work on racks (known to us) is the definition of a kei by Takasaki in 1943, see [**Tak**]. Other early birds are Conway and Wraith [**CW**], and we are indebted to these authors for a copy of their (unpublished) correspondence. They used the name **wrack** for the concept and we have adopted this name, not merely because it is the oldest name, but also because it is a simple English word which (to our knowledge) has no other mathematical meaning. We have however chosen the more common spelling. Rack is used in the same sense as in the phrase "rack and ruin". The context of Conway and Wraith's work is the conjugacy operation in a group and they regarded a rack as the wreckage of a group left behind after the group operation is discarded and only the notion of conjugacy remains. They studied the basic algebra of racks in a special case (the quandle case) but also were aware of the general case and the main topological application (the fundamental rack of a knot in a 3–manifold).

A comprehensive published study of racks in a toplogical context is due to Joyce [**J**]. He invented the name **quandle**. Joyce establishes the basic algebra of quandles, giving several examples, and defines augmented quandles and the associated group (which he calls *Adconj*). He defines the fundamental quandle of a knot in S^3 giving both the topological definition in terms of "nooses" and the definition in terms of the presentation which can be read from a diagram of the knot. He proves the equivalence of the two definitions. His main result is that the fundamental quandle classifies the knot. Joyce's work was largely duplicated independently by Matveev [**Mat**] who used the name **distributive groupoid**. But note that a rack is not a groupoid in the usually accepted sense.

Kauffman [**K**] defines racks in full generality using the name **crystal** for the concept. Kauffman defines the fundamental rack of a knot in S^3 and applies Joyce's theorem to prove that it is a classifying invariant. He also extends Joyce's work on the Alexander quandle to racks and defines an associated R–matrix.

The most extensive algebraic survey of racks is given by Brieskorn [**Br**], who is particularly interested in the context of braids and singularities. In his introduction Brieskorn writes:

> *"Whilst preparing this survey, I found an extremely simple concept unifying many investigations on this subject as well as classical results of E.Artin, A.Hurwitz and W.Magnus. This is the notion of an automorphic set."*

The definition of an automorphic set coincides with that of a rack or crystal. Although "automorphic set" is the mathematically correct terminology for the concept, it has not generally been adopted. Brieskorn's paper contains a wealth of algebraic material about racks, most of which is not relevant to the topological context of our work, except for the connection with braid groups, which we shall examine in section 7.

Winker [**Win**] extends Joyce's work and defines an analogue of the Cayley graph for a quandle and Krüger [**Kr**], independently and simultaneously with [**FR**], defines free products of racks and investigates the automorphism group of the free rack, cf. section 7. There is an interesting connection of racks with computer theory. Roscoe [**Ros**] studies an algebraic object which satisfies just one of the two rack laws (the rack identity) in the context of computer information updates. Finally there is also a connection with some problems in logic, see Dehorney, Jech and Laver **[D, Je, La]**

Content of the notes

These notes review and consolidate much of the previous work on the subject and also contains many new results and new formulations of old results. An outline detailing these new results and the previous work now follows for this chapter.

Section 1 contains the basic definition and some of the examples of racks that we shall need. Most of these examples come from Joyce and Brieskorn. We use the exponential notation for the rack operation, and later a sub-exponential notation for the extra birack operation.

Section 2 contains a review of some of the basic algebra of racks. Most of this section is a reworking for racks of algebra in Joyce's paper. However the section also includes the concept of the free product of two racks and the connection with crossed modules (the associated crossed module). Further basic algebra can be found in Brieskorn paper and in Ryder's Ph.D. thesis [**Ryd**].

In section 3 we consider the fundamental rack of a codimension 2 embedding and its properties. This material contains a reworking in the rack context of material of Matveev and Joyce. It also contains the identification of the associated crossed module as the relative second homotopy group and a calculation of the operator group for the fundamental rack of a classical link.

Section 4 is about presentations. A "Tietze" theorem for rack presentations is given and, in a reworking of material of Joyce and Kauffman, we show how the fundamental rack of a classical link (in S^3) has a finite presentation which can be read in a natural way from the diagram. This section also contains material on presentations of augmented racks.

Section 5 contains the classification theorem which is a generalisation to arbitrary 3–manifolds of the results of Joyce and Kauffman.

In section 6 we make a start on the invariants that can be read from the fundamental rack; see also the theses of Devine [**Dev**], Azcan [**Asc**], Kelly [**Kel**], Lambropoulou [**Lam**] and Ryder [**Ryd**]. New material in this section includes the invariants derived from the (t, s)–rack and matrix racks (examples 6 and 7 of 6.1 and example 3 of 6.3).

Finally in section 7 we explain the connection of the theory of racks with that of braids and in an appendix we give the analogue of Nielsen theory for automorphisms of the free rack. These results allow us to give a criterion for a rack to be a **classical** rack (i.e. isomorphic to the fundamental rack of a framed link in S^3).

1.1. Definitions and Examples

We consider sets X with a binary operation which we shall write exponentially

$$(a, b) \mapsto a^b.$$

There are several reasons for writing the operation exponentially.

(1) The operation is unbalanced and should be thought of as an action, i.e. think of a^b as meaning *the result of* b *acting or operating on* a.

(2) In group contexts exponentiation signifies conjugation. A group with conjugation is one of the principal examples of a rack — indeed this was the source for one strand of the earlier work on racks **[CW]**. A rack is an algebraic object which has just *some* of the properties of a group with conjugacy as the operation.

(3) Finally, and most conveniently, exponential notation allows brackets to be dispensed with, because there are standard conventions for association with exponents. In particular

$$a^{bc} \text{ means } (a^b)^c \text{ and } a^{b^c} \text{ means } a^{(b^c)}.$$

1.1.1 Definition *Racks, Quandles and Keis*

Consider the following axioms, which are related to the Reidemeister moves.

Axiom 1 Given $a \in X$ then $a^a = a$.

Axiom 2 Given $a, b \in X$ there is a unique $c \in X$ such that $a = c^b$.

Axiom 3 Given $a, b, c \in X$ the formula

$$a^{bc} = a^{cb^c}$$

holds. We call this formula the **rack identity (first form).**

A **rack** is a non-empty set X with a binary operation satisfying axioms 2 and 3. A **quandle** is a rack satisfying axiom 1.

Several consequences flow from these axioms.

The second axiom implies that, for each $b \in X$ the function $f^b(x) := x^b$ is a bijection of X to itself, and this fits with the idea that the operation is a (right) action of X on itself.

We shall write $a^{\overline{b}} = (f^b)^{-1}(a)$ for the element c given by axiom 2, but notice that $a^{\overline{b}}$ is *a single symbol for an element of* X. It is not suggested that $\overline{b}$ is itself an element of X; however the notation is suggestive (and intended to be) because now $a^{b\overline{b}} = a^{\overline{b}b} = a$ for all $a, b \in X$. Thus if we identify $\overline{b}$ with b^{-1} then we can give a meaning to any expression of the form x^w where $w = w(a, b, \ldots)$ is a word in $F(X)$ the free group on X, namely the result of repeatedly acting on x by $f^a, (f^a)^{-1}, f^b, (f^b)^{-1}$ etc. The word w is again *not to be regarded as an element of* X, but as an **operator** on X. Shortly, we shall formalise this by introducing the **operator group**.

A **Kei** is a quandle which satisfies

Axiom K Given $a, b \in X$ then $a^{b^2} = a$. In other words $a^{\overline{b}} = a^b$.

The rack identity is a right self-distributive law as can be seen if we temporarily use the notation $a \cdot b$ for a^b:

$$(a \cdot b) \cdot c = (a \cdot c) \cdot (b \cdot c).$$

Axioms 2 and 3 are equivalent to the statement that *right multiplication is an automorphism.*

Substituting $d = a^c$ in the rack identity and then changing d back to a gives the alternative form:

Axiom 3′ Given $a, b, c \in X$ the formula

$$a^{b^c} = a^{\overline{c}bc}$$

holds. This is the **rack identity (second form).**

In other words b^c operates like $\overline{c}bc$, which makes clear the connection between the rack operation and conjugacy in a group.

At this point we can note that given a rack X there is a mirror rack $\overline{X}$ with binary operation $a^{\overline{b}}$.

The Operator Group

In expressions such as $a^{b\overline{c}}$ we refer to a as being at **primary** level and $b, \overline{c}$ as at **operator** level. The second form of the rack identity makes clear that we do not need any "higher" level operators. Expressions involving repeated

operations can always be resolved into one of the form a^w where $a \in X$ is at the primary level and w, lying in the free group $F(X)$ on X, is at the operator level.
In this way we have an action by the group $F(X)$ on X. In general if G acts on X, written $(a, g) \mapsto a \cdot g$ and if $\partial : X \to G$ is a map satisfying $\partial(a \cdot g) = g^{-1}\partial(a)g$ then X has the structure of a rack given by $a^b := a \cdot \partial(b)$. In many situations this is the most convenient method of describing the rack operation. The similarity with crossed modules should be clear.
We shall pursue the notion of a rack with a group G operating in section 2, when we introduce the formal notion of an augmented rack. We shall then be able to formalise the connection with crossed modules.
To make operators precise we define **operator equivalence** by:

$$w \equiv z \iff a^w = a^z \text{ for all } a \in X$$

where $w, z \in F(X)$.
The equivalence classes form the **Operator Group** $Op(X)$ which could also be defined as $F(X)/N$ where N is the normal subgroup

$$N = \{w \in F(X) \mid w \equiv 1\}.$$

1.1.2 Examples of operator equivalence
Since $b^{a^a} = b^{\overline{a}aa}$ (by the rack identity) $= b^a$ for all $a, b \in X$. We have

$$a^a \equiv a \quad \text{for all} \quad a \in X.$$

More generally if a^{b^n} means $a^{bb\ldots b}$ (n repeats) then $a \equiv a^{a^n}$.
In terms of operator equivalence, the rack identity can again be restated:
Axiom 3″ Given $a, b \in X$ we have

$$a^b \equiv \overline{b}ab.$$

This is the **rack identity (third form).**

Orbits and stabilizers
We can now see that a rack is a set X with an action of $F(X)$ (or its quotient $Op(X)$) on X satisfying the rack identity. In section 2 we shall see that there is another group naturally associated to a rack, lying between $F(X)$ and $Op(X)$, called the **associated group**, which therefore also acts on X. The associated group is particularly important because it has a universal property not shared by either $F(X)$ or $Op(X)$.
Since X is a set with a group action we can use all the language of group actions in the context of racks. In particular X splits into disjoint **orbits** and each element has a **stabilizer** (in $F(X)$ or $Op(X)$) associated with it.
1.1.3 Examples of Racks
Example 1 *The Conjugation Rack*
Let G be a group, then conjugation in G i.e. $g^h := h^{-1}gh$ defines a rack operation on G. This makes G into the **conjugation rack** written $\text{conj}(G)$ or alternatively G_{conj}.
The operator group in this rack is the group of inner automorphisms of G and the orbits are the conjugacy classes. Given $g, h \in G$ then $g \equiv h$ if and only if gh^{-1} is in the centre of G.
Example 2 *The Dihedral Rack*
Any union of conjugacy classes in a group forms a rack with conjugation as operation. In particular let R_n be the set of reflections in the dihedral group D_{2n} of order $2n$ (which we regard as the symmetry group of the regular n–gon). Then R_n forms a rack of order n, with operator group D_{2n}, called the **dihedral** rack of order n.
Example 3 *The Core Rack*
The rule $g^h := hg^{-1}h$ also defines a rack operation in a group G called the **core rack**, $\text{core}(G)$, cf. Joyce **[J]**.
Great care is needed working with this rack because composition in the operator group does not correspond to composition in G (g^{hj} has two meanings according as the product hj is taken in G or the operator group).
This is an example of a kei: where $a^2 \equiv 1$ for all $a \in X$, since:

$$g^{h^2} = h(hg^{-1}h)^{-1}h = g, \text{ for all } g \in G,$$

here h^2 means composition in the operator group and the other products are taken in G.
Example 4 *The Reflection Rack*
Let P, Q be points of the plane and define P^Q to be P reflected in Q (i.e. $2Q - P$ in vector notation).
It is elementary to show that this is a rack operation. This example can be generalised by replacing the plane by any geometry with point symmetries satisfying certain general conditions (see Joyce for details). Examples include the natural geometries of S^n and $\mathbb{R}P^n$. Interesting subracks of these latter racks are given by the action of Coxeter groups on root systems, cf. example 10 below.
Example 5 *The Alexander Rack*
Let Λ be the ring of Laurent polynomials $\mathbb{Z}[t, t^{-1}]$ in the variable t. Any $\Lambda-$module M has the structure of a rack with the rule $a^b := ta + (1-t)b$.
For example, letting M be the plane and the action of t multiplication by -1, yields the reflection rack of example 4.

The Quandle Condition
All the above examples have satisfied the identity $a^a = a$ for all $a \in X$, and are therefore quandles.
Example 5 can be generalised to yield a non-quandle rack:
Example 6 *The* $(t,s)-$*Rack*
Let Λ_s be the ring $\mathbb{Z}[t, t^{-1}, s]$ modulo the ideal generated by $s(t+s-1)$. Any Λ_s-module M has the structure of a rack by the rule

$$a^b := ta + sb.$$

This operation satisfies the **Abelian entropy condition**:

$$u^{vw^x} = u^{wv^x}.$$

For explicit representations in terms of matrices see section 6 (6.1 example 6). When s acts like $1-t$ this rack reverts to the Alexander rack, discussed above.
Example 7 *The Cyclic Rack*
Here is a finite rack which is also not a quandle:
The **cyclic rack of order** n, is given by $C_n = \{0, 1, 2, \ldots n-1\}$, the residues modulo n, with operation $i^j := i+1$ mod n for all $i, j \in C_n$.
This example can be generalised: Let X be any G–set and choose a fixed element $g \in G$, then $a^b := a \cdot g$ for all $a, b \in X$ defines a rack structure on X.
Example 8 *There are six different isomorphism classes of racks of order 3:*
1. The **trivial rack** $\{a, b, c \mid x^y = x \text{ for all } x, y\}$.
2. The cyclic rack C_3.
3. The dihedral rack R_3.
4. $\{a, b, c \mid f^a = f^b = f^c = (b, c)\}$ where (b, c) means the symmetry which interchanges b and c and leaves a fixed.
5. $\{a, b, c \mid f^b = f^c = (b, c), f^a = \text{id.}\}$
6. $\{a, b, c \mid f^a = (b, c), f^b = f^c = \text{id.}\}$

Classes 1,3,6 are quandles, whilst 2,4,5 are not.
The last example gives some idea of the rich and varied structure of racks as compared with groups, cf. Ryder [**Ryd**].
Example 9 *The Free Rack*
The **free rack** $\mathrm{FR}(S)$ on a given set S is defined, as a set, to be $S \times F(S)$. We write the pair (a, w) as a^w, i.e.

$$\mathrm{FR}(S) = \{a^w \mid a \in S, w \in F(S)\}.$$

The rack operation is defined by

$$(a^w)^{(b^z)} = a^{w\overline{z}bz}.$$

Axiom 2 of definition 1.1 is easy to check whilst for the rack identity notice

$$(a^w)^{b^z} = a^{w\overline{z}bz} \equiv \overline{w\overline{z}bz}\ a\ w\overline{z}bz = \overline{\overline{z}bz}\ \overline{w}aw\ \overline{z}bz \equiv \overline{b^z}a^w b^z$$

which is the third form of the identity (axiom $3''$).
The operator group is $F(S)$ whilst the set of orbits is in bijective correspondence with the elements of S and all stabilizers are trivial.
The free rack has the universal property that any function $S \to X$, where X is a given rack, extends uniquely to a rack homomorphism $\mathrm{FR}(S) \to X$.
There is also the free quandle $\mathrm{FQ}(S)$ in which the additional conditions

$$a^{aw} = a^{a^{-1}w} = a^w$$

hold.
Example 10. *Coxeter racks.*
Let $(\ ,\ )$ be a symmetric bilinear form on $\mathbb{R}^n$. Then, if S is the subset of $\mathbb{R}^n$ consisting of vectors $\mathbf{v}$ satisfying $\mathbf{v}.\mathbf{v} \neq 0$, there is a rack structure defined on S by the formula

$$\mathbf{u}^{\mathbf{v}} := \mathbf{u} - \frac{2(\mathbf{u},\mathbf{v})}{(\mathbf{v},\mathbf{v})}\mathbf{v}.$$

Geometrically, this is the result of reflecting $\mathbf{u}$ in the hyperplane $\{\mathbf{w}|(\mathbf{w},\mathbf{v}) = 0\}$.
If we multiply the right-hand side of the above formula by -1, then the result geometrically is reflection in the line containing $\mathbf{v}$. In this case the formula

$$\mathbf{u}^{\mathbf{v}} := \frac{2(\mathbf{u},\mathbf{v})}{(\mathbf{v},\mathbf{v})}\mathbf{v} - \mathbf{u}$$

defines a quandle structure on S.
A **root system** is precisely a finite subrack of S which is closed under multiplication by -1 (i.e. closed under *both* rack operations), and then the operator group is the corresponding Coxeter group. For details see Humphreys [**Hum**], and for more information on the rack structures see Azcan [**Azc**], Brieskorn [**Br**].
We can generalise this example in the following way.
Example 11 *Racks defined by Hermitian Forms.*
Let R be a commutative ring with identity and an involutive automorphism $r \mapsto \overline{r}$ called **conjugation.** Let A be an $R-$module with a **Hermitian form**

$$(\ ,\ ) : A \times A \to R.$$

In other words $(\ ,\ )$ is linear in the first variable and $(b,a) = \overline{(a,b)}$.
Let A^* denote those elements of A for which (a,a) is a unit of the ring R.
Let λ, μ be elements of R, such that μ is a unit and λ satisfies $\lambda\overline{\lambda} = 1$. Define the rack operation on A^* by the formula

$$a^b := \mu(a + (\lambda - 1)\frac{(a,b)}{(b,b)}b).$$

Specialisation yields the following examples:
(a) Let $R = \mathbb{R}$ be the reals, let $A = V$ be a vector space over $\mathbb{R}$ and let conjugation be the identity. Let $(\ ,\ )$ be a symmetric bilinear real form defined on V and let $\lambda = -1$ and $\mu = 1$. Then this is the rack structure considered in the last example.
(b) The obvious specialisation of the above to the complex field yields a rack in which the action is complex reflection, see Coxeter [**Cox**].
(c) Take $R = \mathbb{Z}[t, t^{-1}]$ to be the ring of Laurent polynomials with integer coefficients and conjugation defined by $t \mapsto t^{-1}$. Then λ is of the form $\lambda = t^n$ for some integer n.

We are indebted to Tony Carbery for pointing out the following infinite generalisation of the above example.
(d) Let R denote the ring of complex valued continuous functions defined on the unit circle of the complex plane. The conjugation operation in R is given by

$$\overline{f}(z) := \overline{f(z)}$$

where $z \mapsto \overline{z}$ is just the usual conjugation of complex numbers. The set of functions λ satisfying $\lambda\overline{\lambda} = 1$ can be identified with the multiplicative subset of functions from the unit circle to itself.

The last two examples are somewhat mysterious and their applications to knot theory are unknown to us.

1.2. Some Basic Algebra of Racks

In this section we will present some of the algebraic properties of racks needed in the rest of the paper. Further basic algebra will be given in section 4, when we consider presentations of racks and the analogue of the Tietze theorem. See also Brieskorn and Ryder.

Homomorphisms and congruences

There are obvious notions of rack **homomorphism**, **isomorphism** and **subrack**.

An equivalence relation $\sim$ on X is called a **congruence** if it respects the rack operation, i.e.

$$a \sim b,\ c \sim d \implies a^c \sim b^d$$
$$a \sim b,\ c \sim d \implies a^{\overline{c}} \sim b^{\overline{d}}.$$

The equivalence classes form a rack $X/\sim$ with operations defined by $[a]^{[b]} := [a^b]$ and $[a]^{\overline{[b]}} := [a^{\overline{b}}]$, where $[a]$ denotes the equivalence class of a.

A homomorphism $f : X \to Y$ of racks defines a congruence by $a \sim b \iff f(a) = f(b)$. Then the quotient $X/\sim$ is isomorphic to $f(X)$. This is an analogue of the first isomorphism theorem for groups.

The associated group

We have already met the operator group in the previous section. This is an invariant of racks but is not functorial. If we interpret the operation of a rack as conjugation (i.e. read a^w as $w^{-1}aw$) then we obtain a group $As(X)$ called the **associated group**. More precisely let $As(X) = F(X)/K$ where K is the normal subgroup of $F(X)$ generated by the words $a^b b^{-1} a^{-1} b$ where $a, b \in X$. So $As(X)$ is the biggest quotient of $F(X)$ with the property that, when considered as a rack via conjugation, the natural map from $F(X)$ to $As(X)$ is a rack homomorphism.

Given a rack homomorphism $f : X \to Y$, then there is an induced group homomorphism $f_\sharp : As(X) \to As(Y)$; thus we have an **associated group functor** As from the category of racks to the category of groups.

1.2.1 Proposition *Universal Property of the Associated Group*

Let X be a rack and let G be a group. Given any rack homomorphism $f : X \to G_{\mathrm{conj}}$ there exists a unique group homomorphism $f_\sharp : As(X) \to G$ which makes the following diagram commute:

$$\begin{tikzcd}
X \arrow[r, "\eta"] \arrow[d, "f"] & As(X) \arrow[d, "f_\sharp"] \\
G_{\mathrm{conj}} \arrow[r, "\mathrm{id}"] & G
\end{tikzcd}$$

where η is the natural map.

Moreover any group with the same universal property is isomorphic to $As(X)$.

Proof Let $\phi : F(X) \to G$ be the homomorphism defined on the free group on X by f. Then by hypothesis $\phi(a^b b^{-1} a^{-1} b) = 1$ for all $a, b \in X$. It follows that ϕ factors through a unique homomorphism $f_\sharp : As(X) \to G$ of groups and the commutativity of the diagram is clear. Uniqueness of $As(X)$ follows by the usual universal property argument. □

The following corollary is an easy consequence:

1.2.2 Corollary *The functor As is a left adjoint to the conjugation functor. This means there is a natural identification*

$$Hom(As(X), G) \cong Hom(X, \mathrm{conj}(G))$$

of group homomorphisms with rack homomorphisms. □

Example The cyclic rack C_n has operator group $\mathbb{Z}/n$ and associated group $\mathbb{Z}$.

This example makes it clear that the operator group is in general a non-trivial quotient of the associated group. We will now use this fact to make the following definition.

Definition The Excess of a Rack.

Let X be a rack and let N be the subgroup of $F(X)$ which acts trivially on X. Let K be the normal subgroup of $F(X)$ generated by the elements $a^b b^{-1} a^{-1} b$ for all $a, b \in X$. Define the **excess** of the rack X to be

$$Ex(X) = N/K = \ker\{As(X) \to Op(X)\}.$$

In the example above the excess is a copy of the integers.

The associated quandle There is a natural inclusion of the category of quandles in the category of racks and there is a functor from racks to quandles $X \mapsto X_q$. Here X_q is called the **associated quandle** defined as follows: Let $\sim$ be the smallest congruence on X satisfying $a^a \sim a$ for all $a \in X$. Then $X_q := X/\sim$.

This functor is a retraction because it is clearly the identity for a rack which is already a quandle. For explicit examples, consider the Coxeter rack X (section 1 example 10):

$$\mathbf{u}^{\mathbf{v}} := \mathbf{u} - \frac{2(\mathbf{u},\mathbf{v})}{(\mathbf{v},\mathbf{v})}\mathbf{v},$$

defined on the unit sphere $S = (\mathbf{u},\mathbf{u}) = 1$. Then the associated quandle is the **projectivisation** of X, defined by quotienting S by ± 1.

For the (t,s)–rack (section 1 example 6) the associated quandle is the Alexander rack (example 5).

Racks with an explicit group (augmented racks)

The concept of rack can be generalised to make the group action explicit.

An **augmented rack** comprises a set X with an action by a group G, which we write

$$(x,g) \mapsto x \cdot g \quad \text{where} \quad x,\ x \cdot g \in X \quad \text{and} \quad g \in G,$$

and a function $\partial : X \to G$ satisfying the **augmentation identity**:

$$\partial(a \cdot g) = g^{-1}(\partial a)g \quad \text{for all} \quad a \in X, g \in G,$$

which is precisely the same as saying that ∂ is a G–map when the action of G on itself is taken to be conjugation. We can now define an operation of X on itself by defining a^b to be $a \cdot \partial b$. Then the augmentation identity implies

$$\partial(a^b) = (\partial b)^{-1} \partial a \partial b$$

i.e.

$$a^b \equiv \overline{b} a b$$

which is the third form of the usual rack identity (axiom $3''$).

So an augmented rack is an ordinary rack with the extra structure of an explicit operator group.

Note that Joyce used ϵ for the augmentation map ∂. We have chosen to use ∂ instead of ϵ because of the analogy with crossed modules; see the definition of associated crossed module below.

The fundamental rack (to be defined in the next section) has a natural structure as an augmented rack with the fundamental group of the link acting as a group of operators. Here are some further examples.

1.2.3 Examples

(1) *The Lie rack*

We are indebted to Hyman Bass for pointing out the following important class of augmented racks. Let G be a Lie group and $\mathcal{G}$ the associated Lie algebra. Let $\partial : \mathcal{G} \to G$ be the exponential map and let G act on $\mathcal{G}$ via the adjoint action. Then the augmentation identity follows readily from definitions. Therefore the Lie algebra $\mathcal{G}$ is an augmented rack, with group the corresponding Lie group.

(2) *Gauge transformations*
Let E be a principal G–bundle and f a **gauge transformation** of E, that is an automorphism of E as a G–bundle. Then associated to f is an augmented rack structure on E with group G.
Let $p \in E$ then we can write $f(p) = p \cdot \partial(p)$ where $\partial(p) \in G$. This defines the function $\partial : E \to G$. To check the augmentation identity note that

$$\begin{aligned} p \cdot (g\partial(p \cdot g)) &= (p \cdot g) \cdot \partial(p \cdot g) \\ &= f(p \cdot g) \\ &= f(p) \cdot g \qquad \text{since } f \text{ is equivariant} \\ &= (p \cdot \partial(p)) \cdot g \\ &= p \cdot (\partial(p) g) \end{aligned}$$

which implies $g\partial(p \cdot g) = \partial(p)g$ since G acts freely, i.e.

$$\partial(p \cdot g) = g^{-1}\partial(p)g.$$

An augmented rack is a plain rack if we ignore or forget about the explicit group action. Conversely, there is a natural way to regard a plain rack as an augmented rack by taking as group $G = As(X)$ (the associated group) with ∂ the natural map. Thus we can regard the category of racks as a subcategory of the category of augmented racks and then the forgetful functor is a retraction of the larger category onto the smaller.

Crossed modules

A **crossed module** is an augmented rack in which
(1) X is a group
(2) ∂ is a homomorphism
(3) and we have the **crossed module identity**:

$$a \cdot \partial b = b^{-1}ab \text{ for all } a, b \in X,$$

where the left-hand side is the G action and the right-hand side is multiplication in X.
Note that condition (3) implies that the rack operation in X is conjugation. Thus crossed modules correspond precisely to conjugation racks.
Crossed modules occur naturally in topology: the second homotopy group $\pi_2(X, A)$ of a pair of topological spaces is a crossed module with group $G = \pi_1(A)$ (see Whitehead [**Wh**]). We shall use the notation $\widehat{\pi}_2(X, A)$ for this crossed module.

The associated crossed module

The associated group for a plain rack becomes the associated crossed module for an augmented rack. More precisely notice that if X is an augmented rack then its group G acts on $F(X)$ in the obvious way. Moreover

$$\begin{aligned} (a^b) \cdot g = a \cdot (\partial bg) &= a \cdot g(g^{-1}\partial bg) \\ &= a \cdot g\ \partial(b \cdot g) \quad \text{by the augmentation identity} \\ &= (a \cdot g)^{b \cdot g} \end{aligned}$$

Therefore in $F(X)$

$$(a^b b^{-1} a^{-1} b) \cdot g = (a \cdot g)^{(b \cdot g)} (b \cdot g)^{-1} (a \cdot g)^{-1} (b \cdot g)$$

and therefore the action of G on $F(X)$ induces an action on $As(X)$.
Thus $As(X)$ is a G–set and we also have the induced homomorphism $\partial_\sharp : As(X) \to G$. It can be readily checked that this gives $As(X)$ the structure of a crossed module, the **associated crossed module** to the augmented rack X.

Products of Racks

There are many kinds of products which can be defined in the category of racks. We shall only need to consider in detail the following:

The free product Let X, Y be two racks. Define their **free product** $X * Y$ to be the free rack on the disjoint union $X \amalg Y$ quotiented out by the original actions of X and Y.

More precisely $X * Y$ consists of elements of the form x^w or y^w where $x \in X$, $y \in Y$ and $w \in As(X) * As(Y)$ under the equivalence generated by the following:

$$x^{wt} \sim u^t \quad \text{where} \quad x \in X, \; w \in As(X), \; t \in As(X) * As(Y) \quad \text{and} \quad x^w = u \quad \text{in} \quad X$$

and a similar equivalence for Y.

The rack operation on $X * Y$ is defined by the same formula as for the free rack (section 1 example 9). That the operation is well defined follows from the definition of the associated group. For example suppose that $x^w = t$ in X then

$$(z^u)^{x^w} := z^{u\overline{w}xw} = z^{ut} \quad \text{since} \quad \overline{w}xw = t \quad \text{in} \quad As(X).$$

Notice that there are natural inclusions of X and Y in $X * Y$ and that the associated group is the free product:

$$As(X * Y) = As(X) * As(Y).$$

The following lemma implies that the free product is the categorical 'sum' in the category of racks:

1.2.4 Lemma *Let $f : X \to Z$, $g : Y \to Z$ be rack homomorphisms. Then there is a unique extension $f * g : X * Y \to Z$.*

Proof The free product $X * Y$ is generated as a rack by the images of X and Y under the natural inclusions, and the lemma follows. $\square$

Free product of augmented racks

The free product $X * Y$ of augmented racks X, Y with groups G, H is defined in a similar way:

We consider pairs (x, g) where $x \in X$ or Y and $g \in G * H$ with equivalence generated by

$$(x, gt) \sim (u, t) \quad \text{where} \quad x \in X, g \in G, t \in G * H \quad \text{and} \quad x \cdot g = u \quad \text{in} \quad X$$

and a similar equivalence for Y.

The group for $X * Y$ is $G * H$, $\partial_{X*Y} := \partial_X * \partial_Y$ and the action of $G * H$ on $X * Y$ is defined by right multiplication in the second coordinate.

There is again a universal property which we leave the reader to formulate.

Other products

The categorical 'product' for racks is the cartesian product with operation

$$(a, x)^{(b,y)} := (a^b, x^y).$$

There are several other products, for example the disjoint union $X \amalg Y$ where the rack operation is defined by letting Y act trivially on X and vice-versa (Breiskorn [**Br**]). This last product can be generalised by allowing Y to act via any function $Y \to \text{centre}(Op(X))$ and vice versa. Further products are defined by Ryder [**Ryd**].

1.3. The Fundamental Rack of a Link.

This is the most important rack of all and is the *raison d'être* of the whole theory. A (codimension two) **link** is defined to be a codimension two embedding $L : M \subset Q$ of one manifold in another. We shall assume that the embedding is proper at the boundary if necessary, that M is non-empty, that Q is connected and that M is **transversely oriented** in Q. In other words we assume that each normal disc to M in Q has an orientation which is locally and globally coherent.

The link is said to be **framed** if there is given a cross section (called a **framing**) $\lambda : M \to \partial N(M)$ of the normal disk bundle. Denote by M^+ the image of M under λ. We call M^+ the **parallel** manifold to M.

We consider homotopy classes Γ of paths in $Q_0 = \text{closure}(Q - N(M))$ from a point in M^+ to a base point. During the homotopy the final point of the path at the base point is kept fixed and the initial point is allowed to wander at will on M^+.†

The set Γ has an action of the fundamental group of Q_0 defined as follows: let γ be a loop in Q_0 representing an element g of the fundamental group. If $a \in \Gamma$ is represented by the path α define $a \cdot g$ to be the class of the composite path $\alpha \circ \gamma$.

We can use this action to define a rack structure on Γ. Let $p \in M^+$ be a point on the framing image. Then p lies on a unique meridian circle of the normal circle bundle. Let m_p be the loop based at p which follows round the meridian in a positive direction. Let $a, b \in \Gamma$ be represented by the paths α, β respectively. Let $\partial(b)$ be the element of the fundamental group determined by the loop $\overline{\beta} \circ m_\beta \circ \beta$. (Here $\overline{\beta}$ represents the reverse path to β and m_β is an abbreviation for $m_{\beta(0)}$ the meridian at the initial point of β.) The **fundamental rack** of the framed link L is defined to be the set $\Gamma = \Gamma(L)$ of homotopy classes of paths as above with operation

$$a^b := a \cdot \partial(b) = [\alpha \circ \overline{\beta} \circ m_\beta \circ \beta].$$

If L is an unframed link then we can define its **fundamental quandle.** The definition is very similar. Let $\Gamma_q = \Gamma_q(L)$ be the set of homotopy classes of paths from the boundary of the regular neighbourhood to the base point where the initial point is allowed to wander during the course of the homotopy over the *whole* boundary. The rack structure on Γ_q is similar to that defined on Γ.

There is a convenient halfway-house between framed and unframed links: a link L is **semi-framed** if *some* of the components of M are framed. A semi-framed link has a fundamental rack defined by allowing the initial point to wander on the whole boundary of the neighbourhoods of unframed components and on M^+ otherwise. This gives a common generalisation for the rack of a framed link and the quandle of an unframed link, and allows us to make economical statements of results which apply to all cases.

1.3.1 Proposition *The fundamental rack of a semi-framed link satisfies the axioms of a rack.*
The fundamental quandle of an unframed link satisfies the axioms of a rack together with the quandle condition.
In the semi-framed case the fundamental quandle of the corresponding unframed link (i.e. ignore framings), is the associated quandle of the fundamental rack.

Proof The axioms are easy to verify. The inverse action is determined by the class of

$$\overline{\alpha} \circ \overline{m}_\alpha \circ \alpha.$$

To check the rack identity we again use the action of the fundamental group. Using the notation above, $\partial(a^b)$ is represented by the loop

$$\overline{\beta} \circ \overline{m}_\beta \circ \beta \circ \overline{\alpha} \circ m_\alpha \circ \alpha \circ \overline{\beta} \circ m_\beta \circ \beta$$

which is the class of

$$\partial(b)^{-1}\partial(a)\partial(b).$$

In the unframed case note that the element a^a is represented by the path

$$\alpha \circ \overline{\alpha} \circ m_\alpha \circ \alpha \quad \simeq \quad m_\alpha \circ \alpha.$$

However in a homotopy in the definition of Γ_q the initial point is allowed to move along the loop m_α and so the path is homotopic to α which represents a.

The last part of the proposition is obvious. □

Note that if G denotes the fundamental group $\pi_1(Q_0)$ then the set Γ is in fact an augmented rack with group G. We shall use the notation $\widehat{\Gamma}$ for this augmented rack in order to distinguish it from the **plain** fundamental rack Γ. Note that Γ is the underlying plain rack to $\widehat{\Gamma}$.

† This reverses the more usual dog wagging tail convention where the initial point of a path stays fixed. However the tail wagging dog convention fits in more comfortably with operations on the right.

We will now identify the associated group of the fundamental rack of an arbitary codimension two link and the operator group of the fundamental rack of a link of circles in an oriented 3–manifold.

Consider the following fragment of the exact homotopy sequence of the pair (Q, Q_0):

$$\pi_2(Q) \to \pi_2(Q, Q_0) \to \pi_1(Q_0) \to \pi_1(Q).$$

We shall call $\pi_1(Q_0)$ the **fundamental group** of the link and $\ker\{\pi_1(Q_0) \to \pi_1(Q)\} = \mathrm{im}\{\pi_2(Q, Q_0) \to \pi_1(Q_0)\}$ the **kernel** of the link. Further we shall call the relative group $\pi_2(Q, Q_0)$ the **associated group of the link**. Note that if $\pi_2(Q) = 0$ then the associated group and the kernel of the link coincide, and if in addition $\pi_1(Q) = 0$ as in the classical case of links in S^3 then all three groups coincide.

1.3.2 Proposition *The associated group of the fundamental rack $\Gamma(L)$ of a semi-framed link L can be naturally identified with the associated group of L.*

Moreover the associated crossed module of the fundamental augmented rack $\widehat{\Gamma}(L)$ can be identified with the crossed module $\widehat{\pi}_2(Q, Q_0)$ corresponding to the second relative homotopy group $\pi_2(Q, Q_0)$.

Proof Let $a \in \Gamma$, then ∂a is represented by the path $\overline{\alpha} \circ m_\alpha \circ \alpha$ which bounds an obvious 2–disc, namely the meridinal disc at the initial point of α. Thus there is a map $\Gamma \to \pi_2(Q, Q_0)$. Under this map the rack operation corresponds to conjugacy. Therefore it induces a homomorphism $As(\Gamma) \to \pi_2(Q, Q_0)$. We shall show that this is an isomorphism by constructing an inverse map.

Suppose $g \in \pi_2(Q, Q_0)$ is represented by the disc D. After a homotopy we may assume that D meets the neighbourhood N of the link transversely in a finite number of little discs $D_1, \ldots, D_n$. Assign to D_i the sign ϵ_i where $\epsilon_i = +1$ if the orientation of D agrees with the orientation of D_i and -1 if not.

Pick a base point in each ∂D_i. In the case of a framed link let the base point be the intersection of ∂D_i with M^+. Join each of these n base points to the base point $*$ of Q by n paths $\alpha_1, \alpha_2, \ldots, \alpha_n$ in $D - \cup_i\{D_i\}$ which only meet at $*$ and arrive at $*$ in the order $1, 2, \ldots, n$.

Note that this implies that the paths are uniquely determined up to isotopy and possible initial twists about the little disks. Each path α_i determines an element a_i of Γ and this defines a word $a_1^{\epsilon_1} a_2^{\epsilon_2} \cdots a_n^{\epsilon_n}$ in $F(\Gamma)$ and so an element of its quotient $As(\Gamma)$.

In order to check that this element is well defined it is only necessary to see what happens if we change the number of initial twists or the order of the subdiscs or if we change the choice of the disc D by a homotopy.

Now an initial twist changes a_i to $a_i^{a_i}$ (see the end of the proof of proposition 1.3.1) and an interchange of two elements a_i and a_j replaces $a_i a_j$ by $a_j a_i^{a_j}$, or a similar replacement with different signs. None of these affects the value of the product in $As(\Gamma)$.

Now suppose D' is a homotopic choice of disc. By making the homotopy transverse to the link we see that the word in $F(\Gamma)$ changes in two ways: either an interchange of order as above or an introduction or deletion of cancelling pairs aa^{-1} or $a^{-1}a$. Both leave the value in the associated group unchanged. The resulting map is the required inverse.

The last part of the proposition is readily checked. □

Remark The proposition shows that the fundamental augmented rack of a link is a *sharpened form* of the crossed module, $\widehat{\pi}_2(Q, Q_0)$.

To see that the rack really does contain more information than the crossed module consider the example of a knot k in S^2 which is the sum of two knots. The elements of $\pi_2(Q, Q_0)$ represented by meridinal discs across the two connecting arcs coincide, however in general the corresponding elements of the *rack* are different. Indeed it can be shown that the two elements of the rack are never the same if the two knots are both non-trivial.

1.3.3 Corollary *The associated group of the fundamental rack of a link in a 2–connected space can be identified with the fundamental group of the link. In particular the associated group and the fundamental group coincide for classical links in the 3–sphere and for links in a homotopy 3–sphere.* □

Remark The corollary implies that for links in a homotopy 3–sphere the plain fundamental rack Γ and the fundamental augmented rack $\widehat{\Gamma}$ essentially coincide (i.e. coincide under the embedding of racks in augmented racks described in the last section).

Orbits and stabilizers

The orbits in the fundamental augmented rack of a link are in bijective correspondence with the components of the link and the next lemma identifies the corresponding stabilizers.

An element of the fundamental group represented by a loop of the form $\overline{\alpha} \circ \gamma \circ \alpha$ where γ lies in $\partial N(M)$ and α represents the element $a \in \Gamma$ is called a–**peripheral.** The set of a–peripheral elements forms the a–**peripheral subgroup.** If γ lies in the subset M^+ then the class of $\overline{\alpha} \circ \gamma \circ \alpha$ is called a–**longitudinal** and the set of a–longitudinal elements forms the a–**longitudinal** subgroup. If γ lies in the boundary of a normal disc to M then $\overline{\alpha} \circ \gamma \circ \alpha$ is called a–**meridinal.** The set of a–meridinal elements forms the a–**meridinal subgroup.** (The kernel of the link is generated by meridinal elements.)

1.3.4 Lemma *If the element a of the fundamental rack is represented by a path α from ∂N to the base point and if h is an element of the fundamental group which fixes a then h is a–peripheral. If α starts on the neighbourhood of a framed component of M, then h is a–longitudinal.*

Proof Let h be represented by a loop γ. During the course of the homotopy of $\alpha \circ \gamma$ to α the initial point describes a loop δ in ∂N. This implies that γ is homotopic to the loop $\overline{\alpha} \circ \delta \circ \alpha$ and the result follows. □

1.3.5 Corollary *With the notation above, the stabilizer of a in the fundamental group is the a–longitudinal subgroup or the a–peripheral subgroup, according as the component where α starts is framed or unframed.* □

Seifert links and the operator group

We shall finish this section by identifying the operator group for knots and links in 3–manifolds.

Let G be the group of $\widehat{\Gamma}$ (the fundamental group of the link) and let K (the kernel of the link) be the image of $As(\Gamma)$ in G. Define the **action kernel** $J \subset G$ to comprise all elements of G which act trivially on Γ. Recall that the operator group $Op(\Gamma)$ is the quotient of $As(\Gamma)$ by the subgroup of elements which act trivially on Γ. Since the action of $As(\Gamma)$ factors via the action of G, $Op(\Gamma)$ is the quotient of K by the subgroup of K of elements which act trivially on Γ. In other words

$$Op(\Gamma) = \frac{K}{J \cap K}. \tag{3.6}$$

We shall now compute J.

We will need the following definition:

Definition *Seifert Links*

Consider a link in a 3–manifold whose complement has a Seifert circle fibration. An example is a torus link in S^3. We say that a component C of such a link is **framable** if the fibration extends to C and its neighbourhood so that C is a regular fibre. A component is said to be **naturally** framed if it is framable and has the framing given by neighbouring fibres. A **Seifert link** is a link in a 3–manifold whose complement has a Seifert circle fibration such that all framed components are naturally framed. Note that a Seifert link might be unframed and that some framable components might be unframed.

1.3.7 Proposition *Consider a semi-framed link in a 3–manifold whose complement is P^2 irreducible. The action kernel J of the link is non-trivial if and only if the link is a Seifert link. Moreover J can be described explicitly. There are four cases:*

(1) *The link is a Seifert link with at least one framed component, in which case J is the infinite cyclic subgroup of the fundamental group defined by the regular fibres.*

(2) *The link is unframed and Q_0 is $T \times I$ where T is a torus, in which case $J \cong \mathbb{Z}^2$ is the fundamental group.*

(3) *The link is unframed and Q_0 is $K \widetilde{\times} I$ (the twisted I bundle over a Klein bottle), in which case $J \cong \mathbb{Z}^2$ is a subgroup of index 2 in the fundamental group.*

(4) *The link is an unframed Seifert link and Q_0 is neither $T \times I$ nor $K \widetilde{\times} I$, in which case J is the same as in case (1).*

Proof: It is convenient to make the following observation about groups and normal subgroups. Let G be a group which has a subgroup H containing a non trivial element h such that any conjugate $g^{-1}hg$ is in H, where g lies in G. Then the group generated by $g^{-1}hg$ for all $g \in G$ is a non trivial normal subgroup of H and G.

Assume that $J \neq \{1\}$ and choose $h \in J$, $h \neq 1$. By lemma 1.3.4 h is a–peripheral and, for framed components, a–longitudinal *for all* $a \in \Gamma$.
For convenience take the base point $*$ in $\partial N(C)$ where C is a component of the link, and assume that C is framed. Consider elements $a \in \Gamma$ defined by loops α based at $*$. Such an element can be regarded (non-uniquely) as an element $g \in G$, the fundamental group. Now h is a–longitudinal for all such a hence $h = g^{-1}lg$ in G where l is some power of the longitude at $*$. It follows that our observation applies where H is the infinite cyclic subgroup of powers of the longitude. The complementary manifold M therefore has fundamental group containing a normal infinite cyclic subgroup. A result of Waldhausen **[Wal]** shows that M is a Seifert manifold and that l the longitude is a fibre.
A similar argument works if C is unframed. In this case the group H is $\mathbb{Z}^2$. If the resulting normal subgroup is infinite cyclic we can apply the previous argument and deduce that the complementary space is Seifert fibred. If the normal subgroup is $\mathbb{Z}^2$ then we know that it is peripheral and the complementary space is therefore either $T \times I$ where T is a torus or the twisted I bundle over a Klein bottle. The relevant details may be found in Hempel's book **[Hem]** as was kindly pointed out to us by A. Swarup. In either case both are Seifert fibred spaces.
Conversely if the link is Seifert then it is easy to see that the infinite cyclic subgroup defined by the regular fibres acts trivially. If a larger subgroup acts trivially then, by the argument in the previous paragraph, we are in the unframed case and Q_0 is $T \times I$ or $K \tilde{\times} I$. In the first case all elements act trivially while, in the second case, a subgroup of index 2 acts trivially. □

1.3.8 Corollary *The operator group of the (plain) fundamental rack of a semi-framed link in a connected orientable 3–manifold is one of: (a) the kernel of the link, (b) the kernel modulo the integers, (c)* $\mathbb{Z}/2$*, or (d) the trivial group.*
In the last three cases, the link is a Seifert link with possibly some homotopy discs or spheres added by connected sum.

Proof Decompose the complement into irreducible pieces. If two or more pieces are non-simply connected then the fundamental group has no normal subgroup isomorphic to $\mathbb{Z}$ or $\mathbb{Z}^2$, and hence the action kernel is trivial by the proof of the proposition.
Thus if the action kernel is non-trivial then the complement is irreducible, with possibly some homotopy discs or spheres added by connected sum. Remove these connected summands, then by the proposition the complement is a Siefert link and the action kernel is $\mathbb{Z}$ or $\mathbb{Z}^2$. Hence by equation 1.3.6 the operator group is the link kernel modulo $\mathbb{Z}$ or $\mathbb{Z}^2$. In the $\mathbb{Z}^2$ case we must be in case (2) or (3) of the proposition and the quotient is either $\mathbb{Z}/2$ or trivial. □

1.4. Presentations

The main purpose of this section is to explain the natural presentation that can be given to the fundamental rack of a link in a 3–manifold. This will involve explaining several layers of rack presentations. We start with the simplest. Throughout the section, we shall concentrate on framed links (rather than semi-framed links). There are analogues for semi-framed links of most of the results in the section, which are proved in analogous ways. By and large we leave the reader to formulate these parallel results, contenting ourselves with brief comments.

Primary Rack Presentations

A **primary presentation** for a rack consists of two sets S (the generating set) and R (the set of relators). A typical element of R is an ordered pair (x, y), where $x, y \in \mathrm{FR}(S)$, which we shall usually write as an equation: $x = y$ or $(x = y)$.
The presentation defines a rack $[\,S : R\,]$ as follows: Define the congruence $\sim$ on $\mathrm{FR}(S)$ to be the smallest congruence containing R (i.e. such that $x \sim y$ whenever $(x = y) \in R$). Then

$$[\,S : R\,] = \frac{\mathrm{FR}(S)}{\sim}.$$

We can describe $\sim$ more constructively as follows.
Consider the following process for generating relators. Start with the given set R of relators and enlarge R by repeating any or all of the following moves:

(a) Add a **trivial** relator $x = x$ for some $x \in \mathrm{FR}(S)$.
(b) If $(x = y) \in R$ then add $y = x$.
(c) If $(x = y), (y = z) \in R$ then add $x = z$.
(d) If $(x = y) \in R$ then add $x^w = y^w$ for some $w \in \mathrm{FR}(S)$.
(e) If $(x = y) \in R$ then add $t^x = t^y$ for some $t \in \mathrm{FR}(S)$.

Define a **consequence** of R to be any statement which can be generated by a finite number of these moves, and define $\langle R\rangle$ to be the set of consequences of R.
Now a congruence is a relation which is: (1) an equivalence relation and (2) respects the rack operation. If we use $=$ instead of $\sim$ for the congruence, then (1) says that the congruence is closed under moves (a), (b) and (c) whilst (2) says it is closed under moves (d) and (e). It follows that the smallest congruence containing R is precisely the set of consequences, $\langle R\rangle$.
Remark The asociated quandle $[\,S : R\,]_q$ has a presentation, obtained by adding to R the relators $a^a = a$ for all $a \in S$. If $[\,S : R\,]$ is a finite presentation then so is $[\,S : R\,]_q$.
Proof Clearly the new relators hold in the associated quandle, but the new rack *is* a quandle because

$$(a^w)^{a^w} = a^{w\overline{w}aw} = a^{aw} = a^w \text{ since } a^a = a.$$

The "Tietze" Theorem
We shall now prove an analogue for racks of the Tietze move theorem for group presentations.
The two basic moves on presentations are the following:
Tietze move 1 Add to R a consequence (or delete from R a consequence) of the other relators.
Tietze move 2 Introduce a new generator x and a new relator $x = a^w$ (where x does not occur in w), or delete such a pair if x occurs nowhere else in the presentation.
There is an equivalent set of moves which are rather more constructive:
1.4.1 Lemma *Tietze moves 1 and 2 are equivalent to the following moves:*
(1) *Repeat a relator (or delete a repeated relator).*
(2) *Conjugate a relator, i.e. replace for example $a^t = b^w$ by $a^{tz} = b^{wz}$.*
(3) *Substitute at primary level, i.e. if $a = b^w \in R$ then we can replace $c^z = a^t$ by $c^z = b^{wt}$ and we can replace $a^z = c^t$ by $b^{wz} = c^t$.*
(4) *Substitute at operator level, i.e. if $a = b^w \in R$ then we can replace $c^{taq} = d^z$ by $c^{t\overline{w}bwq} = d^z$ or $c^z = d^{taq}$ by $c^z = d^{t\overline{w}bwq}$.*
(5) *Introduce a new generator x and a new relator $x = a^w$ (where x does not occur in w), or delete such a pair if x occurs nowhere else in the presentation.*

Proof
Moves (1) to (5) are all equivalent to, or special cases of, the two Tietze moves, so it suffices to show that the two Tietze moves can be achieved by moves (1) to (5). Since Tietze move 2 is move (5) we have to show that any consequence can be introduced (Tietze move 1). But by definition any consequence can be constructed by the relator moves (a) to (e) (in the definition of rack presentation). So we shall start by proving that any of these can be achieved by moves (1) to (5).
To achieve move (a) (to introduce a trivial relator) use the following trick:
Introduce a new generator t and relation $t = a$. Repeat the relation $t = a$ and substitute to obtain $a = a$. Now delete t and the redundant copy of $t = a$.
To achieve move (b) (to add $y = x$ where $(x = y) \in R$) use another trick:
Repeat $x = y$ and substitute to get $y = y$ and then again to get $y = x$.
Finally moves (c), (d) and (e) are precisely moves (2), (3) and (4) in a different form.
It follows that we can use moves (1) to (5) to replace $[\,S : R\,]$ by $[\,S : R \cup T\,]$ where T contains the required consequence $x = y$. Repeat $x = y$ and then reverse the moves used to generate T to delete it. □
1.4.2 Theorem : Tietze move analogue *Suppose that we have two finite presentations $[\,S : R\,]$ and $[\,S' : R'\,]$ of isomorphic racks then the two presentations are related by a finite sequence of Tietze moves.*

Proof
Identify $[\,S:R\,]$ with $[\,S':R'\,]$ by the isomorphism. We shall start with the presentation $[\,S:R\,]$ and move it into the other presentation.
To avoid confusion we shall use the letters $a_1, a_2, \ldots$ for elements of S and $b_1, b_2, \ldots$ for elements of S'. We shall also use $w_1, w_2, \ldots$ for words in the a_i and $z_1, z_2, \ldots$ for words in the b_i.
Step 1 Since the elements of S' are in the rack, each can be expressed in terms of the generators S, i.e.

$$b_1 = a_{(1)}^{w_{(1)}},\ b_2 = a_{(2)}^{w_{(2)}}, \ldots \tag{$*$}$$

where $a_{(i)} = a_j$ some j. Let Q be the set of statements $(*)$. Use Tietze move 2 to introduce the "new" generators $b_1, b_2, \ldots$ together with the set Q of new relators.
Step 2 Since each statement in R' is true in the rack it is a consequence of R and hence can be introduced by Tietze move 1. Thus we can enlarge the set of relators to $R \cup Q \cup R'$.
Step 3 Since S' generates the rack, we can express each element of S in terms of S':

$$a_1 = b_{(1)}^{z_{(1)}},\ a_2 = b_{(2)}^{z_{(2)}}, \ldots$$

(forming a set of statements Q' dual to Q). Since each of the statements in Q' is true in the rack, it can be introduced as a new relator using move 1 again.
At this point we have reached a symmetrical situation. We have $S \cup S'$ as generating set and $R \cup R' \cup Q \cup Q'$ as relators. We now reverse steps 1 to 3 to delete first Q then R and finally S together with Q'. $\square$

Presentations and the Associated Group
Using the proof of the Tietze theorem we can prove that the associated group of a finitely presented rack $[\,S:R\,]$ has a finite presentation as a group — in fact it has the obvious presentation:
Given a rack presentation $[\,S:R\,]$ then we obtain a group presentation by interpreting the elements of R as group equations (i.e. read a^w as $w^{-1}aw$) yielding the group $<S:R>$. It follows from the Tietze theorem that $<S:R>$ is independent of the presentation of the rack, since each of moves (1) to (5) leaves the group $<S:R>$ unchanged. This also follows from the following result:
1.4.3 Lemma *$<S:R>$ is the associated group $As[\,S:R\,]$.*
Proof We shall prove that $<S:R>$ has the universal property of the associated group. First of all there is a rack homomorphism

$$\phi : [\,S:R\,] \to <S:R>_{\text{conj}}$$

because the congruence which defines $[\,S:R\,]$ comprises all (rack) consequences of R. But examining moves (1) to (5) we see that each rack consequence is a group consequence of R (as group relators) and therefore the identity on S extends to a rack homomorphism ϕ.
Now suppose we are given a rack homomorphism $\psi : [\,S:R\,] \to G_{\text{conj}}$ where G is any group. Consider a typical relator $a^w = b^z$ in in R then $\psi(a^w) = \psi(b^z)$ in G_{conj}, but since ψ is a rack homomorphism this implies

$$\psi(\overline{w})\psi(a)\psi(w)\psi(\overline{z})\psi(\overline{b})\psi(z) = 1 \text{ in } G.$$

But this says that $\psi(R)$ as a *group* relation is true in G, and therefore ψ factors via ϕ. Uniqueness of this factoring is clear since both $[\,S:R\,]$ and $<S:R>$ have the same generating set. $\square$
1.4.4 Remark *Racks and group presentations*
The Tietze theorem gives a way of regarding racks as group theoretic objects, namely equivalence classes of **group presentations of conjugacy type** (presentations with relators all of the form $x^w = y^z$) under our moves (1) to (5) of lemma 4.1, moves which all preserve this class.

Links in S^3
We consider links $L : M \subset S^3$. Since M is a codimension two submanifold of S^3, it is the image of a finite disjoint collection of smooth embeddings of S^1 in S^3. In this case, a framing on a component of M can be 'measured'

because the isotopy class of the framing can be regarded, either as (the isotopy class of) a parallel curve (as in §1) or as an integer (the linking number of the component with its parallel curve), cf. **[FR2]**.

If we project the link in general position onto a plane $\mathbb{R}^3 = S^3 - \mathrm{pt}$ we obtain a **diagram**: a finite collection of arcs and circles, the arcs terminating at crossings, as exemplified in the following figure.

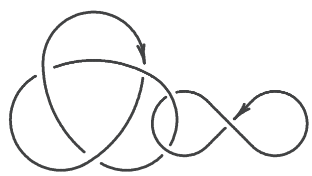


Figure 1.4.1

The chosen orientation on each component of the link is indicated by the arrows in the diagram.

Now a link diagram has a **natural** framing: each component in the diagram has a canonical parallel curve obtained by drawing a curve in the diagram adjacent to the component (indicated as the broken curves in the following figure).

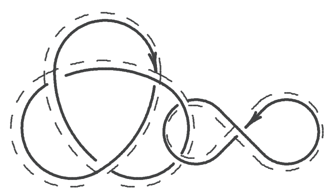


Figure 1.4.2

The following lemma implies that we may suppose that the natural framing and the given framing of L coincide.

1.4.5 Lemma *Given a framed link L in S^3 there is a diagram for L whose natural framing coincides with the given one.*

Proof The Reidemeister Ω_1 moves

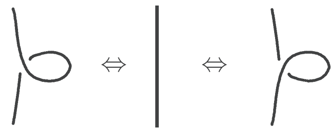


Figure 1.4.3

change the natural framing by ± 1. Hence the diagram can be altered by Ω_1 moves to make the two coincide. □

The rack presentation given by a diagram

A link diagram determines a primary rack presentation by:

(1) Label all arcs (or circles) in the diagram by generators $a, b, c, \ldots$ forming the **generating set** S.

(2) At each crossing write down a relator by the following rule.

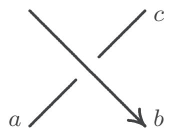


Figure 1.4.4: Write $c = a^b$ or $a = c^{\overline{b}}$.

Note that b crosses a *from the right* as a passes under to become a^b. Note also that the orientation of the under-arc is not used in the rule.

The set of relators gives the **relator set** R.

Extended remark We shall prove shortly that $[S : R]$ is $\Gamma(L)$ (the fundamental rack of L in S^3) which implies that $[S : R]$ is independent of the choice of diagram used to represent L. However, it is worth remarking that this can easily be proved directly, and indeed the *definition* of a rack is tailor made to prove this.

The two rack laws (1.1 axioms 1 and 2) correspond to invariance of $[S : R]$ under the Reidemeister Ω_2 and Ω_3 moves respectively, see the following figures.

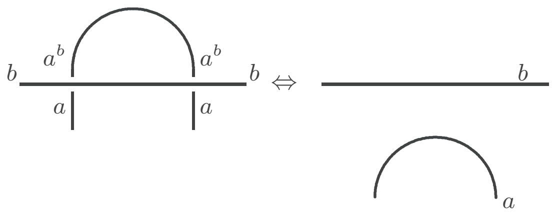


Figure 1.4.5

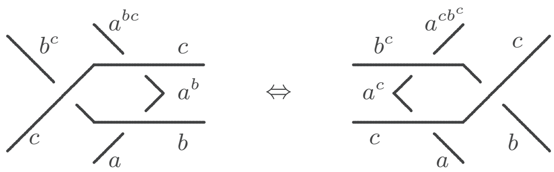


Figure 1.4.6

Now for unframed links, the fundamental quandle is invariant under the final Reidemeister move (the Ω_1 move) by the quandle condition; see the figure.

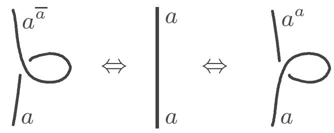


Figure 1.4.7

For framed links, we need a modified version of the Reidemeister move theorem: isotopy classes of framed links correspond to equivalence classes of diagrams under the Ω_2 and Ω_3 moves and the following "double" Ω_1 move:

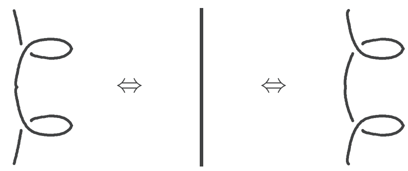


Figure 1.4.8

This follows from the Reidemeister theorem: replace all Ω_1 insertions by double insertions and leave all Ω_1 deletions to the end. Collect all the extra twists on one arc of each component (using Ω_2 and Ω_3 moves). Then the fact that the framings are the same, means that there are (algebraically) the same number of extra twists. But excess pairs of opposite twists can be cancelled using the double Ω_1 move or the following sequence of Ω_2 and Ω_3 moves:

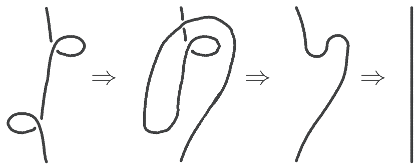


Figure 1.4.9

The figure below shows how the rack laws imply that $[\,S : R\,]$ is independent of the double Ω_1 move. The critical observation on the right is that $c = a^{a\overline{c}}$ implies $c \equiv a$ and also $a^{a\overline{c}} = a$:

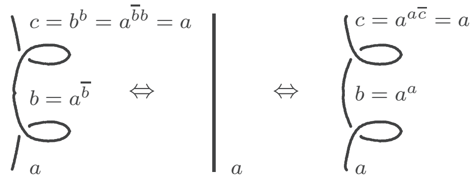


Figure 1.4.10

This completes the extended remark.

1.4.6 Examples

Example 1 *The unknotted circle with framing n. Shown here with $n = 4$.*

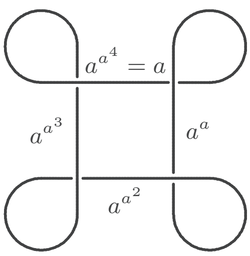


Figure 1.4.11

In the diagram, we have simplified the labels, by making obvious substitutions. The presentation gives the rack

$$[\,a \;:\; a^{a^n} = a\,]$$

i.e. the cyclic rack C_n. Notice that the circle with framing $-n$ has isomorphic fundamental rack. Thus extra structure will be needed to cope with orientations (see end of section 5).

Example 2 *The Hopf link.*

With both framings 0

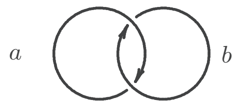


Figure 1.4.12The Hopf link

the rack is

$$[\, a, b \; : \; a^b = a, b^a = b \,].$$

With framings n and m the rack becomes

$$[\, a, b \; : \; a = a^{a^n b}, b = b^{ab^m} \,].$$

Imposing the two quandle relations: $a^a = a, b^b = b$, makes these two racks identical (the fundamental quandle of the unframed Hopf link).

Example 3 *The left hand trefoil knot (framing -3).*

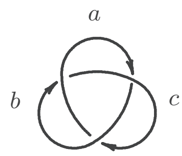


Figure 1.4.13The trefoil knot

The fundamental rack is

$$[\, a, b, c \; : \; a^b = c, \; c^a = b, \; b^c = a \,]$$

or

$$[\, a, b \; : \; a^{ba} = b, \; b^{\bar{b}ab} = a \,].$$

Remark Note that the presentation of the fundamental rack has deficiency zero in contrast to the deficiency of the fundamental group which is one.

Example 4 *The Borromean rings.*

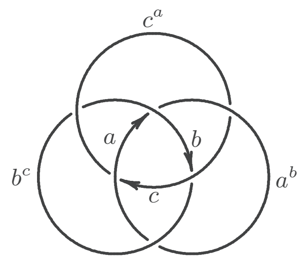


Figure 1.4.14The Borromean rings

The fundamental rack is

$$[\, a, b, c \; : \; c^{a\bar{b}\bar{a}b} = c, \; a^{b\bar{c}\bar{b}c} = a, \; b^{c\bar{a}\bar{c}a} = b \,].$$

1.4.7 Theorem *Let D be a diagram for a framed link L in S^3 and $X = [\, S : R \,]$ the rack presented by D, then $X = \Gamma(L)$, in particular $\Gamma(L)$ has a finite presentation.*

Proof We shall define rack homomorphisms

$$\lambda : \Gamma(L) \to X \qquad \mu : X \to \Gamma(L)$$

such that $\lambda \circ \mu = \mu \circ \lambda = \mathrm{id}$.

Definition of λ An element of $\Gamma(L)$ is represented by a path γ from a point $p \in M^+$ to $*$. Project γ in general position onto the plane of the diagram D and then read from γ an element of $\mathrm{FR}(S)$ as follows. Suppose that the initial point of γ lies on the arc labelled by the generator a and suppose that γ subsequently passes *under* arcs labelled $b, c, d \ldots$ then associate to γ the element $a^{b^\epsilon c^\epsilon d^\epsilon} \ldots$ where $\epsilon = +1$ if the arc labelled b crosses γ in the right-hand sense and $\epsilon = -1$ otherwise. For an illustration see the following figure.

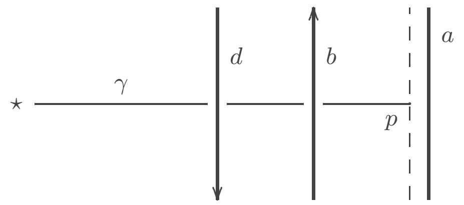


Figure 1.4.15Read $\lambda(\gamma) = a^{b\overline{d}}$.

To prove that λ is well defined we have to check that if we change γ by a homotopy then we get the same element of X.

There are three types of critical stages during the homotopy where the expression for $\lambda(\gamma)$ changes. These are illustrated in the following three typical pictures:

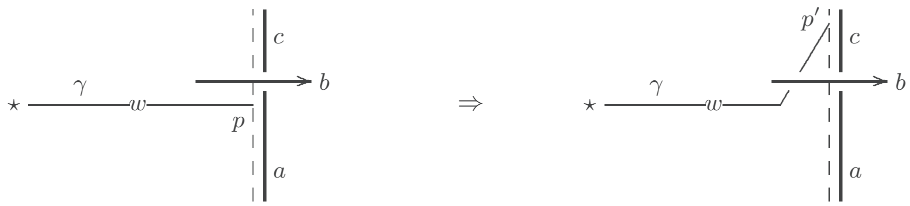


Figure 1.4.16Critical stage type 1

In this picture $c = a^b$ in X. Moving p to p' the value of λ changes to $c^{\overline{b}w} = a^{b\overline{b}w} = a^w$ i.e. we get the same element of X.

Figure 1.4.17Critical stage type 2

In the left-hand picture we read $a^{w_1\overline{e}ew_2}$ whilst in the right-hand picture we read $a^{w_1w_2}$, which is the same element of X.

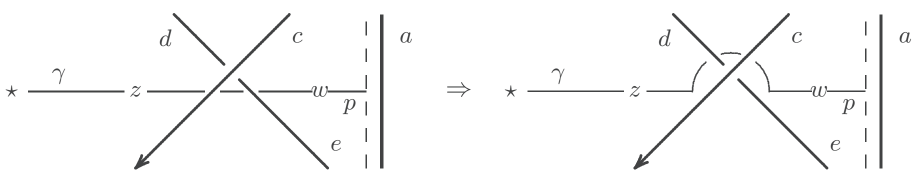


Figure 1.4.18Critical stage type 3

In the left-hand picture we read $a^{we\overline{c}z}$ whilst in the right- hand picture we read $a^{w\overline{c}dz}$. But $e = d^c$ which implies $e \equiv \overline{c}dc$ i.e. $\overline{c}d \equiv e\overline{c}$. These are the same element of X.

Thus $\lambda : \Gamma \to X$ is well defined.

Definition of μ. We start by defining μ on the free rack $\mathrm{FR}(S)$. In what follows we shall misuse notation and write $\mu(a)$ for both the class and the path which represents it. First define $\mu(a)$ where $a \in S$ to be any path from the parallel curve to the arc labelled a to the base-point *over* all other arcs of the diagram. Next suppose that $\mu(x)$ is defined; we will define $\mu(x^c)$ and $\mu(x^{\overline{c}})$ where $c \in S$. This is done by post-composing the path for x with a loop that starts at the base-point, goes over the other arcs to near the arc labelled c, once around this arc in the positive sense for x^c and negative for $x^{\overline{c}}$ and back to the base-point over the other arcs:

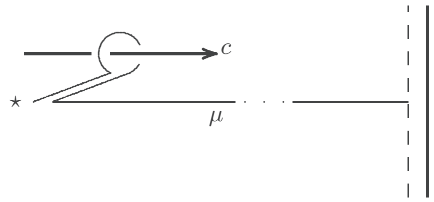


Figure 1.4.19Defining μ

This defines μ on $\mathrm{FR}(S)$. To prove that μ is well-defined we have to check that if $a^w \sim b^z$ in the congruence generated by R then $\mu(a^w)$ is homotopic to $\mu(b^z)$.

But examining moves (1) to (5) of lemma 4.1 we see that the only non-trivial part to be checked is that if a^w is altered by either a primary or a secondary substitution using relators of X then $\mu(a^w)$ is altered by a homotopy.

Primary substitution. Replace c^w by a^{bw} where $c = a^b$ is a relator.

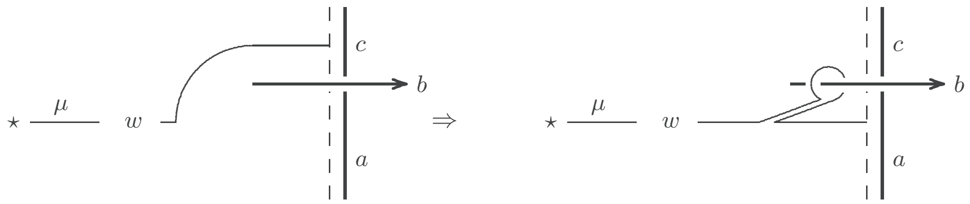


Figure 1.4.20

The picture shows a typical situation, and the required homotopy can be seen.

Secondary substitution. Replace t^{wcz} by $t^{w\overline{b}abz}$ where $c = a^b$.

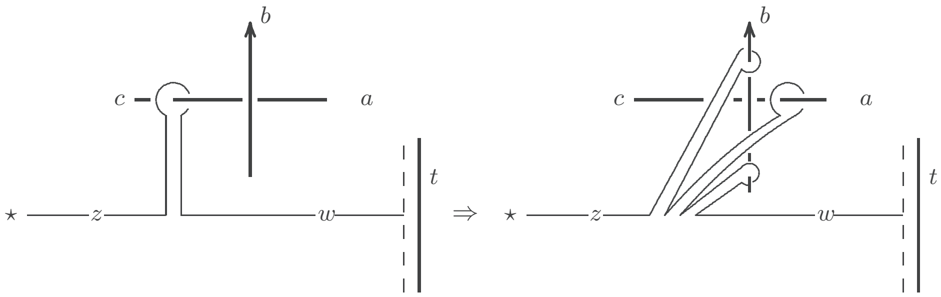


Figure 1.4.21

The picture again shows a typical situation, and the required homotopy can be seen.

We have defined

$$\lambda : \Gamma(L) \to X \qquad \mu : X \to \Gamma(L).$$

It is clear from the definitions that $\lambda \circ \mu = \mathrm{id}_X$ whilst to see that $\mu \circ \lambda = \mathrm{id}_\Gamma$ we take an arbitrary path γ and deform it into $\mu \circ \lambda(\gamma)$ by pulling "feelers" back to the base-point, as illustrated in the following figure.

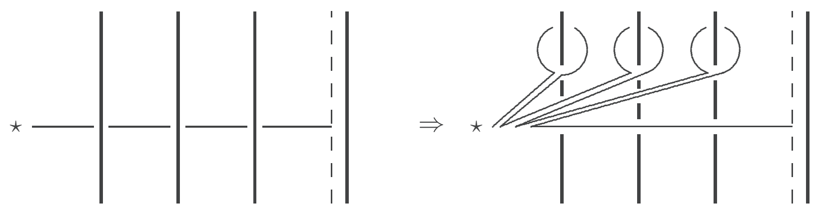


Figure 1.4.22

□

Remarks

(1) If we regard the relators R in the diagram presentation as group relators, then we obtain the Wirtinger presentation of the fundamental group. This verifies the fact we already know: the associated group is the fundamental group of the link, (Corollary 1.3.3).

(2) A similar analysis can be carried out for an embedding of M^n in S^{n+2}: we obtain a "diagram" by projecting onto $\mathbb{R}^{n+1}$ in general position and regarding top dimensional strata (n–dimensional sheets) as "arcs" to be labelled by generators and $(n-1)$–dimensional strata (simple double manifolds) as "crossings" to be labelled by relators. In general position a homotopy between paths only crosses the $(n-1)$ strata and a proof along the lines of the theorem can be given that this determines a finite presentation of the fundamental rack.

(3) There is a general process for obtaining a (not necessarily finite) presentation for *any* codimension 2 embedding by using an analogue of the edge-path presentation for the fundamental group. We leave the details to the interested reader.

We now turn to presentations of the fundamental augmented rack for links in general 3–manifolds. We shall need to enlarge the concept of presentation and this is the content of the remainder of the section.

We shall consider two stages of generalisation. The first (allowing operator relations) does not essentially change the class of racks being considered.

Operator relations

The concept of rack presentation can be generalised by allowing relations which apply only at operator level. For example here is an alternative presentation for the cyclic rack C_n using an operator relation:

$$[a \ : \ a^n \equiv 1].$$

A presentation for a rack with operator relations comprises three sets: a set S of generators, a set R_P of primary relators (as in the first definition of presentation, above) and a set R_O of **operator relators** which are words $w \in F(S)$, to be understood as relations at the operator level $w \equiv 1$.

This concept appears to be more general than the earlier one but in fact it is not:

1.4.8 Lemma *Let S be a set of generators. An operator relator is equivalent to n primary relators where $n = |S|$.*

Proof We shall show that the operator relator $w \equiv 1$ is equivalent to the n primary relators

$$a_1^w = a_1,\ a_2^w = a_2,\ \ldots, a_n^w = a_n \quad \text{where} \quad S = \{a_1, a_2, \ldots, a_n\}. \tag{$*$}$$

Since $w \equiv 1$ implies each of the primary relators $a_i^w = a_i$ it suffices to prove the converse, i.e. that $x^w = x$ is a consequence of $(*)$ for each $x \in \mathrm{FR}(S)$. Write $x = a_j^t,\ t \in F(S)$ say and use induction on the length of t.

Suppose that $t = t_1 a_k^\varepsilon,\ \varepsilon = \pm 1$ and for definiteness suppose that $\varepsilon = +1$. By induction

$$a_j^{t_1} = a_j^{t_1 w}$$

is a consequence of $(*)$. Then using the relator moves we have the following consequences:

$$\begin{aligned} x = a_j^t &= a_j^{t_1 a_k} \\ &= a_j^{t_1 w a_k} \quad \text{(move (d))} \\ &= a_j^{t_1 w a_k^w} \quad \text{(move (e) using } a_k^w = a_k) \\ &= a_j^{t_1 w \overline{w} a_k w} \quad \text{(definition)} \\ &= a_j^{t_1 a_k w} = a_j^{tw} = x^w. \end{aligned}$$

The case $\varepsilon = -1$ is similar. □

General presentations

The final generalisation of presentations is to allow operator **generators** as well:

Definitions Given sets S, T the **extended free rack** $\mathrm{FR}(S,T)$ is defined by

$$\mathrm{FR}(S,T) := S \times F(S \cup T) = \{a^w \mid a \in S, w \in F(S \cup T)\}$$

with rack operation given by

$$(a^w)^{(b^z)} = a^{wz^{-1}bz}.$$

The proof that this is a rack is formally identical to the case of the usual free rack (1.3 example 9).
A **general presentation** of a rack comprises four sets: S_P, S_O the **primary** and **operator** generators and R_P, R_O the **primary** and **operator** relators, where elements of R_P are statements of the form $a^w = b^z$ where $a^w, b^z \in \mathrm{FR}(S_P, S_O)$, and elements of R_O are words $w \in \mathrm{FR}(S_P, S_O)$
Now given a general presentation define the congruence $\sim$ on $\mathrm{FR}(S,T)$ to be the smallest congruence containing:
(1) $x \sim y$ if $(x = y) \in R_P$
(2) $z^x \sim z^y$ if $(x = y) \in R_P$
(3) $z^w \sim z$ if $w \in R_O$.
Then the rack generated by the presentation is defined to be:

$$[\, S_P, S_O : R_P, R_O \,] := \frac{\mathrm{FR}(S,T)}{\sim}.$$

It is worth examining a simple example in some detail because the operator generators in general introduce a non-finiteness in any possible primary generating set.
Example $S_P = \{a\}$ $S_O = \{u\}$ $R_P = R_O = \emptyset$.
Here

$$X = [\, S_P, S_O : R_P, R_O \,] = \mathrm{FR}(\{a\}, \{a, u\}) = \{a^w \mid w \in F(a,u)\}$$

the rack structure is given by

$$(a^w)^{(a^z)} = a^{wz\overline{a}z}.$$

So as a *set* X can be identified with the free group $F(a,u)$ but the rack structure is not conjugacy.
Notice that a^w is in the same orbit as a^z if and only if w and z have the same total degree in u and the set of orbits is in bijective corespondence with the set of cosets of $\mathrm{Ker}(F(a,u) \to F(u))$ in $F(a,u)$. Therefore X has infinitely many orbits and hence cannot have a finite primary presentation.

Presentations of augmented racks
The example makes it clear that a general presentation has operator structure not implied by the rack structure, thus a general presentation fits naturally with the idea of augmented racks.
Let $X = [\, S_P, S_O : R_P, R_O \,]$ be a general presentation and let G be the *group* presented by $< S_P \cup S_O, R_P \cup R_O >$. Then there is a natural map $\partial : X \to G$ and G acts on X by the formula for the rack operation. Therefore X is an augmented rack, which we denote

$$\widehat{X} = [\, S_P, S_O : R_P, R_O \,]_G.$$

We call $\widehat{X}$ the **augmented rack presented by** $[\, S_P, S_O : R_P, R_O \,]$.
Note: Do not confuse G with the associated group $As(X)$. In the simpler case without operator generators, we can see from lemma 4.8 that G and $As(X)$ are in general different (in this simpler case G is a quotient of $As(X)$: the operator relations do not become trivial in $As(X)$, but *central*, see the first line of the proof of the lemma).
In the general case, even if the presentation of X is finite, $As(X)$ does not necessarily have a finite presentation: it is generated by S_P and all conjugates $w^{-1}aw$ where $a \in S_P$ and $w \in F(S_O)$ with relators R_P and commutators of generators by elements of R_O. However there is an important special case in which the rack (and hence the associated group) **does have** a finite primary presenation, given in the lemma below.
Note that since $\partial : X \to G$ is a rack homomorphism from X to G_{conj}, lemma 2.1 gives a homomorphism $\partial_\sharp : As(X) \to G$.
1.4.9 Lemma *Suppose that the presentation of X is finite and that $\partial_\sharp$ is onto, then X has a finite primary presentation.*

Proof We shall show how to replace one operator generator by a finite number of primary generators; the result then follows from lemma 4.8.
Let $t \in S_O$. Since $\partial_\sharp$ is onto, we can write t as an element of $As(X)$ as a product of elements of X, i.e.

$$t = a_1^{w_1} a_2^{w_2} \dots a_n^{w_n} \quad \text{in} \quad As(X).$$

Since the operator group is a quotient of the associated group, this implies that

$$t \equiv a_1^{w_1} a_2^{w_2} \dots a_n^{w_n}.$$

Introduce n new primary generators, $b_1, b_2, \dots, b_n$ together with n primary relators $b_i = a_i^{w_i}$ for $i = 1, 2, \dots, n$. Then we can substitute $b_1 b_2 \dots b_n$ for t at operator level and the operator generator t is now redundant and can be deleted. □

Augmented presentations There is also the useful concept of an **augmented presentation** of an augmented rack. This comprises an explicit group G, a set S of generators, a function $\partial : S \to G$ and a set R of relators which are statements of the form $x = y$ where $x, y \in F(S)$, which respect ∂, i.e. such that $\partial x = \partial y$ in G.
The presentation defines an augmented rack $[\,S : R\,]_G$ by defining $\sim$ on $\mathrm{FR}(S, G)$ to be the smallest congruence containing:
(1) $x \sim y$ if $(x = y) \in R$
(2) $z^x \sim z^y$ if $(x = y) \in R$
(3) $z^{gh} \sim z^k$ if $gh = k$ in G
and setting

$$[\,S : R\,]_G = \frac{\mathrm{FR}(S, G)}{\sim}.$$

If G is finitely presented, then we can convert this to a general presentation, by adding the presentation of G as operator generators and relators.
Remark There are Tietze type theorems for all the more general classes of presentations, which we shall leave the reader to formulate and prove.

The fundamental augmented rack of a link in a 3–manifold

We finish the section by explaining how to read a presentation of the fundamental augmented rack of a link L in a closed orientable 3–manifold from a diagram. Recall that such a manifold can be obtained by surgery on a framed link in S^3. Thus we can represent L by a diagram in which some of the curves (which we think of as 'red' curves) are the surgery curves, and others ('black' curves) are the actual link components.

We label the black arcs by primary generators and the red arcs by operator generators, then at each crossing where the underarc is black we read a primary relator by the usual rule and at each crossing where the underarc is red we read an operator relator. Finally, for each red curve we read a further operator relator by reading round the curve and noting undercrossings. The whole process is illustrated by the example in the following diagram, where the 'red' surgery curve has been drawn with broken lines.

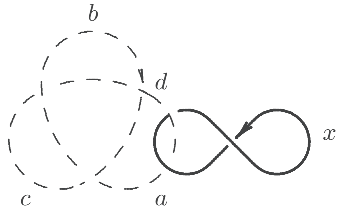


Figure 1.4.23

The presentation is

$$[\{x\},\{a,b,c,d\}:\{x=x^{\overline{x}d}\},\{b\equiv\overline{d}cd,c\equiv\overline{b}db,a\equiv\overline{c}bc,a\equiv\overline{x}dx,\overline{c}\overline{d}\overline{b}x\equiv 1\}]$$

where the first four operator relations come from the crossings with 'red' underarcs and the final operator relation is obtained by reading undercrossings round the red curve.

We leave the details of the proof of the following theorem to the reader.

1.4.10 Theorem *The fundamental augmented rack $\widehat{\Gamma}(L)$ is the augmented rack presented by any diagram for L.*

Sketch of proof The group G corresponding to the presentation is the group $\pi_1(Q_0)$ of the link. That the rack is the fundamental rack is then proved in a similar way to theorem 1.4.7. □

1.4.11 Corollary *The fundamental rack of a link in a 3–sphere has a finite primary presentation.*

Proof This follows from lemma 1.4.9, and the remark below corollary 1.3.3.

1.5. The Main Classification Theorem.

The main result of this section is the classification theorem (theorem 5.1 below) which states that the fundamental augmented rack of an irreducible link in a closed connected 3–manifold is a complete invariant for both the link *and* the ambient 3–manifold.

Since for simply connected 3–manifolds the augmented rack and the plain rack coincide, we deduce that the plain fundamental rack is a complete invariant in this case.

There is also a classification theorem for more general links in 3–manifolds, including any link in S^3, which involves the concept of an **oriented** rack. This will be considered at the end of the section.

Definition A link $L : M \subset Q^3$ is **irreducible** if Q is a closed connected 3–manifold and $Q_0 = \text{closure}(Q - N(M))$ is P^2 irreducible.

Remarks

(1) For a link in S^3 irreducibility is the same as being **non-split**, i.e. L is not the disjoint separated union of two non-trivial sublinks.

(2) For a general 3–manifold irreducibility is equivalent to Q_0 being sufficiently large (because a 3–manifold with boundary tori is sufficiently large if and only if it is P^2 irreducible). Thus other reasonable names for an irreducible link would be **non-split** or **sufficiently large**.

(3) Note however that Q^3 *need not be irreducible*: it is well known that *any* closed connected 3–manifold contains an irreducible link, in fact an irreducible knot. For example, as Dale Rolfson remarked to us, the spine of any open book decomposition of a 3–manifold is an irreducible link.

(4) Irreducibility can be detected algebraically from the fundamental rack. This follows from the following lemma.

1.5.1 Lemma *A semi-framed link $L : M \subset Q^3$ in a closed connected 3–manifold which contains no homotopy 3–sphere summands is reducible if and only if the fundamental augmented rack $\widehat{\Gamma}(L)$ is a non-trivial free product. If Q^3 is a homotopy sphere, then the result is also true with $\widehat{\Gamma}(L)$ replaced by the plain rack $\Gamma(L)$.*

Proof If the link is reducible, then $\widehat{\Gamma}(L)$ is $X * Y$ where X and Y are non-trivial and are the augmented racks of the connected summands; this can easily be checked from definitions.

Conversely, if $\widehat{\Gamma}(L)$ is a non-trivial free product, then $G = \pi_1(Q_0)$ is a non-trivial free product. It then follows from a standard result in 3–manifolds [12; theorem 7.1] that Q_0 is a connected sum, i.e. L is reducible.

If Q is a homotopy sphere then $\pi_1(Q_0)$ is the associated group of the plain rack $\Gamma(L)$ (corollary 3.3), and the proof is similar to the proof for the augmented rack. □

We now give the main result of this section:

1.5.2 Classification theorem

The fundamental augmented rack is a complete invariant for irreducible semi-framed links in closed, connected 3–manifolds.

More precisely suppose that $L : M \subset Q$ and $L' : M' \subset Q'$ are two irreducible semi-framed links in closed, connected 3–manifolds and suppose that the fundamental augmented racks are isomorphic:

$$\widehat{\Gamma}(L) \cong \widehat{\Gamma}(L').$$

Then there is a homeomorphism $Q \to Q'$ carrying M to M' as semi-framed submanifolds.

Proof We shall use Waldhausen's classification theorem for P^2 irreducible, sufficiently large 3–manifolds, see Hempel [**Hem**]; theorem 13.6.

We need a couple of observations.

Observation 1 The orbits of $\widehat{\Gamma}(L)$ are in one–one correspondence with the components of M.

Observation 2 A choice of **base path system** — that is a base point for Q_0 and one on each component of ∂Q_0, in the parallel manifold if appropriate, together with base paths from the base points on ∂Q_0 to the one for Q_0 — is the same as a choice of representatives $\alpha_1, \alpha_2, \ldots, \alpha_t$ of elements of $\widehat{\Gamma}(L)$ one from each orbit.

Thus a choice $x_1, x_2, \ldots, x_t$ of elements of $\widehat{\Gamma}(L)$, one from each orbit, is equivalent to a choice of base path system, up to equivalence generated by homotopy through base path systems.

Now assume that L is framed and choose a base path system for Q_0 and let $x_1, x_2, \ldots, x_t$ be the corresponding choice of elements of $\widehat{\Gamma}(L)$. Then we can read off the following items of the corresponding π_1 system:

$\pi_1(Q_0)$ (the group for $\widehat{\Gamma}(L)$).

The x_i–meridinal subgroup (generated by ∂x_i) for $i = 1, 2, \ldots, t$.

The x_i–longitudinal subgroup, namely its stabiliser in $\pi_1(Q_0)$, see 1.3.5.
Now the isomorphism of $\widehat{\Gamma}(L)$ with $\widehat{\Gamma}(L')$ carries these items to corresponding items in the π_1 system for Q_0' given by the base path system determined by the images of $x_1, x_2, \ldots, x_t$. It follows from Waldhausen's theorem that there is a homeomorphism $Q_0 \to Q_0'$ realising this isomorphism of π_1 systems and the correspondence of listed items. Since meridinal subgroups go to meridinal subgroups, this homeomorphism extends to a homeomorphism $Q \to Q'$ carrying M to M', and since longitudinal subgroups go to longitudinal subgroups, this carries M to M' as framed submanifolds.
The proof in the general (semi-framed) case is the same with the longitudinal subgroups replaced by peripheral subgroups on the unframed components. □

Remark The proof of theorem 1.5.2 makes it clear that the fundamental augmented rack of a link is an algebraic gadget which encapsulates all the information in the fundamental group and peripheral group structure of the complement, without the need for any unnatural choice of base path system. Thus it is the precise algebraic input for Waldhausen's theorem when applied to link complements in closed 3–manifolds.

1.5.3 Lemma
Suppose $L : M \subset Q$ is a semi-framed link and that $p : Q' \to Q$ is a covering. Let $M' = p^{-1}(M)$ and $L' : M' \subset Q'$. Then $\Gamma(L) \cong \Gamma(L')$.

Proof This follows at once from the path lifting property of covering spaces. □

Orientations for racks

The proof of theorem 1.5.2 fails for reducible links because the fundamental rack does not determine the orientation of the components of M. Although the longitudinal *subgroups* are invariant under isomorphism , specific *longitudes* are not. For example consider the disjoint union of two trefoil knots K and K' in S^3. If K say is reflected (changing right-hand to left-hand trefoil or vice versa) then the link changes but the fundamental rack remains the free product of two trefoil racks.
Moreover the conclusion of theorem 1.5.2 only gives a homeomorphism between the links which may not respect orientations of the components of the link or the ambient space. Consider an oriented link L. Then L has an inverse L^* where the orientations of each component are reversed. In section 2 we considered inverted racks where the new binary operation is $a^{\overline{b}}$. The fundamental rack of the inverse link L^* is the inverted rack $\Gamma(L)^*$ of the fundamental rack $\Gamma(L)$. If the orientation of space is reversed the mirror link $\overline{L}$ is obtained. The fundamental rack of the link $\overline{L^*}$ is isomorphic to the fundamental rack of the link L under an isomorphism induced by the space reversing homeomorphism. It follows that the fundamental rack is not a complete invariant for oriented links L which are not equivalent to their inverted mirror image.
We can avoid these difficulties and extend the theorem to general framed links in S^3 by introducing orientations for racks:

1.5.4 Definition An **orientation** for the fundamental rack Γ of a framed link in S^3, is a choice, for each component (orbit of Γ), of generator of the (cyclic) stabiliser.
An oriented rack carries the extra information which enables the orientation of the components to be recreated from the algebra. Using oriented racks, the main theorem extends to arbitrary framed links in 3–manifolds.

1.5.5 Theorem *The oriented fundamental augmented rack is a complete invariant for oriented and semi-framed links in closed connected 3–manifolds which contain no homotopy 3–sphere summands.*

Proof Decompose the fundamental rack $\Gamma(L)$ into a free product of indecomposable racks. By lemma 1.5.1 this corresponds to the decomposition of Q_0 into its connected summands. A similar decomposition applies to $\Gamma(L')$. Now apply theorem 1.5.2 to each piece, and then the resulting homeomorphism carries each piece of L to the corresponding piece of L', with determinate orientations. Thus the homeomorphism can be pieced together along the separating spheres to yield the required homeomorphism. □

Remark The result can be extended to 3–manifolds which are not closed under the extra condition that each connected summand of Q_0 meets $\partial N(M)$.

1.6. Invariants of Links

We have shown that the fundamental rack is a complete invariant for irreducible links in S^3. It follows that, theoretically at least, all invariants of such links can be derived from the fundamental rack. In this section we shall briefly indicate how some old and new invariants are defined in terms of the fundamental rack.
The subject of rack invariants is enormous. Because a rack is such a simple algebraic object (as simple as a group), there are an enormous number of "naturally occurring" racks, some of which we have mentioned in previous sections. Each such rack gives rise to link invariants, and it follows that it is absurdly easy to define new (or apparently new) invariants in this way.
We shall here consider two ways to define invariants:
(1) **Representation invariants**
(2) **Functorial invariants**
Representation invariants are defined by considering rack homomorphisms (representations) to 'known' racks, and functorial invariants are defined by transforming the fundamental rack (by a functor) into a one of a class of racks with more easily computable invariants.
A third method of defining invariants is given by the rack space of the fundamental rack. Any topological invariant of this space is *a fortiori* a link invariant. These invariants are strongly connected with the concepts of **cobordism of links** and the **rack space**, see [**FRB1, FRB2**], investigated in our future paper on the rack space.
Throughout the section, L will denote a semi-framed link and Γ its fundamental rack.

Representation invariants
Let X be any fixed rack. Then the set $\Omega = \mathrm{Hom}(\Gamma, X)$ of representations (i.e. rack homomorphisms) of Γ in X is a link invariant. If X has any extra structure then this set inherits similar extra structure. For example if X is a **topological rack** (i.e. X is a topological space, the rack operation is continuous in both variables, and $a \mapsto a^b$ is a homeomorphism for each $b \in X$) then Ω is also a topological space the **representation space** of Γ in X.
Now suppose that L is a classical link (i.e. a link in S^3) and suppose that D is a diagram for L. Then a representation $\rho \in \Omega$ of Γ in X can interpreted in terms of the diagram D as a **labelling** of D. In other words each arc of D is labelled by an element of X so that at each crossing the labels satisfy the rule $c = a^b$ where a, b, c are indicated as in figure 1.4.4
Any extra structure on Ω has the obvious interpretation in terms of labels.
If X is a finite rack, then the set Ω of representations can be enumerated in a systematic manner by enumerating all labellings satisfying the above rule. We conjecture that there exists a countable sequence $\{X_i\}$ of finite racks such that the sequence $\{\mathrm{Hom}(\Gamma, X_i)\}$ distinguishes all irreducible classical links.
Now if X is a quandle then there is always the **trivial** representation which is obtained by labelling each arc by the same element. Therefore the crudest invariant with a quandle X is the existence or otherwise of a non-trivial representation. If X is not a quandle there may be no representation at all and the crudest invariant is the existence or otherwise of *any* representation.

1.6.1 Examples

Example 1 *Conjugation racks*
Let X be a conjugation rack (i.e. a union of conjugacy classes in a group G) then, by corollary 1.2.2, representations of Γ in X are in bijective correspondence with representations (homomorphisms) of the associated group in G. In the case when L is a classical link, the associated group is the fundamental group and this case has been extensively studied in the literature, see for example the book of Burde and Zieschang [**BZ**]. The next example gives a specific case.

Example 2 *The Dihedral Rack.*
Let L be a classical link and let X be the dihedral rack R_n. Representations in R_n may be described as follows: let the arcs of any diagram of L be coloured with the n colours $0, 1, \ldots, n-1$ such that at each crossing if x_a, x_b, x_c are the three colours assigned to the arcs labelled a, b, c in figure 6.1 then the following equation holds;

$$x_c \equiv 2x_b - x_a \mod n$$

If n is prime it is well known that these equations have a non-constant solution if and only if n divides $\Delta(-1)$, the **determinant** of L.
If $n = 3$ this is the well known property of being 3–colourable. For instance the determinant of the trefoil is 3 and of course the trefoil is 3–colourable.
In general a representation into any finite rack could be interpreted as a suitable "colouring scheme" for the diagram.
Example 3 *The Alexander Rack*
Let Λ be the ring of Laurent polynomials $\mathbb{Z}[t, t^{-1}]$ in the variable t with integer coefficients. Any $\Lambda-$module has the structure of a quandle with the rule $a^b := ta + (1-t)b$. The equations needed for a representation to this quandle are

$$x_c = tx_a + (1-t)x_b$$

where the unknowns x correspond as before to figure 6.1. Let $f(t)$ be an irreducible polynomial over the integers. Then the equations above have a non-trivial solution with t a root of f if and only if f divides the Alexander polynomial $\Delta(t)$. We shall consider a more substantial Alexander invariant later in the section.
Example 4 *The Dodecahedral Rack*
If X is the reflection rack whose elements are the edges of a dodecahedron, then the existence of a representation can used to distinguish knots which have determinant ± 1 and so have no non trivial representation to R_n. An example is 10_{124}, see Joyce **[J]** for details and Azcan **[Azc]** for generalisations using Coxeter groups.
All the above examples have used quandles. The remaining three examples of representation invariants use non-quandle racks.
Example 5 *The Cyclic Rack*
Consider the cyclic rack C_n of order n. Define the **total writhe** of a component of a link to be the framing number of the component plus the sum of its linking numbers with the other components. It is easy to see that the link has a representation to C_n if and only if the total writhe of each component is divisible by n.
Example 6 *The (t, s)–Rack*
The (t, s)–rack is a generalisation of the Alexander rack defined in 1.3 example 6. The two dimensional real plane has the structure of a Λ_s module if t, s act linearly as the matrices

$$\begin{pmatrix} u & 0 \\ 0 & u \end{pmatrix} \quad \begin{pmatrix} -u & -u \\ 1 & 1 \end{pmatrix}$$

where $u \neq 0$.
Let X denote this rack. If we seek a non-trivial representation of Γ in X, then we find a number of linear equations in u have to be satisfied. For instance if we take the standard diagram representing the trefoil and with writhe 3 then the existence of a representation depends on the solution of the equations

$$\begin{array}{ccccccccc} ux_1 & - & y_1 & - & uz_1 & - & uz_2 & = & 0 \\ ux_2 & - & y_2 & + & z_1 & + & z_2 & = & 0 \\ uy_1 & - & z_1 & - & ux_1 & - & ux_2 & = & 0 \\ uy_2 & - & z_2 & + & x_1 & + & x_2 & = & 0 \\ uz_1 & - & x_1 & - & uz_1 & - & uz_2 & = & 0 \\ uz_2 & - & x_2 & + & y_1 & + & y_2 & = & 0. \end{array}$$

The polynomial which is the determinant of the matrix of these equations is an invariant of the framed knot.
A more general Λ_s–module structure on a 2–dimensional vector space is given by the matrices

$$\begin{pmatrix} u & 0 \\ 0 & u \end{pmatrix} \quad \begin{pmatrix} -u-x & -y(u+x) \\ y^{-1}(1+x) & 1+x \end{pmatrix}$$

where u and y are non-zero. This leads to a 3–variable polynomial invariant. We have not yet decided whether these "new" polynomials contain any really new information.
These polynomials are invariants of the *framed* knot or link. However there are corresponding invariants of the *unframed* knot or link, see remark 6.3 below
Example 7 *Matrix racks*
There is a way of associating a rack to any matrix or affine group and this leads to whole families of new and computable polynomial invariants for knots and links.

The general construction is this: Given a set X with an action by the group G, we can define a rack structure on $G \times X$ by the formula

$$(g,x)^{(h,y)} := (h^{-1}gh,\ x \cdot h). \qquad (*)$$

If we apply this in the case when G is a matrix or affine group and X the corresponding vector space, then the formula gives a rack structure on (some subset of) a linear space.
A simple example is given by the group of dilations acting on the plane:

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b \\ c \end{pmatrix} \qquad a,b,c \in \mathbb{R} \quad a \neq 0$$

then the rack structure is given by

$$(a,b,c,x,y)^{(d,e,f,z,t)} = (a, db - ae + e, dc - af + f, dx + f, dy + e)$$

where a,b,c represent the element of the group as in the equation above, and x,y are the coordinates of the point of the plane.
The formula can be used to define several polynomial invariants including the multi-variable Alexander polynomial (in fact this polynomial comes from an even simpler example: the group of affine transformations of a 1–dimensional space, see example 3 below 1.7.6 and Devine [**Dev**] for details).

1.6.2 Remarks

(1) When using a non-quandle rack, the invariants found for a classical link will depend in general on the *framing* of the link (i.e. the writhe of the diagram). However there is a way to define an infinite family of corresponding invariants of the *unframed* link: we choose arbitrarily integers corresponding to the components of the link and then we choose to frame the link with the unique framing given by setting the writhes equal to these integers. For each choice of integers, we have in this way an invariant of the unframed link.

(2) The general construction for a rack given in $(*)$ above can be further generalised. Let X be a set on which a *rack* Y acts (i.e, for each $y \in Y$ we have a bijection $x \mapsto x \cdot y$ of X such that y^z acts like $\overline{z}yz$, where $x \cdot \overline{z}$ means the pre-image of x under the action of z), then almost the same formula gives a rack structure on $Y \times X$

$$(y,x)^{(t,u)} := (y^t, x \cdot t).$$

Functorial Invariants

Invariants of links can be obtained by applying a functor to the fundamental rack. One advantage of this method of defining invariants is that they automatically apply to arbitrary racks and not just to classical racks.

1.6.3 Examples

Example 1 *The Associated Group*

The associated group functor is an example. However, as we have seen in proposition 1.3.2, this leads to an existing topological invariant $\pi_2(Q, Q_0)$, the associated group of the link.

Example 2 *The Alexander module*

The following is, we believe, a new invariant which generalises the definition of the Alexander polynomial of a knot or link.

Let Λ be the ring of Laurent polynomials $\mathbb{Z}[t, t^{-1}]$ in the variable t with integer coefficients. As we have seen above any Λ–module has the structure of a quandle with the rule $a^b := ta + (1-t)b$. This means that $\mathbb{MOD}$, the category of Λ–modules is a subcategory of $\mathbb{RACKS}$, the category of racks. There is a functor

$$A : \mathbb{RACKS}\mathbf{a}\mathbb{MOD}$$

the **Alexander functor**, which is a left adjoint to the inclusion defined as follows;

If X is a rack let $A(X)$ be the Λ–module with generators $\{u_a | a \in X\}$ and relations

$$u_{a^b} = tu_a + (1-t)u_b \qquad a, b \in X.$$

In the case of classical links the Alexander functor takes the fundamental rack into the usual Alexander module. However our invariant is defined for an arbitrary codimension 2 link.

For the generalisation to Alexander modules with many variables see Devine.

Example 3 *The (t, s)–module*

The Alexander module can be generalised to give a non-quandle module by replacing Λ by Λ_s in the last example and modifying the relations to

$$u_{a^b} = tu_a + su_b \qquad a, b \in X.$$

Example 4 *Verbal Groups*

The associated group functor can be generalised.

Let X be a rack. We consider the group with generators $\{g_x | x \in X\}$ and relations

$$g_{a^b} = w(g_a, g_b)$$

where $w(x, y)$ is a fixed word in two variables. In order for the group to be well defined, w has to satisfy two conditions which are the analogues of the rack laws. Examples are $w = y^{-1}xy$ which yields the associated group and $w = yx^{-1}y$ which yields the **associated core group**. Kelly **[Kel]** has proved that these are the *only* such examples. However, as Kelly shows, there is a further generalisation: replace g_x by a fixed set of say n generators and w by n words in $2n$ variables; there are then many new examples. For instance in the case $n = 2$, Kelly found, in a computer search, more than 500 new examples including several infinite families.

There is much work to be done to categorise and to calculate these new invariants.

1.7. Racks and Braids

In this section we shall explore the relationship between racks and braids. We shall show that there is a faithful representation of the braid group B_n on n strings in the automorphism group $Aut(\mathrm{FR}_n)$ of the free rack on n elements. This representation can be used to define and to calculate link invariants. Moreover a classical result of Artin can be adapted to characterise the image of the represention, and we shall apply this result to give a characterisation of the fundamental rack of a link in a 3–manifold.

We call a rack **classical** if it is the fundamental rack of a link in S^3. We shall give separate characterisations for classical racks and for the fundamental augmented racks of links in general oriented 3–manifolds. These results complement the main classification theorem of section 5.

The braid groups

Let $L_n : \{P_1, \dots, P_n\} \subset D^2$ be a fixed link comprising n points in the interior of the 2–disc. Let L_n^+ be the framed version of L_n, which we can think of as comprising n standard little discs which are reduced copies of D^2.

A **braid** on n strings is an equivalence class of links $\beta : A \subset D^2 \times I$, where A comprises n arcs each of which meets every level $D^2 \times \{t\}$ in a single point, and such that $\beta \cap (D^2 \times \{i\}) = L_n$ for $i = 0, 1$. The equivalence is isotopy through similar links. Similarly a **framed braid** is an equivalence class of framed links of the same type, such that $\beta \cap (D^2 \times \{i\}) = L_n^+$ for $i = 0, 1$.

It is well known that braids on n strings form a group the **braid group** B_n, with composition defined by stacking two braids one above the other. Similarly, framed braids form the **framed braid group** FB_n. A braid or framed braid determines a permutation $\pi \in S_n$ of $\{P_1, \dots, P_n\}$ by following the strings from top to bottom, and this defines surjective homomorphisms $B_n \to S_n$ and $FB_n \to S_n$.

Now an unframed braid has a standard framing defined by transporting the little discs down the strings, keeping them parallel throughout. In general a framed braid can be regarded as a braid with an integer attached to each string, which counts the total rotation of the little disc as the string is traversed from top to bottom. The standard framing corresponding to zeros. In composition the integers on the two pieces of the string are added.

Thus FB_n is the **wreath product** $\mathbb{Z} \wr B_n$ that is the semi-direct product of $\mathbb{Z}^n$ with B_n where the action of B_n on $\mathbb{Z}^n$ is given by permuting the factors using the homomorphism of B_n to S_n.

The representations of B_n and FB_n

Let β be a framed braid. We shall consider the fundamental racks $\Gamma(L_n^+)$ and $\Gamma(\beta)$. Now a braid can be 'unbraided' and any framing can be 'untwisted' hence, as a link, $\beta \cong L_n^+ \times I$. Therefore the fundamental racks are isomorphic. Moreover the fundamental rack $\Gamma(L_n^+)$ can be identified with the free rack $\mathrm{FR}_n = \mathrm{FR}\{a_1, a_2, \dots, a_n\}$ on n generators as pictured:

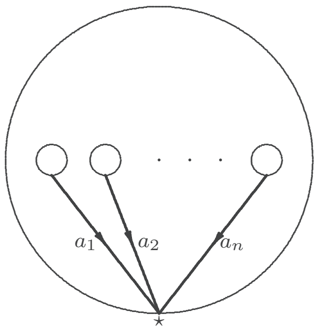


Figure 1.7.24 Generators of $\Gamma(L_n^+)$

Thus we have isomorphisms

$$\mathrm{FR}_n = \Gamma(L_n^+) \xrightarrow{i_0} \Gamma(\beta) \xleftarrow{i_1} \Gamma(L_n^+) = \mathrm{FR}_n$$

where i_0, i_1 are induced by inclusions of L_n^+ in β at $D^2 \times 0$, $D^2 \times 1$ respectively. The composition $i_1^{-1} i_0$ is an isomorphism

$$\beta_* : \mathrm{FR}_n \to \mathrm{FR}_n$$

of the free rack.

It is easy to check that $\beta \mapsto \beta_*$ is a homomorphism

$$\lambda : FB_n \to Aut(\mathrm{FR}_n).$$

In theorem 1.7.3 below we shall show that λ is a faithful representation of the braid groups in the automorphism group of the free rack.

Example

Let σ_i be the braid which is the simple right handed interchange of the ith and $(i+1)$st strings and which keeps the other strings fixed. Let σ_i have the standard framing.

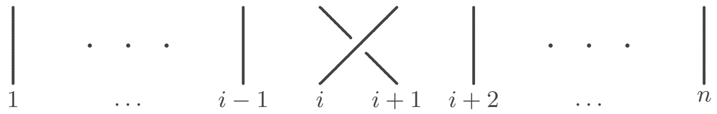


Figure 1.7.25 The braid σ_i

Then σ_i induces the automorphism σ_{i*}

$$\begin{cases} a_i & \mapsto & a_{i+1}^{a_i} \\ a_{i+1} & \mapsto & a_i \\ a_j & \mapsto & a_j \quad j \neq i, i+1. \end{cases}$$

The automorphism β_* for any braid β with standard framing can now be calculated since β can be written as a word in the σ_i with σ_i^{-1} represented by a left handed interchange.

The general form of an automorphism of FR_n is $a_i \mapsto a_{\pi(i)}^{w_i}$ where π is a permutation of the set $\{1, 2, \ldots, n\}$ and w_i for $i = 1, 2, \ldots, n$ are words in the free group $F(a_1, \ldots, a_n)$, which satisfy certain conditions obtained by considering the analogue of Nielsen theory for racks. This theory is given in an appendix. We shall not need to consider these conditions in detail in this section. In general the above formula defines a monomorphism of FR_n to itself, (indeed the same formula defines a monomorphism of the free group) and this is also proved in the appendix, see corollary 1.8.6.

1.7.1 Example Let $\beta = \sigma_1^{-1}\sigma_2\sigma_1^{-1}$ be the following illustrated braid, again with standard framing:

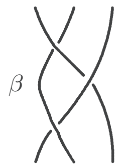


Figure 1.7.26 $\sigma_1^{-1}\sigma_2\sigma_1^{-1}$

then β_* is given by

$$\pi = (13) \quad w_1 = a_2^{-1}a_1^{-1}a_2 \quad w_2 = a_2^{-1}a_1a_2a_3a_2^{-1}a_1^{-1}a_2 \quad w_3 = a_2.$$

For a braid with non-standard framing, we decompose the braid into interchanges and twists (the identity braid but with one string, say the i–th, framed ± 1). The automorphism for a twist on the i–th string is

$$\begin{cases} a_i & \mapsto & a_i^{\pm a_i} \\ a_j & \mapsto & a_j \quad j \neq i. \end{cases}$$

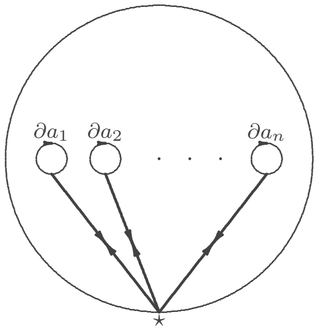
Later in the section we shall show how to read the automorphism β_* from the diagram, without decomposing β into elementary interchanges and twists.

Similar considerations apply to unframed braids, and we have the representation $\overline{\lambda} : B_n \to Aut(\mathrm{FQ}_n)$ where FQ_n is the free quandle on n generators. An element of $Aut(\mathrm{FQ}_n)$ is again defined by a permutation π and n words w_i but now the words w_i are only determined up to premultiplication by powers of $a_{\pi(i)}$ (because $a_{\pi(i)}^{a_{\pi(i)}} = a_{\pi(i)}$ in a quandle). Note that FQ_n is the associated quandle to FR_n and that the natural homomorphism $Aut(\mathrm{FR}_n) \to Aut(\mathrm{FQ}_n)$ is given by using the same permuation π and words w_i. Further we can regard $Aut(\mathrm{FQ}_n)$ as a subgroup of $Aut(\mathrm{FR}_n)$ by choosing to premultiply w_i by the unique power of $a_{\pi(i)}$ which makes the total power of $a_{\pi(i)}$ in w_i zero. A general element of $Aut(\mathrm{FR}_n)$ is then an element of $Aut(\mathrm{FQ}_n)$ together with an integer for each w_i giving the total degree of $a_{\pi(i)}$. Thus we have another wreath product

$$Aut(\mathrm{FR}_n) = \mathbb{Z} \wr Aut(\mathrm{FQ}_n)$$

and the homomorphism $\lambda : FB_n \to Aut(\mathrm{FR}_n)$ carries one wreath product structure into the other.

Finally we have connections with the free group F_n on n generators. The associated groups of both $\Gamma(L_n)$ and $\Gamma(L_n^+)$ can be identified with the fundamental group $\pi_1(D_0^2)$ of the disc minus the n little discs, which is the free group $F_n = F(\partial a_1, \partial a_2, \ldots, \partial a_n)$ on the n loops illustrated:



We shall usually use the symbols $a_1, a_2, \ldots, a_n$ for the generators of F_n, rather than the more accurate $\partial a_1, \partial a_2, \ldots, \partial a_n$, whenever no confusion is likely to arise. Given a possibly framed braid β, then β_* induces an automorphism $\beta_{*\sharp}$ of the free group F_n which can be described in the same way as β_* using the fundamental groups of β and L_n in place of the fundamental racks. Thus we have further representations of B_n and FB_n in the automorphism group $Aut(F_n)$ of the free group.
In summary we have the following commuting diagram.

$$\begin{array}{ccccc}
B_n & \xrightarrow{\overline{\lambda}} & Aut(\mathrm{FQ}_n) & \longrightarrow & Aut(F_n) \\
\downarrow & & \downarrow & & \\
\mathbb{Z}\wr B_n & \xrightarrow{\lambda} & \mathbb{Z}\wr Aut(\mathrm{FQ}_n) & & \\
\| & & \| & & \\
FB_n & & Aut(\mathrm{FR}_n) & &
\end{array}$$

A classical result of Artin shows that $B_n \to Aut(F_n)$ is injective. In fact all the maps in the diagram are injective.
Remark It can be readily seen that the homomorphism $Aut(\mathrm{FQ}_n) \to Aut(F_n)$ is injective. This is because the image of the automorphism $a_i \mapsto a_{\pi(i)}^{w_i}$ in $Aut(F_n)$ is given by $a_i \mapsto w_i^{-1}a_{\pi(i)}w_i$ and then $a_{\pi(i)}$ and w_i are determined by this word in the free group up to premultiplication of w_i by a power of $a_{\pi(i)}$.

The Artin condition

Definition The permutation $\pi \in S_n$ and words $w_i \in F_n$, $i = 1, 2, \ldots n$ are said to satisfy the **Artin condition** if the identity

$$\prod_{i=1}^{n} a_i = \prod_{i=1}^{n} w_i^{-1} a_{\pi(i)} w_i$$

holds in the free group. In this case, we say that the homomorphism $g : \mathrm{FR}_n \to \mathrm{FR}_n$ defined by $a_i \mapsto a_{\pi(i)}^{w_i}$ is **Artin**.
In theorem 1.7.3 below, we shall show that Artin homomorphisms are in fact automorphisms.
For example in 1.7.1 above

$$\begin{aligned}
&\prod_{i=1}^{n} w_i^{-1} a_{\pi(i)} w_i \\
&\quad = (a_2^{-1}a_1a_2\ a_3\ a_2^{-1}a_1^{-1}a_2)(a_2^{-1}a_1a_2a_3^{-1}a_2^{-1}a_1^{-1}a_2\ a_2\ a_2^{-1}a_1a_2a_3a_2^{-1}a_1^{-1}a_2)(a_2^{-1}\ a_1\ a_2) \\
&\quad = a_1a_2a_3.
\end{aligned}$$

A similar definition works for FQ_n since the words $w_i^{-1}a_{\pi(i)}w_i$ are well defined by the automorphism.
1.7.2 Lemma *An automorphism $g \in Aut(FR_n)$ or $Aut(FQ_n)$ determined by a braid or framed braid is Artin.*
Proof The element $\partial a_1 \partial a_2 \cdots \partial a_n$ of the free group regarded as an element of $\pi_1(D_0^2)$ is represented by the boundary ∂D^2 of the disc. But, from the definition of the induced automorphism of the free group, this element maps to itself, since ∂D^2 can be homotoped down $\partial D^2 \times I$ outside the braid, from top to bottom.

But the right hand side of the Artin condition is precisely the image of this element under the induced automorphism. □

1.7.3 Theorem *An Artin homomorphism is an automorphism. Moreover the homomorphism* $\lambda : FB_n \to Aut(FR_n)$ *defined above is injective and the image of* λ *consists precisely of Artin automorphisms.*

Remark By the commuting diagram given earlier, the theorem is equivalent to the classical result of Artin, see Birman [**Bir**]; theorem 1.9, where a combinatorial proof can be found. However, the rephrasing of the result in terms of racks has an advantage because it leads to a simple geometric proof along the lines of the main result of section 5.

Proof To prove the theorem we need to show that if g is an Artin homomorphism of FR_n to itself then there is a unique braid β such that $\beta_* = g$. We shall start by proving the there is a homeomorphism h of D^2_0 (i.e. D^2 with n little discs around $P_1, P_2, \ldots, P_n$ removed) such that $h|\partial D^2 = \mathrm{id}$ and $h_* = g$. Moreover h is unique up to isotopy. To prove this we shall use the 2–dimensional version of Waldhausen's theorem [**Wal**]; theorem 13.1. A homeomorphism h of D^2_0 is determined up to isotopy by its effect on $\pi_1(D^2_0) = F_n$ and the peripheral structure. But the elements $a_1, a_2, \ldots, a_n \in \Gamma$ define a base path system for D^2_0 (see the proof of theorem 1.5.2) and their images $g(a_1), g(a_2), \ldots, g(a_n)$ another base path system. Moreover by corollary 1.8.6 g induces a monomorphism $g_\sharp$ of $\pi_1(D^2_0)$ to itself and we have the peripheral structure given by the inner loops $\partial a_1, \partial a_2, \ldots, \partial a_n$ and the outside loop $\partial a_1 \partial a_2 \cdots \partial a_n$ which is the image in $As(\Gamma)$ of $a_1 a_2 \cdots a_n$. Thus the n inner loops are mapped by $g_\sharp$ to corresponding inner loops in the other system and the Artin condition says precisely that the outer loop maps to itself. Thus by the 'Waldhausen' theorem quoted there is a homeomorphism h, unique up to isotopy, such that $h|\partial D^2 = \mathrm{id}$ and $h_* = g$.

The connection with braids is well known. We extend h to a homeomorphism $D^2 \to D^2$ by inserting the little discs (which are permuted by parallel translations). Then there is an isotopy relative to the boundary of the identity to h and this isotopy restricted to the little discs gives a braid β such that $\beta_* = g$. This sets up an isomorphism between the braid group and the group of isotopy classes of homeomorphisms of D^2 which satisfy the same conditions as h, therefore β is unique. □

Reading the automorphism from the braid

We now give the promised recipe for reading the automorphism β_* from the braid β. Assume that we have a diagram for the braid and assume first that the framing is standard. Orient the strings of the braid downward and label the fixed points at the top of the braid $P_1^+, P_2^+, \ldots, P_n^+$. In a similar fashion label the fixed points at the bottom of the braid $P_1^-, P_2^-, \ldots, P_n^-$.

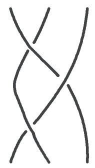
Starting at the bottom of the braid label the arc starting at P_j^- by $a_j, j = 1, 2, \ldots, n$. Now continue up the braid and label arcs which start at a crossing points using the rules for labelling arcs given in section 4. The labels are all elements of the free rack $\mathrm{FR}(a_1, a_2, \ldots, a_n)$.

Suppose now that the top arc of the string which started at P_j^- finishes at P_i^+ with label a_j^w. Then put $w_i = w$ and $\pi(i) = j$. In other words the label on P_i^+ is $a_{\pi(i)}^{w_i}$, and the automorphism can be read from the top labels, as we see from the following example.

Example We shall check the rule for the braid of example 1.7.1:

$$\beta_*\left\{\begin{array}{ccc} a_1 & a_2 & a_3 \\ \downarrow & \downarrow & \downarrow \\ a & b & c \end{array}\right.$$

β



$$a = a_3^{\overline{a_1^{a_2}}} = a_3^{\overline{a_2 a_1} a_2}$$

$$b = a_2^{a} = a_2^{\overline{a_2} a_1 a_2 a_3 \overline{a_2 a_1} a_2}$$

$$c = a_1^{a_2}$$

$$\beta_* \quad \left\{\begin{array}{ccc} a_1 & a_2 & a_3 \\ \downarrow & \downarrow & \downarrow \\ b & c & a_1^{a_2} \end{array}\right.$$

$$b = a_3^{\overline{a_1^{a_2}}} = a_3^{\overline{a_2 a_1} a_2}$$

β

$$c = a_2^{b} = a_2^{\overline{a_2} a_1 a_2 a_3 \overline{a_2 a_1} a_2}$$

$a_1 \quad a_2 \quad a_3$

For general framings, we correct the framing of the diagram by inserting little twists (see lemma 4.5), and use the same method. By the results of section 4 (see in particular figures 3 to 6) it makes no difference where the little twists are inserted, or how the braid is respresented as a diagram.

To prove that the method gives the correct result, we make the following observations:

(1) It gives the correct result for a braid which is a simple interchange (σ_i) or a single twist. This is readily checked by hand.

(2) The method gives a homomorphism $\mu : B_n \to Aut(\mathrm{FR}_n)$.
To see this consider the effect of stacking the braid β' on top of the braid β. If the i–th point at the bottom of β is labelled a_i and at the top is labelled $a_{\pi(i)}^{w_i}$, then the labels at the top of the combined braid are obtained from those for β' by substituting $a_{\pi(i)}^{w_i}$ for a_i. But this is precisely how the composition $\beta'_* \circ \beta_*$ of the two automorphism of FR_n is formed.

Since any framed braid can be decomposed into simple interchanges and twists, it now follows that $\lambda = \mu$, i.e. the method gives the correct result.

We can now give our characterisation of classical racks.

1.7.4 Theorem: Characterisation of classical racks

A rack is the fundamental rack of a framed link in S^3 if and only if it has a primary presentation of the form

$$[\, a_1, a_2, \ldots, a_n \ : \ a_1 = a_{\pi(1)}^{w_1}, a_2 = a_{\pi(2)}^{w_2}, \ldots, a_n = a_{\pi(n)}^{w_n} \,]$$

where $\pi, w_1, w_2, \ldots, w_n$ satisfy the Artin condition.

Proof Let β be a framed braid. The **closure** $C(\beta)$ of β is the link in S^3 obtained by joining the top of the braid to the bottom by n arcs 'round the back' with standard framing.

By a theorem of Alexander (see e.g. [2]) any framed link $L : M \subset S^3$ can be represented as the closure $C(\beta)$ of a braid (in fact this can be done so that the framing of L is given by the standard framing of β though we shall not need to use this fact).

Now we can read a presentation for the fundamental rack $\Gamma(L)$ from the diagram for $C(\beta)$ by the methods of section 4. Moreover, by using the same lebels as in the above discussion, we see that $\Gamma(L)$ has the presentation $[\ldots, a_i, \ldots \mid \ldots, a_i = a_{\pi(i)}^{w_i}, \ldots]$ where $a_i \mapsto a_{\pi(i)}^{w_i}$ is the atomorphism of FR_n induced by β. Thus $\Gamma(L)$ has a presentation of the required form by theorem 7.3.

Conversely suppose we are given a rack Γ with a presentation of this form then, again by theorem 7.3, there is a braid β which induces the automorphism $a_i \mapsto a_{\pi(i)}^{w_i}$. Then the closure of β has fundamental rack isomorphic to Γ.
□

Remarks

(1) The theorem has content. For example any rack whose associated group has torsion is not classical.

(2) There is a similar characterisation of the fundamental racks of semi-framed or unframed links in S^3. Here the presentation has the same form but with extra relations $a_j^{a_j} = a_j$ corresponding to the unframed components. The proof is essentially the same.

(3) Markov's theorem (see e.g. [2]) can be combined with the theorem to give an algebraic classification of classical links. By Markov's theorem, any two braids which have isotopic closures are related by a series of moves. The resulting presentations of the fundamental rack are therefore also related by a series of moves through presentations of the same form. We leave the details to an interested reader.

Links in general 3–manifolds

We now adapt the last result to give an algebraic characterisation for the fundamental augmented rack of a link in any closed orientable 3–manifold. As at the end of section 4, we shall regard the 3–manifold as given by surgery on a link in S^3 and by Alexander's theorem we can represent a diagram for the surgery curves together with the actual link as the closure of a braid β, with t 'red' strings and n 'black' strings, where the closures of the 'red' strings represent the surgery curves and the 'black' the link curves. By a suitable conjugacy of the braid, we may suppose that the t 'red' strings start (and finish) at the t right-most positions.

Let $\pi \in S_n$ be the permutation given by the black strings and $\sigma \in S_t$ the permutation given by the red strings. Denote by $\pi|\sigma$ the permutation in S_{n+t} of all $n+t$ strings.

We need to decompose σ into its cycle decomposition. By a further conjugacy of the braid we can suppose that this decomposition is in fact of the form

$$\sigma = (1, 2, \ldots, l_1)(l_1 + 1, l_1 + 2, \ldots, l_2) \cdots (l_{p-1} + 1, \ldots, l_p).$$

Suppose that β induces the automorphism

$$\begin{aligned} a_i &\mapsto a_{\pi(i)}^{w_i} \quad i = 1, 2, \ldots, n \\ b_i &\mapsto b_{\sigma(i)}^{z_i} \quad i = 1, 2, \ldots, t \end{aligned}$$

of the free rack FR_{n+t}, where we have used $b_1, b_2, \ldots, b_n$ for the last t generators in order to distinguish the surgery strings from the genuine link strings. Note that w_i and z_i are words in $F_{n+t} = F(a_1, \ldots, a_n, b_1, \ldots, b_t)$.

We can now read off an augmented presentation of the fundamental augmented rack of L using the recipe given at the end of section 4.

S_P The primary generators $\{a_1, a_2, \ldots, a_n\}$.

S_O The operator generators $\{b_1, b_2, \ldots, b_t\}$.

R_P The primary relators $\{a_i = a_{\pi(i)}, i = 1, 2, \ldots, n\}$.

R_O The operator relators $\{b_i \equiv b_{\sigma(i)}, i = 1, 2, \ldots, t\}$, together with the p further operator relators

$$\{z_{l_1} z_{l_1 - 1} \cdots z_1, \;\; z_{l_2} z_{l_2+1} \cdots z_{l_1+1}, \;\; \ldots, \;\; z_{l_p} z_{l_p+1} \cdots z_{l_{p-1}+1}\}.$$

The last set of operator relators are the ones which come from reading around the surgery curves noting undercrossings.

The following theorem is proved in a similar way to the last theorem:

1.7.5 Theorem: Characterisation of fundamental augmented racks in 3–manifolds

An augmented rack is the fundamental rack of a framed link in a closed oriented 3–manifold if and only is it has a presentation of the form listed above such that the permutation $\pi|\sigma$ and the words $w_1, w_2, \ldots, w_n, z_1, z_2, \ldots, z_t$ satisfy the Artin condition.

Remarks

(1) Again there is a semi-framed version of the theorem, which we leave the reader to formulate.

(2) Again the theorem can be combined with a 'Markov' theorem to give an algebraic classification of links in terms of moves through presentations of the same type. The appropriate theorem here is the extension of Markov's theorem to general 3–manifolds contained in Lambropoulou's thesis [21].

Invariants

The connections between braids and racks can be used to define and calculate invariants of classical links. As with all discussions of invariants in this paper, we shall content ourselves here with a brief outline of the methods and return to discuss the subject in greater depth in future papers.

The key idea is an extension of an idea of Brieskorn [3; proposition 3.1]. Given any rack X, there is an action of the automorphism group $Aut(\mathrm{FR}_n)$ on X^n as follows. Let $f \in Aut(\mathrm{FR}_n)$ and let $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in X^n$. Define $j_{\mathbf{x}}$ to be the unique rack homomorphism $\mathrm{FR}_n \to X$ such that $a_i \mapsto x_i$ for each i. Then the action f_* of f on X^n is given by

$$f_*(\mathbf{x}) := (j_{\mathbf{x}}(f(a_1)), j_{\mathbf{x}}(f(a_2)), \ldots, j_{\mathbf{x}}(f(a_n))).$$

For an automorphism $f \in Aut(\mathrm{FR}_n)$ which comes from a braid, f_* has a simple interpretation: label the strings at the bottom by the elements $x_1, x_2, \ldots, x_n \in X$, use the usual rules to carry the labels up the braid. Then the labels at the top are $j_{\mathbf{x}}(f(a_1)), j_{\mathbf{x}}(f(a_2)), \ldots, j_{\mathbf{x}}(f(a_n))$. That the two descriptions of the action coincide follows from the method of reading the automorphism from the braid given earlier in the section.

Now let β be a fixed framed braid and let β_* be the corresponding automorphism of FR_n. Consider the closure $C(\beta)$ and let Γ be the fundamental rack of $C(\beta)$. Then a representation of Γ in X is a labelling of $C(\beta)$ by elements of X i.e. an n–tuple $(x_1, x_2, \ldots, x_n) \in X^n$ such that $x_i = j_{\mathbf{x}}(f(a_i))$ for each i, in other words *a fixed point of* f_*. The following result now follows quickly.

1.7.6 Proposition *The set of representations $\Omega = Hom(\Gamma, X)$ is in natural bijection with the fixed point set of f_*.*

Remark This result can be extended to a larger class of racks. Let $f \in Aut(\mathrm{FR}_n)$ and let $\sim$ be the smallest congruence on FR_n such that $x \sim f(x)$ for all $x \in \mathrm{FR}_n$. Define the **almost classical rack** FR_n/f to be FR_n quotiented out by the congruence $\sim$.

The proposition applies to almost classical racks, because a representation ρ of FR_n/f in X determines the n–tuple $\mathbf{x} = (\rho a_1, \rho a_2, \ldots, \rho a_n)$ such that $f_*(\mathbf{x}) = \mathbf{x}$, and conversely.

The usefulness of the result is best demonstrated by examples:

Example 1 *The extended Burau representation*

Let Λ be the ring of Laurent polynomials $\mathbb{Z}[t, t^{-1}]$ in the variable t with integer coefficients and with rack operation $a^b = ta + (1-t)b$. Then the action of $Aut(\mathrm{FR}_n)$ on $X^n = \Lambda^n$ is an extension of the Burau representation and gives a representation $B : Aut(\mathrm{FR}_n) \to GL_n(\Lambda)$.

Now let $f = \lambda(\beta)$ as before, then the fixed point set of f_* is defined by the equation $B(f)\mathbf{x} = \mathbf{x}$ (where $B(f)$ is an $n \times n$ matrix with entries in Λ). It follows that the eigenspace of $B(f)$ with eigenvalue 1 is an invariant of the link. From this we can obtain the usual Alexander polynomial of the link from the codimension one minors.

In fact substantially more invariants of $B(f)$ than just this eigenspace are invariants of the link, since the matrix determines the Alexander module, see example 4 below.

Example 2 *The (t, s)–rack*

We shall describe the extension of the above Burau representation to the ring Λ_s explicitly. The automorphism group of the free rack $Aut(\mathrm{FR}_n)$ is generated by permutation of the generators and **elementary** isomorphisms:

$$\begin{cases} a_i & \mapsto & a_i^{a_j} \\ a_k & \mapsto & a_k \quad k \neq i. \end{cases}$$

The corresponding matrices are permutation matrices and "elementary" matrices obtained from the unit matrix by replacing the i–th diagonal entry by t and the (i, j)–th entry by s. If $i = j$ then replace the i–th diagonal entry

by $s+t$. In explicit examples, for instance the (2×2) matrix examples of 6.2 example 6, the entries are regarded as blocks and the blocks s and t replaced by the appropriate smaller matrices.
Again the eigenspace contains polynomial information and again we can deduce further invariants, see example 5 below.
Example 3 *A sample matrix rack*
Let X be the matrix rack (6.2 example 7) obtained by considering the action of the 1–dimensional affine group $x \mapsto ax+b$ on $\mathbb{R}^1$. Then the rack structure on X is given by

$$(a,b,x)^{(c,d,y)} = (a, cb-ad+d, cx+d).$$

The corresponding representation $Aut(\mathrm{FR}_n) \to X^n = \mathbb{R}^{3n}$ can be quickly written down explicitly as in the last example, and is multilinear.
Several multivariable polynomials can be read from this representation including the multivariable Alexander polynomial (Devine [8]).
Remark In all these examples there is link information contained in the eigenspaces other than that with eigenvalue 1, because they correspond to representations in the appropriate projectivised linear rack.

Functorial invariants
So far we have used the connection with braids to indicate how to read off *representation* invariants of almost classical racks. But *functorial* invariants can also be read. The observation that we need is the following: Suppose we are given a functor

$$\mathcal{F} : \mathbb{RACKS}\mathbf{a}\mathbb{RACKS}$$

and an automorphism $f \in Aut(\mathrm{FR}_n)$. Let $\mathcal{F}(n)$ denote $\mathcal{F}(\mathrm{FR}_n)$, then by functoriality we have the induced automorphism $f_\flat : \mathcal{F}(n) \to \mathcal{F}(n)$. Therefore if $\Gamma = \mathrm{FR}_n/f$ then $\mathcal{F}(\Gamma) = \mathcal{F}(n)/f_\flat$.
Example 4 *The Alexander module*
The relevant functor here was defined in 6.3 example 2. Here $\mathcal{F}(n)$ is the free n–dimensional Λ module and $f_\flat$ is the Burau matrix $B(f)$. It follows that all the invariants of $B(f)$ which are invariants of the module $\mathcal{F}(n)/B(f)$ are invariants of the link. For example we could take all the polynomials in the Smith normal form.
Example 5 *The (t,s)–module*
Once again we can generalise to the (t,s)–rack. Here $B(f)$ is replaced by the (t,s) matrix described explicitly in example 2 above (or in particular matrix representations, by the block matrix obtained by substitution) and again many polynomials invariants of the link can be read. For example we can take one of the variables to be the "variable" and the rest to be "fixed" and then consider the Smith normal form, which gives us in general n multivariable polynomials.
1.7.7 Remark *Racks and R–matrices*
The representation of $Aut(\mathrm{FR}_n)$ as permutations of X^n defined earlier restricts to a representation of B_n. In detail this is defined as follows. There is a bijection $T : X^2 \to X^2$ defined by $T(a,b) := (b, a^b)$ and further bijections $T_i : X^n \to X^n$ for all $n \geq 2$ and $1 \leq i < n$ defined by $T_i := I_{i-1} \times T \times I_{n-i-1}$ where I_i is the identity map on X^i. The representation $B_n \to \mathrm{Perm}(X^n)$ of the braid group is then given by $\sigma_i \to T_i$.
Now suppose that X is in fact a based module over a ring Λ then this representation is given by an "R–matrix" with entries in Λ (cf. Kauffman [17] for details in the case of the Alexander rack).
The other examples of racks which are also modules given earlier in the paper also define R–matrices in this way. This suggests the possibility of using the theory of racks to define 3–manifold invariants in the spirit of Reshetikhin and Turaev [**RT**].

2. Welds and welded links

2.8. Introduction

In this chapter we look at welded braids forming a group called the braid permutation group. This is an extension of the braid group and by analogy we have welded links which extend the class of classical links.

We consider the subgroup of the automorphism group of the free group generated by the braid group and the permutation group. This is proved to be the same as the subgroup of automorphisms of permutation–conjugacy type and is represented by generalised braids (braids in which some crossings are allowed to be "welded"). As a consequence of this representation there is a finite presentation which shows the close connection with both the classical braid and permutation groups. The group is isomorphic to the automorphism group of the free quandle and closely related to the automorphism group of the free rack. These automorphism groups are strongly connected with invariants of classical knots and links in the 3–sphere.

Let F_n denote the free group of rank n with generators $\{x_1, \ldots, x_n\}$ and let $\mathrm{Aut}F_n$ denote its automorphism group. Let $\sigma_i \in \mathrm{Aut}F_n,\ i = 1, 2, \ldots, n-1$ be given by

$$\begin{cases} x_i & \mapsto & x_{i+1} \\ x_{i+1} & \mapsto & x_{i+1}^{-1} x_i x_{i+1} \\ x_j & \mapsto & x_j \qquad j \neq i,\ i+1. \end{cases}$$

and let $\tau_i \in \mathrm{Aut}F_n,\ i = 1, 2, \ldots, n-1$ be given by

$$\begin{cases} x_i & \mapsto & x_{i+1} \\ x_{i+1} & \mapsto & x_i \\ x_j & \mapsto & x_j \qquad j \neq i,\ i+1. \end{cases}$$

As we saw in the first chapter the elements $\sigma_i,\ i = 1, 2, \ldots, n-1$ generate the **braid subgroup** B_n of $\mathrm{Aut}F_n$ which is well known to be isomorphic to the classical braid group of on n strings, and the elements $\tau_i,\ i = 1, 2, \ldots, n-1$ generate the **permutation subgroup** P_n of $\mathrm{Aut}F_n$ which is a copy of the symmetric group S_n of degree n. We shall call the subgroup BP_n of $\mathrm{Aut}F_n$ generated by both sets of elements σ_i and $\tau_i,\ i = 1, 2, \ldots, n-1$ the **braid–permutation group** and this is the subject of this chapter.

This group is interesting for a variety of reasons.

Firstly, the elements of BP_n can be pictured in a way that extends the well-known pictures for elements of the braid group. These are braids in which some of the crossings are "welded" and the welded and unwelded crossings interact in an intuitively simple way.

Secondly, there is a simple characterisation of the automorphisms in BP_n. Let $\pi \in S_n$ be a permutation and w_i, $i = 1, 2, \ldots, n$ be words in F_n. Then the assignation

$$x_i \mapsto w_i^{-1} x_{\pi(i)} w_i$$

determines a homomorphism of F_n to itself which is in fact injective (see 2.7). If it is also surjective (and hence an automorphism) then we call it an automorphism of **permutation–conjugacy type**. The automorphisms of permutation–conjugacy type define a subgroup PC_n of $\mathrm{Aut}F_n$ which is in fact precisely BP_n.

Thirdly, the pictorial definition of BP_n as equivalence classes of welded braids leads to the following finite presentation which includes the standard presentations for both B_n and S_n.

The *generators* are $\sigma_i,\ \tau_i,\ \ i = 1, \ldots, n-1$, as above, and the *relations* are:

$$\begin{cases} \sigma_i \sigma_j & = \sigma_j \sigma_i \quad |i-j| > 1 \\ \sigma_i \sigma_{i+1} \sigma_i & = \sigma_{i+1} \sigma_i \sigma_{i+1} \end{cases}$$

The braid relations

$$\begin{cases} \tau_i^2 & = 1 \\ \tau_i \tau_j & = \tau_j \tau_i \quad |i-j| > 1 \\ \tau_i \tau_{i+1} \tau_i & = \tau_{i+1} \tau_i \tau_{i+1} \end{cases}$$

The permutation group relations

$$\begin{cases} \sigma_i\tau_j & = \tau_j\sigma_i \quad |i-j|>1 \\ \tau_i\tau_{i+1}\sigma_i & = \sigma_{i+1}\tau_i\tau_{i+1} \\ \sigma_i\sigma_{i+1}\tau_i & = \tau_{i+1}\sigma_i\sigma_{i+1} \end{cases}$$

The mixed relations

Fourthly, BP_n is isomorphic to the automorphism group $\mathrm{Aut}FQ_n$ of the free quandle of rank n, and is closely related to the automorphism group $\mathrm{Aut}FR_n$ of the free rack of rank n (the latter group is the wreath product of BP_n with the integers) and these groups are strongly connected with invariants of classical knots and links in the 3–sphere.

Finally the interpretation of elements of BP_n as welded braids suggests a natural relationship with the **singular braid group** SB_n, of Birman et al., which has applications to the theory of Vassiliev invariants, [**Bir2**], [**Kh**]. Indeed these groups have a common subgroup (the braid group) and a common quotient.

The rest of this chapter is organised as follows. In section 2 we prove that the subgroups BP_n and PC_n are identical. This is done by using a variant of the Nielsen theory for free groups. In section 3 we describe welded braid diagrams and the group of welded braids (equivalence classes of diagrams under allowable moves). This is isomorphic to the group presented above and in section 4 we prove that it is isomorphic to BP_n. The proof is diagramatic: using the results of section 2 and 3 we have to show that a welded braid which induces the identity automorphism of F_n can be changed to the trivial braid by allowable moves. Finally in section 5 we interpret the results of previous sections in terms of racks and quandles.

The presentation for BP_n and the proof of this presentation given in this paper are closely related to results for $\mathrm{Aut}FR_n$ given by Krüger [**Kr**]. However his context is rather different from ours and in particular there is no interpretation in terms of diagrams. The results in section 2 of this paper are similar to results in [**Ko**] : our automorphisms of permutation–conjugacy type are called pseudo-conjugations in [**Ko**].

2.9. Automorphisms of permuation–conjugacy type

The groups BP_n and PC_n were defined in section 1. The main result of this section is the equality of these two groups:

2.9.1 Theorem $\quad BP_n = PC_n$

Since the generators σ_i, τ_i of BP_n are automorphisms of permuation–conjugacy type, BP_n is a subgroup of PC_n; thus we have to show that any automorphism of permuation–conjugacy type is a product of the "elementary" automorphisms σ_i, τ_i. It is convenient to use rather different elementary automorphisms and it is also convenient to use exponential notation for conjugacy.

Notation From now until the end of section 4, the notation x^y will mean the product $y^{-1}xy$.

Elementary automorphisms of permuation–conjugacy type

We shall call the automorphisms $p_{i,k}, s_\pi \in PC_n$ given as follows, **elementary** automorphisms.

$$p_{i,k} : \begin{cases} x_i & \mapsto & x_i^{x_k} \\ x_j & \mapsto & x_j \quad j \neq i \end{cases}$$

$$s_\pi : x_i \mapsto x_{\pi(i)} \quad \pi \in S_n.$$

It is easy to see that $p_{i,k}, s_\pi$ are alternative generators for BP_n indeed

$$\sigma_i = p_{i+1,i}\tau_i \quad \text{and} \quad \tau_i = s_{t_i}$$

where t_i is the transposition $(i, i+1)$ and

$$p_{i,k} = \tau_i\tau_{i+1}\ldots\tau_{k-1}\sigma_{k-1}\tau_{k-2}\ldots\tau_i \quad \text{if} \quad i<k$$

$$p_{i,k} = \tau_{i-1}\tau_{i-2}\ldots\tau_{k+1}\sigma_k\tau_k\ldots\tau_{i-1} \quad \text{if} \quad i>k.$$

(Note that products in these formulae are read from left to right.)

Thus to prove the theorem we have to prove that any automorphism in PC_n is a product of the above elementary automorphisms.

We shall use an adaptation of Nielsen theory following the treatment given in Lyndon and Schupp [**LS**] pages 4 to 17. See also Humphries [**Hum2**] where a similar theory is developed and applied to the problem of complex monodromy.

Denote the reduced length of the word $w \in F_n$ by $l(w)$. We shall consider sets $\mathbf{u} = \{u_1, \ldots, u_k\}$ of distinct elements of F_n. We use the notation $\mathbf{u}^{-1} = \{u_1^{-1}, \ldots, u_k^{-1}\}$ for the (disjoint) set of inverses and $\mathbf{u}^{\pm 1} = \mathbf{u} \cup \mathbf{u}^{-1}$.

A set $\mathbf{u}$ is called **Nielsen reduced** if the following conditions hold.

N0 If $u \in \mathbf{u}$ then $u \neq 1$.

N1 If $u,v \in \mathbf{u}^{\pm 1}$ and $uv \neq 1$ then $l(uv) \geq \max\{l(u),l(v)\}$.

N2 If $u,v,w \in \mathbf{u}^{\pm 1}$ and $uv \neq 1$ and $vw \neq 1$, then $l(uvw) > l(u) - l(v) + l(w)$.

2.9.2 Lemma *Let $\mathbf{u} = \{u_1, \ldots, u_k\}$ be Nielsen reduced. If $w = v_1 \cdots v_r$ where $v_i \in \mathbf{u}^{\pm 1}$ and all $v_i v_{i+1} \neq 1$, then $l(w) \geq r$.*

Proof For each $v \in \mathbf{u}^{\pm 1}$ let v_0 be the longest initial segment of v that cancels in any product uv where $u \in \mathbf{u}^{\pm 1}$ and let v_1 be the longest final segment of v that cancels in any product vw where $w \in \mathbf{u}^{\pm 1}$. Note that $v_1 = (v^{-1})_0^{-1}$. Then we can write $v = v_0 m v_1$ where by N2 $l(m) \geq 1$.

So in the product $w = v_1 \cdots v_r$ there is always at least an irreducible subword $m_1 \cdots m_r$ and the result follows. □

2.9.3 Corollary *Let $\mathbf{u} = \{u_1, \ldots, u_k\}$ be Nielsen reduced and suppose in addition that $\mathbf{u}$ generates F_n. Then $\mathbf{u}^{\pm 1} = \mathbf{x}^{\pm 1}$ where $\mathbf{x}$ is the basis $\{x_1, \ldots, x_n\}$ of F_n. (Note in particular that $k = n$.)*

Proof Let the basis element x_i be written as a product $x_i = v_1 \cdots v_r$ where $v_i \in \mathbf{u}^{\pm 1}$ and all $v_i v_{i+1} \neq 1$. Then $1 = l(x_i) \geq r$. So r is forced to be unity and $x_i = v_j$ for some j. □

2.9.4 Definition *Sets of PC-type*

Consider a permutation $\pi \in S_n$ of $\{1, 2, \ldots, n\}$ and n words w_i for $i = 1, 2, \ldots, n$ in the free group F_n. Corresponding to this data is the set of n words $x_{\pi(i)}^{w_i}$ in F_n, where $i = 1, 2, \ldots, n$, obtained by permuting and conjugating the generators. We shall call such a set of words a set of **permutation–conjugacy type**, or **PC–type** for short. We shall also use this name for the set obtained by inverting some of the elements of this set.

Let $\mathbf{u} = \{u_1, u_2, \ldots, u_n\}$ be a set of words of PC–type; we shall consider the following two **double Nielsen transfomations** which preserve PC–type:

T1 replace u_i by $u_j^{-1} u_i u_j$ where $j \neq i$;

T2 replace u_i by $u_j u_i u_j^{-1}$ where $j \neq i$.

2.9.5 Lemma *Let $\mathbf{u} = \{u_1, u_2, \ldots, u_n\}$ be a set of words of PC–type. Suppose that $l(u_i u_j) < l(u_i)$ then either $l(u_j^{-1} u_i u_j) < l(u_i)$ or $l(u_i u_j u_i^{-1}) < l(u_j)$.*

Proof We first observe that we cannot have $l(u_i) = l(u_j)$ because then $l(u_i u_j) < l(u_i)$ implies that at least half of u_i, u_j cancel in the product and the middle letter of u_i cancels with that of u_j. But this is impossible since the words are of PC–type and their middle letters are different generators.

Now assume that $l(u_j) < l(u_i)$. We shall show $l(u_j^{-1} u_i u_j) < l(u_i)$.

Write $u_i = w^{-1} x w$ where x is one of the generators $x_1, x_2, \ldots x_n$, or an inverse. Then since w has length greater than half of u_j, more than half of u_j cancels with w, i.e.

$$w = ya \quad u_j = a^{-1} z \quad \text{where} \quad l(z) < l(a)$$

Therefore

$$l(u_j^{-1} u_i u_j) = l(z^{-1} a a^{-1} y^{-1} x y a a^{-1} z) = l(z^{-1} y^{-1} x y z) < l(a^{-1} y^{-1} x y a) = l(u_i).$$

In the case when $l(u_j) > l(u_i)$ then $l(u_i u_j) < l(u_j)$ and we can show in a similar way that $l(u_i u_j u_i^{-1}) < l(u_j)$. □

2.9.6 Lemma *Let $\mathbf{u} = \{u_1, u_2, \ldots, u_n\}$ be a set of words of PC–type. Then $\mathbf{u}$ can be carried by a sequence of moves T1 and T2 above to a set $\mathbf{v}$ of words of PC–type which is Nielsen reduced.*

Proof The condition N0 is automatically satisfied so assume that $\mathbf{u}$ does not satisfy N1. Then there is a pair $u,v \in \mathbf{u}^{\pm 1}$ such that $l(uv) < l(u)$ and $uv \neq 1$. Then by the last lemma there is a transformation T1 or T2 which reduces $\sum l(u_i)$. Therefore if we apply T1 and T2 until $\sum l(u_i)$ is minimum the condition N1 will hold.

Now consider a triple $u,v,w \in \mathbf{u}^{\pm 1}$ such that $uv \neq 1, vw \neq 1$. Then by N1 we have $l(uv) \geq l(v)$ and $l(vw) \geq l(v)$. It follows that that part of v which cancels in the product uv is no more than half of v. Likewise that part of v which cancels in the product vw is also no more than half of v. So we can write in reduced form $u = ap^{-1}, v = pbq, w = q^{-1}$. Notice that $b \neq 1$ because v is one of a set of PC–words and hence has odd reduced length. So $uvw = abc$ is reduced and $l(uvw) = l(u) - l(v) + l(w) + l(b) > l(u) - l(v) + l(w)$. □

2.9.7 Corollary *A set of words* $\mathbf{u} = \{u_1, u_2, \ldots, u_n\}$ *of PC–type forms the basis of a free subgroup of* F_n *of rank* n *and hence the endomorphism of* F_n *defined by* $x_i \mapsto u_i$ *is injective.*

Proof By the last lemma we can assume without loss that our set of words is Nielsen reduced and the result now follows from lemma 2.2. □

Proof of theorem 2.1

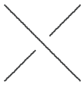
As remarked near the beginning of the section we have to prove that if f is an automorphism of F_n of permutation–conjugacy type then f is a product of elementary automorphisms.

Let $f(x_i) = u_i$ then $\{u_i\}$ is a set of words of PC–type. Now composition of f with the elementary automorphism $p_{i,k}$ for $i \neq k$ realises a double Nielsen transformation on $\{u_i\}$. (If $i = k$ the automorphism has no effect.)

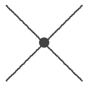
Therefore by lemma 2.6 we may assume that $\{u_i\}$ is Nielsen reduced. But by corollary 2.3 this implies that these words are a permutation of the words x_i, $i = 1, \ldots, n$ and therefore f is the elementary automorphism s_π for suitable π. □

2.10. Welded braids

A **welded braid diagram** on n strings is a set of n monotone arcs from n points on a horizontal line at the top of the diagram down to a similar set of n points at the bottom of the diagram. The arcs are allowed to cross each other either in a "crossing" thus:

 or 

or in a "weld" thus:



An example (on three strings) is illustrated in figure 2.10.27 below:

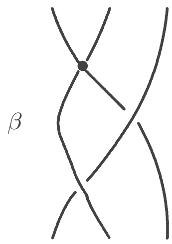


Figure 2.10.27 A welded braid diagram

It is assumed that the crossings and welds all occur on different horizontal levels; thus a welded braid diagram determines a word in the **atomic** diagrams illustrated and labelled in figure 2.10.28

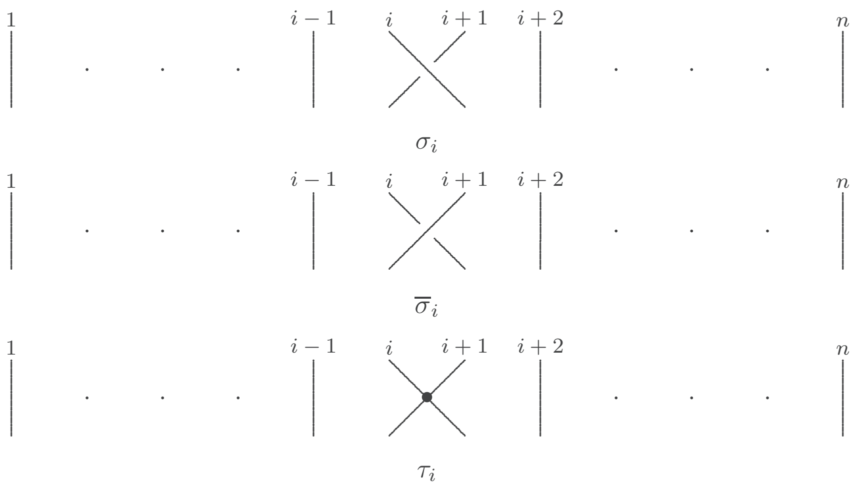


Figure 2.10.28 The three atomic braids

Convention Braids are read *from top to bottom* and words are read *from left to right.*
Thus figure 2.10.27 determines the word $\tau_1\overline{\sigma}_2\sigma_1$ which is read: first τ_1 then $\overline{\sigma}_2$ then σ_1.
We identify braid diagrams which determine the same word in the atomic diagrams (in other words diagrams which differ by a planar isotopy through diagrams).
The set of welded braid diagrams (on n strings) forms a semi-group WD_n with composition given by "stacking": if β_1 and β_2 are diagrams then $\beta_1\beta_2$ is the diagram obtained by placing β_1 above β_2 so that the bottoms of the arcs of β_1 coincide with the tops of the arcs of β_2. There is an identity in WD_n, namely the diagram with no crossings.
The notation for atomic diagrams is intended to be confused with the notation for the generators of BP_n because we now consider the homomorphism $\Phi : WD_n \to \mathrm{Aut}F_n$, with image BP_n, defined by mapping σ_i and τ_i to the automorpisms with the same names and $\overline{\sigma}_i$ to σ_i^{-1}.
Thus a welded braid diagram β determines an automorphism $\Phi(\beta)$ of F_n (of permutation–conjugacy type). There is a convenient way to read $\Phi(\beta)$ from the diagram β which we now describe.

Reading the automorphism from the diagram
Label the strings at the bottom of the diagram by the generators $x_1, x_2, \ldots, x_n$ in order and continue to label the subarcs between crossings moving up the diagram and using the following rules (figure 2.10.29):

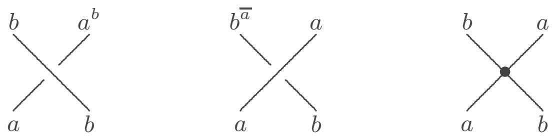
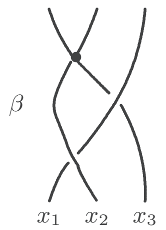


Figure 2.10.29 Rules for labelling subarcs

In figure 2.10.29 we have used the convention that $\overline{x}$ means x^{-1}. This is useful to avoid double exponents. The top of the strings are thus labelled by words $w_1, w_2, \ldots, w_n$ and the automorphism determined by the diagram is given by $x_i \mapsto w_i$ for $i = 1, \ldots, n$. An example of this process in given in figure 2.10.30, where the automorphism corresponding to the diagram in figure 2.10.29 is calculated.

$$\Phi(\beta)\left\{\begin{array}{ccc} x_1 & x_2 & x_3 \\ \downarrow & \downarrow & \downarrow \\ w_1 & x_2 & x_1^{x_2} \end{array}\right.$$



$$w_1 = x_3^{\overline{x_1^{x_2}}} = x_3^{\overline{x}_2\overline{x}_1x_2}$$

Figure 2.10.30

To see that this process gives the correct result it is merely necessary to observe that it is correct for elementary braids (which follows at once from the rules in figure 2.10.29) and that the process gives a homomorphism $WD_n \to \mathrm{Aut}F_n$. To see this consider the effect of stacking the braid β' on top of the braid β. If the i–th point at the bottom of β is labelled x_i and at the top is labelled w_i, then the labels at the top of the combined braid are obtained from those for β' by substituting w_i for x_i. But this is precisely how the composition $\Phi(\beta')\Phi(\beta)$ of the two automorphism of F_n is formed.

Allowable moves on diagrams
We now consider the local changes that can be made to a welded braid diagram β which leave the automorphism $\Phi(\beta)$ unchanged. Four such moves are given in figure 2.10.31

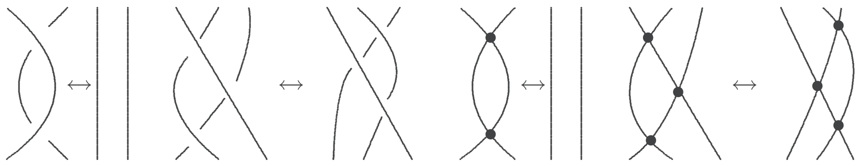


Figure 2.10.31

These are moves involving two or three strings which do not mix welded and unwelded crossings. It is easy to check that these leave the induced automorphism unchanged. The first two moves in figure 2.10.31 are Reidemeister moves from knot theory. If we identify welded braid diagrams that differ by these Reidemeister moves then we can embed the group of braids into the semigroup of welded braid diagrams. If we also identify those braids that differ by the second two moves then we can give welded braid diagrams the structure of a group.

However there are more local changes that do not change the assigned automorphism. In figure 2.10.32 we present two of them.

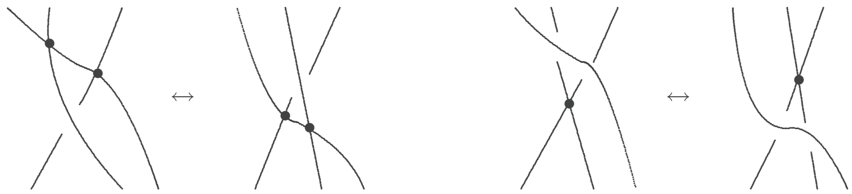


Figure 2.10.32

In addition there are moves of a general class involving four strings which allow non-interferring crossings to be reordered. One such move is illustrated in figure 2.10.33.

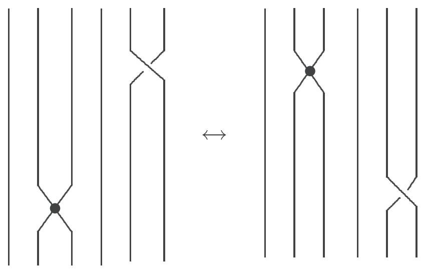


Figure 2.10.33

We call the moves illustrated in figure 2.10.31-33 (and moves similar to figure 2.10.33) **allowable moves** and we define a **welded braid** to be an equivalence class of welded braid diagrams under allowable moves. Welded braids on n strings form a group which we shall denote WB_n with composition given (as in WD_n) by stacking. The inverse of a welded braid β is obtained by reflecting β in a horizontal line. That this is an inverse follows from the first and third moves in figure 2.10.31.

There are other local moves of the same type as those given above which also do not change the induced automorphism. For instance the moves obtained from the moves in figure 2.10.32 by reflection in horizontal or vertical lines. However, as the reader can readily check, these other moves can all be obtained as suitable sequences of the allowable moves given in figures 2.10.31-33 . In figure 2.10.34 below, we show as an example how this can be done for a move of the same type as the first in figure 2.10.32

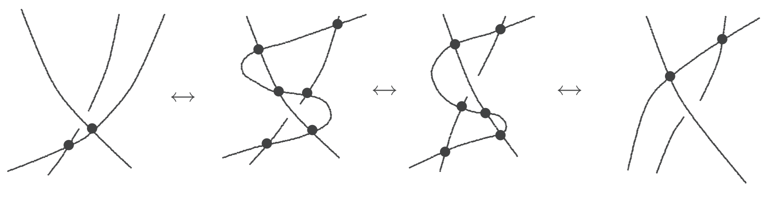


Figure 2.10.34

All the allowable moves and consequent moves are similar to Reidemeister moves, and therefore one could be forgiven for thinking the following. Welded braids can be defined as the 2–dimensional projection of 3–dimensional welded braids and two welded braid diagrams give the same automorphism of F_n if they are projections of equivalent welded braids — where the equivalence among 3–dimensional welded braids is a natural equivalence, for example isotopy in 3–space.

That this is not so can be seen by observing that the local change pictured in figure 2.10.35 **changes** the assigned automorphism.

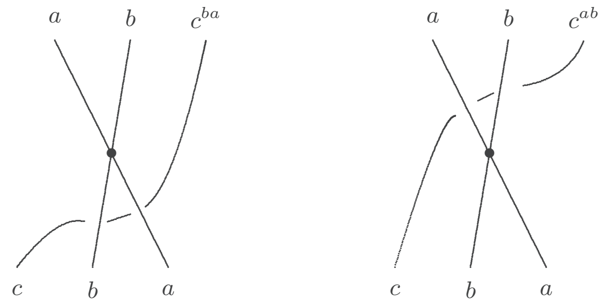


Figure 2.10.35 Two welded braids which do not induce the same element of $\mathrm{Aut}F_n$.

There is however a way to regard welded braids as 3–dimensional objects if desired. Think of the braids as comprising strings embedded in half of 3–space, namely the half *above* the plane of the diagram, and think of the strings as free to move above the plane except at the welds, which are to be regarded as small spots of glue holding the strings down onto the plane. Thus strings are not allowed to move behind welds, as happens in figure 2.10.35

We can now state the main result of the paper:

2.10.2 Main Theorem *Two welded braid diagrams determine the same automorphism of F_n if and only if they can be obtained from each other by a finite sequence of allowable moves.*

The proof of the theorem comprises the next section. We finish this section by remarking that the theorem implies the presentation for BP_n given in section 1. It is not hard to see that the group WB_n of welded braids has this presentation. The elementary braids σ_i and τ_i generate WB_n and the allowable moves each correspond to one of the relations listed in section 1. For example the moves in figure 2.10.32 correspond to the second and third "mixed relations" respectively.

Indeed welded braids can be regarded as a convenient way of illustrating this presentation.

Now we have the homomorphism $\Phi : WD_n \to \mathrm{Aut}F_n$ with image BP_n. Since allowable moves leave the induced automorphism unaltered, this factors, via a homomorphism $WB_n \to \mathrm{Aut}F_n$, with the same image. The theorem implies that this homomorphism is injective and hence WB_n and BP_n are isomorphic. Moreover the generators of WB_n and BP_n (with the same names) correspond, hence BP_n has the presentation given.

æ

2.11. Proof of the main theorem

In order to prove the main theorem (Theorem 2.10.2) it is sufficient to show that any welded braid which induces the identity automorphism of the free group F_n can be reduced to the identity braid by a finite sequence of allowable moves.

Let $C = \{x_i^w \mid w \in F_n\}$ denote the set of conjugates of generators of F_n. If $a = x_i^w$ is in C let $L(a) = l(w)$ where w is chosen to have minimal length $l(w)$. Note that for example $x_i^{x_i w} = x_i^w$.

If a, b belong to C then we write $a \prec b$ if $L(b^a) < L(b)$ and we write $\overline{a} \prec b$ if $L(b^{\overline{a}}) < L(b)$. The following technical lemmas will be useful.

Let a, b, c be arbitary elements of C.

2.11.1 Lemma *If $a \prec b$ $(\overline{a} \prec b)$ and $a = x_i^w$ for some w then $b = x_j^{u\overline{x_i}w}$ $(b = x_j^{ux_iw})$ for some u where $u\overline{x}_iw$ (ux_iw) is reduced.*

Proof If $b = x_j^v$ then $b^a = x_j^{v\overline{w}x_iw}$. If $l(v\overline{w}x_iw) < l(v)$ it follows that v ends in $\overline{x}_iw$. The case when $\overline{a} \prec b$ is similarly proved. □

2.11.2 Lemma *If $a \prec b$ and $b \prec c$ then $a \prec c$ and $a \prec c^b$.*

Proof If $a = x_i^w$ then by the above $b = x_j^{u\overline{x}_i w}$ and $c = x_k^{v\overline{x}_j u\overline{x}_i w}$. So $c^a = x_k^{v\overline{x}_j uw}$, $c^{ba} = x_k^{vuw}$ and $c^b = x_k^{vu\overline{x}_i w}$. Using the fact that $l(\overline{x}_i w) > l(w)$ the result follows.

2.11.3 Lemma *If $a \prec b$ and $\overline{b} \prec c$ then $a \prec c$ and $a \prec c^{\overline{b}}$.*

Proof The proof is similar to 4.2. $\square$

2.11.4 Lemma *If $a \prec b$, $c \prec b$ and $a \neq c$ then $L(a) \neq L(c)$. If $L(a) < L(c)$ then $a \prec c$ and $L(b^{ca}) < L(b)$.*

Proof Let $a = x_i^w$, $c = x_j^u$ in reduced form. From the hypotheses it follows that the exponent of b ends in $\overline{x}_i w$ and also $\overline{x}_j u$. Because $a \neq c$ $L(a) \neq L(c)$. If $L(a) < L(c)$ it follows that $u = s\overline{x}_i w$ and $b = x_k^{v\overline{x}_j s\overline{x}_i w}$, $c = x_j^{s\overline{x}_i w}$. Then $c^a = x_j^{sw}$ so $L(c^a) < L(c)$. We have $b^{ca} = x_k^{vsw}$ which gives the second conclusion. $\square$

2.11.5 Lemma *If $\overline{a} \prec b$, $c \prec b$ and $L(a) < L(c)$ then $\overline{a} \prec c$ and $L(b^{c\overline{a}}) < L(b)$.*

Proof The proof is similar to 4.4. $\square$

2.11.6 Lemma *The following three pairs of conditions are contradictory.*

(1) $a \prec b$ and $b \prec a$.

(2) $a \prec b$ and $\overline{b} \prec a$.

(3) $a \prec b$ and $\overline{a} \prec b$.

Proof For the pairs (1) and (2) lemmas 4.2 and 4.3 respecively imply that $a \prec a$ which is impossible. For (3) lemma 4.1 implies that $b = x_j^{u_1\overline{x}_i w} = x_j^{u_2 x_i w}$. A simple argument using the minimal length representation of the exponent shows that $b = x_i^w = a$ which is again a contradiction. $\square$

We can assign a non negative integer $L(\kappa)$, the **length** to any endomorphism κ of F_n of PC–type by the formula: $L(\kappa) = \sum L(\kappa(x_i))$. Clearly $L(\kappa) = 0$ if and only if κ is a permutation of the generators.

Now let β be a welded braid diagram which represents the identity automorphism. We assume that all (welded and unwelded) crossings have distinct y–coordinates. Let us suppose that the unwelded crossings occur with y–coordinates: $y_1, \ldots, y_{k-1}$. We choose the notation so that

$$0 = y_0 < y_1 < y_2 < \ldots < y_{k-1} < y_k = 1$$

where y_0 is the top level of the braid and y_k is the bottom level of the braid. *Remember the y–coordinate increases as we go down the page.* Any horizontal line whose height is none of these critical values and also does not meet a weld will divide the welded braid β into an upper welded braid β' and a lower welded braid β'' so that $\beta = \beta'\beta''$. If t is the height of the horizontal line define $L(t) = L(\Phi(\beta''))$. Notice that the function $L(t)$ changes only at the critical values. The values of L for $0 < t < y_1$ and for $y_{k-1} < t < 1$ will be 0. If $k > 1$ let us take one of the maximal L–valued intervals, say $[y_s, y_{s+1}]$. The value of s cannot be 0 or $k-1$. We will show that we can change the welded braid by allowable moves so that the value of L is reduced within this interval and is unaltered elsewhere. Let ς_i and ς_j be the unwelded crossings at level y_s and y_{s+1} respectively. Then the braid between levels y_s and y_{s+1} has only welded crossings in other words it is a permutation τ say. We will endevour to reduce the number of welds in τ to a minimum.

Let x be the number of strings which are involved with ς_i and ς_j. The integer x can take the values $2, 3$ or 4. By allowable moves that do not essentially alter the function L we can assume that the remaining $n - x$ strings do not have any welded crossings in the interval $[y_s, y_{s+1}]$. For example figures 6 and 8 show how to move a welded string past an unwelded crossing.

Now we consider the three cases $x = 2, 3, 4$. We will assume that we have used allowable the moves to minimize the number of welds in the interval $[y_s, y_{s+1}]$.

- **Case** $x = 2$

The maximum number of welds occuring now is 1. The possible cases up to a simple symmetry are displayed in figure 2.11.36. The remaining $n - 2$ strings are not shown.

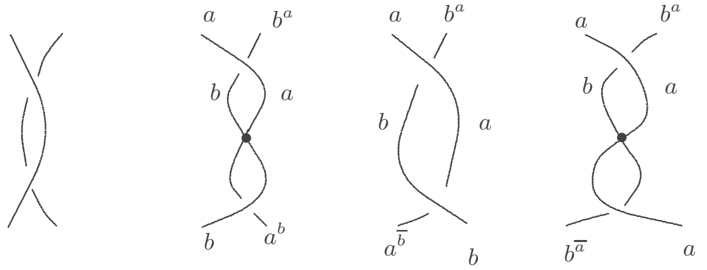


Figure 2.11.36

The first case can easily be simplified. It is not difficult to show with the help of the lemmas above that the other three cases can not occur if the interval defines a maximum of the function L. For example the second picture above gives the conditions $a \prec b$ and $b \prec a$ which contradicts lemma 2.11.6 (1) (see the notation of the picture).

- **Case** $x = 3$

The number of possibilities to be considered can be reduced up to a simple symmetry to the following 9 illustrated in figure 2.11.37 below.

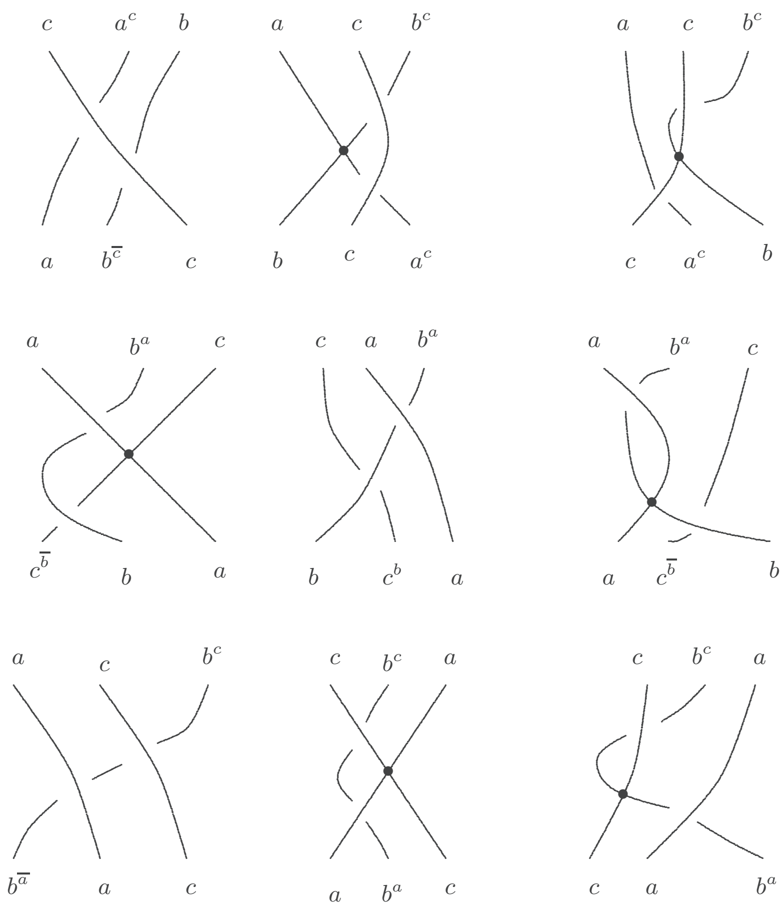


Figure 2.11.37

These can be grouped into three sets of 3 by the following description. The first 3 occur where the overcrossing string at the bottom is an overcrossing string at the top. For the next 3 the overcrossing string at the bottom is an undercrossing string at the top and for the last 3 the undercrossing string at the bottom is the undercrossing string at the top.

The fourth possibility is that the undercrossing string at the bottom is an overcrossing string at the top. However this becomes the middle case by reflection in a horizontal line.

We now change the first six cases by allowable moves so that they look like the following illustrated in figure 2.11.38

below.

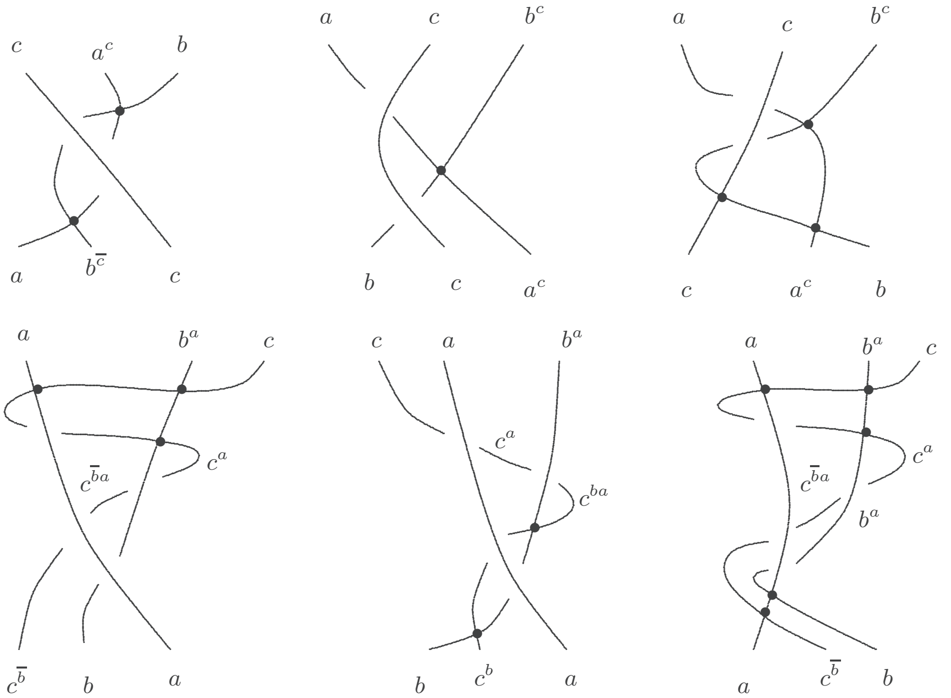


Figure 2.11.38

The relevant change to the last three depends on the length of a and c. If the lengths $L(a)$ and $L(c)$ are equal then Lemma 4.4 gives a contradiction. We give the appropriate changes if $L(a) < L(c)$ in figure 2.11.39 The reader can easily construct the changes in the opposite case.

After these changes have been completed it only remains to check that the value of L has been decreased in the interval. This can be done using the above lemmas and may safely be left to the reader. For convenience we give the notation in the pictures so that they fit the notations of the lemmas.

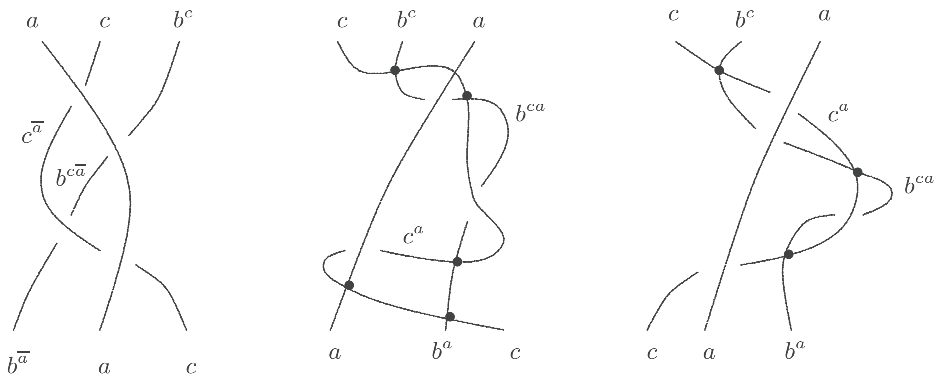


Figure 2.11.39

• **Case** $x = 4$

By minimality there is only one possibility which is illustrated in figure 2.11.40 below.

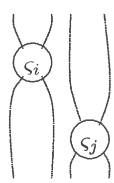


Figure 2.11.40

We change the situation by interchanging the heights of ς_i and ς_j. This has the curious effect of decreasing the value of L in the interval as the reader may verify.

We now repeat this argument until the value of L is constant (which must be zero). This implies that all crossings are welded and therefore the induced automorphism is a permutation. Since this is the identity permutation we can reduce the welded braid to the identity by a sequence of the two right hand moves in figure 2.10.31
This completes the proof of the main theorem. □

2.12. Racks and Quandles

In this section we turn to the theory which originally prompted our interest in the subgroups of $\mathrm{Aut}\,F_n$ studied in the previous sections and show how it is related to the theory of racks. We have throughout written a^b for the conjugate $b^{-1}ab$ in a group. This leads naturally to the binary operation of a rack which can be thought of as generalised conjugation.
Note that *from now on exponential notation is no longer used as a shorthand for conjugation in groups.*
If S is the finite set $\{x_1, \ldots, x_n\}$ then we will use the notation

$$FQ_n = FQ(\{x_1, \ldots, x_n\})$$

for the free quandle on $\{x_1, \ldots, x_n\}$ and

$$FR_n = FR(\{x_1, \ldots, x_n\})$$

for the free rack on $\{x_1, \ldots, x_n\}$.
Let $\partial : FQ_n \to F_n$ from the free quandle to the free group be defined by $\partial(a^w) = w^{-1}aw$. Then $\partial : FQ_n \to F_n$ is injective and any automorphism $\phi : FQ_n \to FQ_n$ induces an automorphism $\phi_\sharp : F_n \to F_n$ such that the diagram below commutes.

$$\begin{tikzcd}
FQ_n \arrow[r, "\phi"] \arrow[d, "\partial"'] & FQ_n \arrow[d, "\partial"] \\
F_n \arrow[r, "\phi_\sharp"] & F_n
\end{tikzcd}$$

The correspondence $\phi \to \phi_\sharp$ embeds $\mathrm{Aut}\,FQ_n$ as a subgroup of $\mathrm{Aut}\,F_n$.
$\mathrm{Aut}\,FQ_n$ is isomorphic to BP_n. We can see this using the previous methods as follows. Let $\phi : FQ_n \to FQ_n$ be an automorphism. Suppose $\phi(x_i) = x_{\pi(i)}^{w_i}$. Then the set $\{x_{\pi(i)}^{w_i}\}$ is of PC–type and generates F_n.
However such sets $\{x_{\pi(i)}^{w_i}\}$ determine PC_n and conversely. So $\mathrm{Aut}\,FQ_n$ is isomorphic to PC_n but we already know that PC_n is BP_n.
Now consider the group $\mathrm{Aut}\,FR_n$. We will see shortly that this is a wreath product of $\mathrm{Aut}\,FQ_n$ with the integers.
Let σ_i and τ_i where $i = 1, 2, \ldots, n-1$ be the elements of $\mathrm{Aut}\,FQ_n$ and $\mathrm{Aut}\,FR_n$ which are defined in an exactly similar way to their namesakes in $\mathrm{Aut}\,F_n$. The **elementary automorphisms** $p_{i,k}$ and s_π of FQ_n and FR_n are also defined in exactly the same way as for F_n, see section 2. The methods of section 2 show that these elementary automorphism generate the two automorphism groups.
Now let ρ_i, where $i = 1, 2, \ldots, n$, be the elements of $\mathrm{Aut}\,FR_n$ defined by

$$\begin{cases} x_i & \mapsto & x_i^{x_i} \\ x_j & \mapsto & x_j \quad j \neq i. \end{cases}$$

In other words ρ_i is $p_{i,i}$. It is easily seen that

$$\rho_i\sigma_i = \sigma_i\rho_{i+1},\ \rho_i\tau_i = \tau_i\rho_{i+1} \text{ and } \rho_i\rho_j = \rho_j\rho_i.$$

Let η: $\mathrm{Aut}FR_n \to \mathrm{Aut}FQ_n$ be the natural map. Then the kernel of η is R the subgroup of $\mathrm{Aut}\,FR_n$ generated by the ρ_i. Clearly R is isomorphic to the lattice group $\mathbb{Z}^n$. Thus $\mathrm{Aut}\,FR_n$ is a semi-direct product of $\mathrm{Aut}\,FQ_n$ with R; moreover the action on R permutes the factors, in other words $\mathrm{Aut}\,FR_n$ is the permutation wreath product of $\mathrm{Aut}\,FQ_n$ with $\mathbb{Z}$.

In terms of presentations we have the following.

2.12.1 Theorem *The group $AutFQ_n$ has a finite presentation of the following form. The generators are σ_i and τ_i where $i = 1, 2, \ldots, n-1$.*
The relations are

$$\begin{cases} \sigma_i\sigma_j & = \sigma_j\sigma_i \quad |i-j| > 1 \\ \sigma_i\sigma_{i+1}\sigma_i & = \sigma_{i+1}\sigma_i\sigma_{i+1} \end{cases}$$

The braid relations

$$\begin{cases} \tau_i^2 & = 1 \\ \tau_i\tau_j & = \tau_j\tau_i \quad |i-j| > 1 \\ \tau_i\tau_{i+1}\tau_i & = \tau_{i+1}\tau_i\tau_{i+1} \end{cases}$$

The permutation group relations

$$\begin{cases} \sigma_i\tau_j & = \tau_j\sigma_i \quad |i-j| > 1 \\ \tau_i\tau_{i+1}\sigma_i & = \sigma_{i+1}\tau_i\tau_{i+1} \\ \sigma_i\sigma_{i+1}\tau_i & = \tau_{i+1}\sigma_i\sigma_{i+1} \end{cases}$$

The mixed relations for $AutFQ_n$

The group $AutFR_n$ has the same generators and relations but in addition the generators ρ_i, where $i = 1, 2, \ldots, n$ and relations

$$\rho_i\rho_j = \rho_j\rho_i.$$

The ρ_i commute

$$\begin{cases} \rho_i\sigma_i & = \sigma_i\rho_{i+1} \\ \rho_i\tau_i & = \tau_i\rho_{i+1} \end{cases}$$

The mixed relations for $AutFR_n$

□

Final remarks *The theory of welded braids developed in this paper has a corresponding theory of welded knots and links. For example gluing the top of a welded braid to the bottom yields a welded link in half 3–space. This can be generalised to yield a theory of welded links in any 3–manifold with boundary. This theory is important because any such link has a fundamental rack: for a glued braid this can be calculated by quotienting the free rack by the automorphism which the welded braid determines. (This automorphism can be computed from the welded braid by the same method as we gave in section 3 for computing the induced automorphism of the free group.) The theory of racks can then be applied to yield invariants for these generalised links.*

3. Biquandles and Virtual Links

In this chapter we define a birack and a biquandle, generalizing the notion of a rack and a quandle. This gives rise to natural invariants of virtual knots and braids. Some of the properties of biracks and biquandles are explained in this paper. Applications to particular virtual braids and links are given.

3.13. Introduction

The biquandle and birack [**FRS, CS**] are algebras associated with a link diagram that are invariant (up to isomorphism) under the generalized Reidemeister moves for virtual knots and links. The basic idea for the birack was first given in [**FRS**]. For historical documents see, http://www.maths.sussex.ac.uk////Staff/RAF/Maths/. Drinfeld in [**Dr**] asked for set theoretic solutions to the Yang Baxter equations and this is what a birack provides. Work related to this can be found in [**ESS, EGS, LYZ, LYZ2, S**]. The definition given here describes this notion geometrically.

The operations in a biquandle are motivated by the formation of labels for the *semi-arcs* of the diagram and the implied invariance under the moves. A semi-arc of a diagram corresponds to an arc running from a classical crossing to the next classical crossing ignoring virtual crossings. We will give the abstract definition of the biquandle before a discussion of these knot theoretic issues hoping that the reader is sufficiently motivated to wade through the consequent algebra.

The operations in a biquandle follow from the labelling of a diagram. In contradistinction to the classical case the overcrossing arc has two labels, one on each side of the crossing. In a presentation of a *biquandle* there is a generator labeling each semi-arc of the diagram. The relations amongst the generators now follow from the two conditions imposed by each crossing.

A diagram of a classical knot or link can be described by the Gauss code, see [**K**]. However not all Gauss codes can be realised as real knots or links by such a diagram. Their realization is dependent on the introduction of *virtual crossings*. These are crossings which are neither above or below in space but just indicate that the journey of the arc intersects the journey of another arc. The labelling on the arc is not altered by this encounter. It turns out that the birack and biquandle is the correct algebraic generalization of the rack and quandle [**FR, J, M**] in dealing with this generalization of classical links. In section 5 we give detailed definitions of, and motivations for virtual links.

3.14. Switches: Definition and Examples

Let X be a set. Denote by $P_n(X)$ the group of permutations of the n-fold cartesian product, X^n. We will shorten $P_1(X)$ to $P(X)$. A **switch** *on X is defined to be an element $S \in P_2(X)$ satisfying the following braid relation in $P_3(X)$,*

$$(S \times id)(id \times S)(S \times id) = (id \times S)(S \times id)(id \times S).$$

Note that this is a more general notion than that considered in [**BF**] where only 2×2 matrices are considered.

Let $S_1 = S \times id$ and $S_2 = id \times S$. Then the relation can now be written

$$S_1 S_2 S_1 = S_2 S_1 S_2.$$

It follows that a switch on X defines a representation of the braid group B_n into the group $P_n(X)$ by sending the standard generator σ_i to $S_i = (id)^{i-1} \times S \times (id)^{n-i-1}$. If in addition the relation $S^2 = id$ holds then we get a representation of the symmetry group on n objects into the group $P_n(X)$. Odesskii in [**O**] calls these twisted transpositions. These are the correct switches to use for what Kauffman calls **flat virtuals** and Turaev calls virtual strings. That is, virtual knots in which the positive and negative real crossings are indistiguishable.

3.15. Examples of switches

0. Let $S = id \times id$ be the identity.

1. Let $T : X^2 \to X^2$ be defined by $T(a,b) = (b,a)$. This switch is called the **twist**.

There is a representation of the virtual braid group using the twist T in an analogous fashion to that of any other switch S. As before, the generator σ_i corresponding to a real crossing is sent to $S_i = (id)^{i-1} \times S \times (id)^{n-i-1}$ and the generator τ_i corresponding to a virtual crossing is sent to $T_i = (id)^{i-1} \times T \times (id)^{n-i-1}$.

If we now put $S_{12} = S_1T_1$, $S_{13} = T_1S_2T_2T_1$ and $S_{23} = S_2T_2$ then the fundamental relation becomes

$$S_{12}S_{13}S_{23} = S_{23}S_{13}S_{12}.$$

This may be more familiar to some readers as the Yang-Baxter equation from quantum group theory, albeit using tensor products.

Here we are using the algebraic (left to right) notation for the composition of operators. So $(AB)(x) = B(A(x))$. This is in order to marry with the rule for labelling arcs in section 6.

The reader may easily check the above equation using the fact that both S and T are switches and in addition

$$T_i^2 = 1 \text{ and } T_1T_2S_1 = S_2T_1T_2$$

2. Let $(a,b) \to a^b$ be a rack action on a set X. Then $S(a,b) = (b, a^b)$ is a switch. This is called the **rack** switch.

3. Let $X = G$ be a group. Then $S(g,h) = (g^{-1}h^{-1}g, h^2g)$ defines the **Wada** switch on G. This example is due to Wada. The inverse is given by $S^{-1}(g,h) = (hg^{-2}, hg^{-1}h^{-1})$ as may be easily verified.

4. Let G be a group with a $\mathbb{Z} \times \mathbb{Z}$ action written $(n,m) \cdot g = g^{\lambda^n \mu^m}$. Then $S(a,b) = (b^\mu, (b^{\lambda\mu})^{-1}a^\lambda b)$ defines the **Silver–William's** switch, [**SW**].

5. Let X be a module over a commutative ring. The linear isomorphism $S : X^2 \to X^2$ given by

$$S(a,b) = (\mu b, \lambda a + (1 - \mu\lambda)b)$$

with λ, μ invertible elements of the ring, defines a switch on X. This switch is the abelianisation of the Silver–William's switch above and is, with the identity, essentially the only interesting linear ones. Full results in this case can be found in [**S**]. We call the above the **Alexander switch** since putting $\mu = 1$ defines the Alexander rack.

6. We can generalise the above to the non-commutative case. Suppose we have a ring R and an R-module M on which we would like to define a switch. Suppose the switch is defined by the 2×2 matrix with entries in R,

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

So $S(a,b) = (Aa + Bb, Ca + Db)$. Then the Yang-Baxter equations imply the seven equations

$$\begin{array}{rl} A = A^2 + BAC & [B,A] = BAD \\ [C,D] = CDA & D = D^2 + CDB \\ [A,C] = DAC & [D,B] = ADB \\ [C,B] = & ADA - DAD \end{array}$$

where $[X,Y]$ denotes the commutator $XY - YX$.

It is proved in [**BF**] that only the first four equations are needed. We also want the matrix to be invertible, and if the switch is to define a birack (see later) then B, C must both be invertible. In fact it then follows that the conditions define an algebra with two generators and one relation. This will be the subject of the next chapter.

The equations have a solution in quaternions, the **Budapest** birack, of the form $A = D = 1 + i,\ B = -j,\ C = j$. This defines a representation of the virtual braid group and *a fortiori* the braid group into the group of invertible matrices with quaternionic entries. This is just one of a family of solutions discovered by Andrew Bartholomew and Peter Croyden using a computer search. The properties of these switches may be found in [**BF**]. In fact, as we will see in the next chapter, all quaternionic solutions have a simple geometric solution.

7. Let $X = \mathbb{Z}^2$ and let $x^+ = \max(0, x)$ and $x^- = \min(0, x)$. The Dynnikov switch is defined by $(a_1, b_1, a_2, b_2) \to (a_1', b_1', a_2', b_2')$ where

$$\begin{array}{ll} a_1' = a_1 - b_1^+ - (b_2^+ + d)^+ & \qquad a_2' = a_2 - b_2^- - (b_1^- - d)^- \\ b_1' = b_2 + d^- & \qquad b_2' = b_1 - d^- \end{array}$$

and $d = a_1 + b_1^- - a_2 - b_2^+$.

The idea behind this switch would take too much time to explain here but see [**D**] for some details.

3.16. The Switch Identities

A switch S on a set X defines 2 binary operations on X by the rule

$$S(a, b) = (b_a, a^b).$$

The operations b_a and a^b are called the **down, up** operations respectively. The notation generalizes the exponential notation for racks. In that case the up operation is the rack operation and the down operation, where $b_a = b$, is trivial. This notation avoids the need for brackets. So for example $a^b{}_c$ means $(a^b)_c$. Another notation for the down and up operations is the one introduced by Kauffman as

$$a_b = a\ \underline{\mathrm{b}}\rvert \qquad \textit{and} \qquad a^b = a\ \overline{\mathrm{b}}\rvert$$

Note that the Kauffman notation keeps all the elements at the same level [**K2**]. For example

$$(a^b)^c = a^{bc} = a\ \overline{\mathrm{b}}\rvert\ \overline{\mathrm{c}}\rvert$$

and

$$a^{(b^c)} = a^{b^c} = a\ \overline{\mathrm{b}\ \ \overline{\mathrm{c}}\rvert}\rvert$$

Given a switch it is not difficult to prove that the Yang–Baxter equations imply the following identities,

i) **Up Interchanges** $\qquad a^{bc} = a^{c_b b^c}$

ii) **Down Interchanges** $\qquad a_{bc} = a_{c^b b_c}$

iii) **The Rule of Five** $\qquad a_b{}^{c_b a} = a^c{}_{b^c a}$

These have been called the Wada identities in [**D**] because of [**W**], but see also [**FRS**].

Since the switch S is a bijection there are two more binary operations called the **up-bar** and **down-bar** operations. These are defined by

$$S^{-1}(a, b) = (b^{\overline{a}}, a_{\overline{b}}).$$

Note that the bar operations need not be the inverses to the unbarred operations.

The four operations (up, down, up-bar and down-bar) satisfy

$$b = b^{\overline{a} a_{\overline{b}}} = b^{a \overline{a_b}} = b_{a \overline{a^b}} = b_{\overline{a} a^{\overline{b}}}$$

for all a, b in X. These identities are called the **partial inverses**.

Remark: The notation Kauffman uses for the **up-bar** and **down-bar** operations, is given by

$$a^{\overline{b}} = a\,\lceil\overline{\mathrm{b}} \qquad \text{and} \qquad a_{\overline{b}} = a\,\lfloor\underline{\mathrm{b}}$$

With this notation the partial inverses identities are written

$$b = b\,\lceil\overline{\mathrm{a}}\ \ \overline{\mathrm{a}\ \ \underline{\mathrm{b}\rfloor}\ \rceil} = b\ \overline{\mathrm{a}\rceil}\lceil\overline{\mathrm{a}\ \ \underline{\mathrm{b}\rfloor}} = \underline{\mathrm{a}\rfloor}\lfloor\underline{\mathrm{a}\ \ \ \overline{\mathrm{b}\rfloor}} = b\,\lfloor\underline{\mathrm{a}}\ \ \underline{\mathrm{a}\lceil\overline{\mathrm{b}}\ \rfloor}$$

3.17. Biracks and Biquandles

The up, down operations define two endomorphisms of X, indexed by X called the **up, down maps** according to the rules $f^a(x) = x^a$ and $f_a(x) = x_a$ for each $a \in X$.

Consider a switch S on X. We say that the pair (X, S) define a **strong birack** *if the following two conditions hold.*

i) The up map $f^a : X \to X$ is a permutation in $P(X)$ for every a in X and we use the notation $x^{a^{-1}} = (f^a)^{-1}(x)$ for the inverse permutation and call it the **inverse up** *operation.*

ii) The down map $f_a : X \to X$ is a permutation in $P(X)$ for every a in X and we use the notation $x_{a^{-1}} = (f_a)^{-1}(x)$ for the inverse permutation and call it the **inverse down** *operation.*

If f^a and f_a are only surjective then the pair (X, S) is called a **weak birack**.

For this paper we will call a strong birack a birack.

Examples of biracks are given by the Twist, Silver-Williams, Alexander, Wada and Budapest switches. The identity and Dinnikov switches do not define biracks as may be easily verified.

Note that in the paper **[K2]** and **[KR]** only the condition that f_a and f^a are surjective is used.

For a birack we can define the following endomorphisms of X^2:

$$S^+_-(a,b) = (b^{a_{b^{-1}}}, a_{b^{-1}}) \qquad S^-_+(a,b) = (b^{a^{-1}}, a_{b^{a^{-1}}})$$

called the **sideways** operations.

The following lemma defines an analogous notation for the barred operations.

3.17.1 lemma *For a birack the functions $x \to x^{\overline{a}}$ and $x \to x_{\overline{a}}$ are bijective. The inverses are written $x \to x^{\overline{a}\ ^{-1}}$ and $x \to x_{\overline{a}\ ^{-1}}$. In terms of the previous notation they are given by the following formulæ*

$$x^{\overline{a}\ ^{-1}} = x^{a_{x^{-1}}} \quad \textit{and} \quad x_{\overline{a}\ ^{-1}} = x_{a^{x^{-1}}}$$

Proof *The formulæ follow from the partial inverses considered above.* □

3.17.2 lemma *Let x, b, c be elements of a birack X. Then the following equalities hold*

$$\begin{array}{llll} i)\ x_{(b_c)^{-1}} = & x_{c^{-1}b^{-1}c^b} & ii)\ x_{(c^b)^{-1}} = & x_{b_c c^{-1} b^{-1}} \\ iii)\ x^{(b^c)^{-1}} = & x^{c^{-1}b^{-1}c_b} & iv)\ x^{(c_b)^{-1}} = & x^{b^c c^{-1} b^{-1}} \end{array}$$

Proof. We will only prove i) because the other cases are proved in a similar way. So, for i) take an element a in the birack, such that $x = a_{bc}$. We can do so because we are working with a birack. In fact $a = x_{c^{-1}b^{-1}}$. Now, applying the down interchange birack identity to a, b and c we have that $a_{bc} = a_{c^b b_c}$. From this and because it is a birack we get that $a_{bc(b_c)^{-1}} = a_{c^b}$. In terms of x this means:

$$x_{(b_c)^{-1}} = a_{c^b} = x_{c^{-1}b^{-1}c^b}$$

. $\square$

Again by analogy define

$$\overline{S}\,{}^{+}_{-}(a,b) = (b_{\overline{a^{\bar{b}^{-1}}}}\ ,\ a^{\bar{b}^{-1}}) = (b_{\overline{a^{b_{a^{-1}}}}}\ ,\ a^{b_{a^{-1}}})$$

and

$$\overline{S}\,{}^{-}_{+}(a,b) = (b_{\bar{a}^{-1}}\ ,\ a^{\overline{b_{\bar{a}^{-1}}}}) = (b_{a^{b^{-1}}}\ ,\ a^{\overline{b_{a^{b^{-1}}}}})$$

3.17.3 Theorem *Suppose that the switch S on X defines a birack. Then the following conditions hold.*

i) $TS^-_+TS^+_- = S^+_-TS^-_+T = id^2$.

ii) $\overline{S}\,{}^{-}_{+}S^+_- = \overline{S}\,{}^{+}_{-}S^-_+ = id^2$.

iii) The sideways maps S^+_-, S^-_+, $\overline{S}\,{}^{-}_{+}$ *and* $\overline{S}\,{}^{+}_{-}$ *are in* $P_2(X)$.

Proof. Number i) is a straight forward calculation from which we conclude that S^+_- and S^-_+ are permutations in $P_2(X)$. Number ii) is also a calculation. Perhaps the best way of seeing this is to show that $\overline{S^-_+} = TS^-_+T$ using the partial inverse identities. Number iii) follows from i) and ii). $\square$

Definition *We say that the birack (X, S) is a* **biquandle** *if the following identities hold,*

$$a^{a^{-1}} = a_{a^{a^{-1}}} \text{ and } a_{a^{-1}} = a^{a_{a^{-1}}}$$

for every a in X. The Alexander, Silver-Williams, Wada and Budapest biracks are all biquandles. A rack birack is a biquandle if and only if the rack involved is a quandle.

3.17.4 Theorem *For birack X the following statements are all equivalent,*

i) X is a biquandle,

ii) Given an element a in X, then there exists a (unique) x in X such that $x = a_x$ and $a = x^a$ and there exists a (unique) y in X such that $y = a^y$ and $a = y_a$.

iii) The four sideways operations $S^{\pm}_{\mp}$ and $\overline{S}^{\pm}_{\mp}$ all leave the diagonal of X^2 invariant.

Proof. Clearly by taking $x = a^{a^{-1}}$ and by taking $y = a_{a^{-1}}$ statements i) and ii) are equivalent. Now consider statement iii). Let Δ denote the diagonal in X^2. Suppose X is a biquandle. Then $S^-_+(a,a) = (a^{a^{-1}}, a_{a^{a^{-1}}}) = (a^{a^{-1}}, a^{a^{-1}})$ and so $S^-_+(\Delta) \subset \Delta$. On the other hand $S^-_+(x_{x^{-1}}, x_{x^{-1}}) = (x,x)$ for any x in X and so $S^-_+(\Delta) = \Delta$. Conversely if this equation holds then so does the first biquandle identity. The second biquandle identity follows by consideration of S^+_-. But if any one of these operations leave the diagonals invariant then so do the other three operations by theorem 4.5 i) and ii). $\square$

Definition *We say that the (weak) birack* (X, S) *is a* **weak biquandle** *if given an element* a *in* X*, then there exists* x *in* X *such that* $x = a_x$ *and* $a = x^a$ *and there exists* y *in* X *such that* $y = a^y$ *and* $a = y_a$*.*

Note that a biquandle becomes weak if x, y above are not unique. This brings us to the following question. Are weak biquandles the same as biquandles?

The following question can also be asked. Does there exist a birack such that $x^{x^{-1}} = x_{x^{x^{-1}}}$ for all x in X but $a_{a^{-1}} \neq a^{a_{a^{-1}}}$ for some a in X or conversely?

These questions were answered in the negative in [**Stan**]

The definition given in [**K2**] and [**KR**] of a biquandle is that of a weak biquandle.

3.18. Virtual Braids and Knots

Recall that classical knot theory can be described in terms of knot and link diagrams. A **diagram** is a 4-regular plane graph (with extra structure at its nodes representing the crossings in the link) represented on a plane and implicitly on a two-dimensional sphere S^2. Recall from chapter 1 that two such diagrams are **equivalent** if there is a sequence of moves of the types indicated in part (A) of Figure 1 (The Reidemeister Moves) taking one diagram to the other. These moves are performed locally on the 4-regular plane graph (with extra structure) that constitutes the link diagram.

Virtual knot theory is an extension of classical knot theory, see [**K**]. In this extension one adds a **virtual crossing** (See Figure 1) that is neither an over-crossing nor an under-crossing. We shall refer to the usual diagrammatic crossings, that is those without circles, as **real** crossings to distinguish them from the virtual crossings. A virtual crossing is represented by two crossing arcs with a small circle placed around the crossing point. The arcs of the graph joining classical crossings are called the **semi-arcs** of the diagram.

In addition to their application as a geometric realization of the combinatorics of a Gauss code, virtual links have physical, topological and homological applications. In particular, virtual links may be taken to represent a particle in space and time which dissappears and reappears. A virtual link may be represented, up to stabilisation, by a link diagram on a surface, [**Ku**]. Finally an element of the second homology of a rack space can be represented by a labelled virtual link, see [**FRS**]. Since the rack spaces form classifying spaces for classical links the study of virtual links may give information about classical knots and links.

The allowed moves on virtual diagrams are a generalization of the Reidemeister moves for classical knot and link diagrams. We show the classical Reidemeister moves as part (A) of Figure 1. These classical moves are part of virtual equivalence where no changes are made to the virtual crossings. Taken by themselves, the virtual crossings behave as diagrammatic permutations. Specifically, we have the flat Reidemeister moves (B) for virtual crossings as shown in Figure 1. In Figure 1 we also illustrate a basic move (C) that interrelates real and virtual crossings. In this move an arc going through a consecutive sequence of two virtual crossings can be moved across a single real crossing. In fact, it is consequence of moves (B) and (C) for virtual crossings that an arc going through any consecutive sequence of virtual crossings can be moved anywhere in the diagram keeping the endpoints fixed and writing the places where the moved arc now crosses the diagram as new virtual crossings. This is shown schematically in Figure 2. We call the move in Figure 2 the **detour**, and note that the detour move is equivalent to having all the moves of type (B) and (C) of Figure 1. This extended move set (Reidemeister moves plus the detour move or the equivalent moves (B) and (C)) constitutes the move set for virtual knots and links.

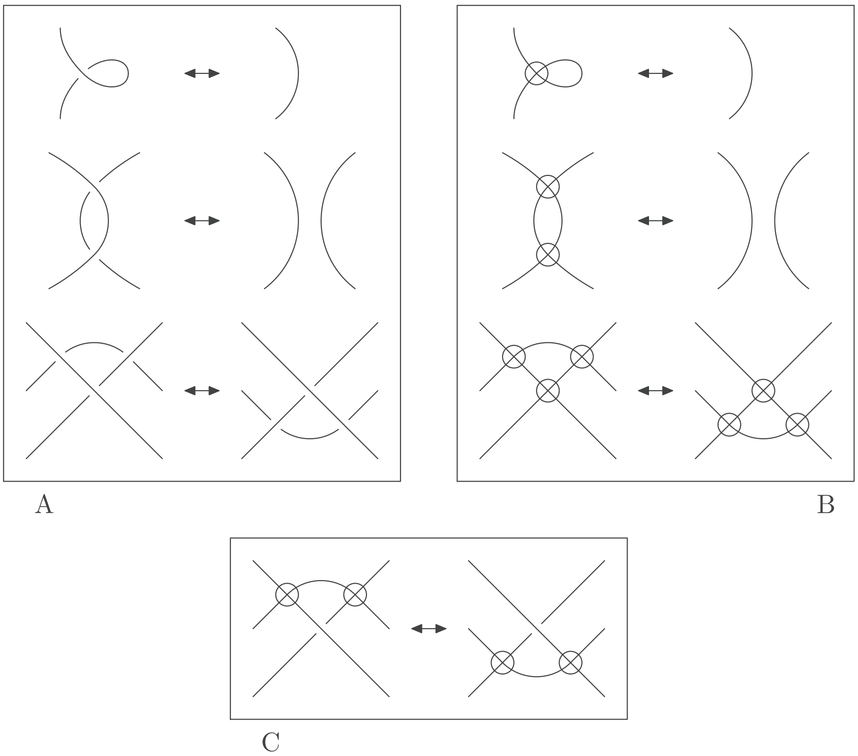


fig1: Generalized Reidemeister Moves for Virtual Knots

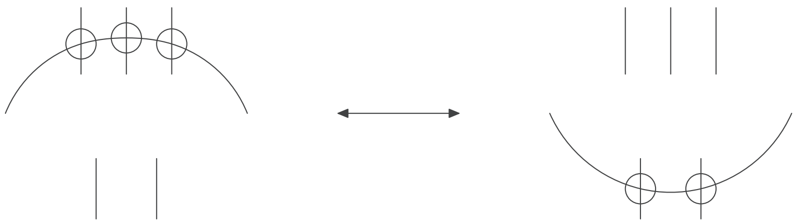


fig2: Schema for the Detour Move

There is a useful topological interpretation for this virtual theory in terms of embeddings of links in thickened surfaces. See [**KK**], [**Ku**]. Regard each virtual crossing as a shorthand for a detour of one of the arcs in the crossing through a 1-handle that has been attached to the 2-sphere of the original diagram. The two choices for the 1-handle detour are homeomorphic to each other (as abstract surfaces with boundary a circle) since there is no a priori difference between the meridian and the longitude of a torus. By interpreting each virtual crossing in this way, we obtain an embedding of a collection of circles into a thickened surface $S_g \times \mathbb{R}$ where g is the number of virtual crossings in the original diagram L, S_g is a compact oriented surface of genus g and $\mathbb{R}$ denotes the real line. Thus to each virtual diagram L we obtain an embedded disjoint union of circles in $S_{g(L)} \times \mathbb{R}$ where $g(L)$ is the number of virtual crossings of L. We say that two such surface embeddings are *stably equivalent* if one can be obtained from another by isotopy in the thickened surfaces, homeomorphisms of the surfaces and the addition or subtraction of empty handles. Then we have the

3.18.1 Theorem *Two virtual link diagrams are equivalent if and only if their correspondent surface embeddings* *are stably equivalent,* **[KK]**, **[Ku]**.

The surface embedding interpretation of virtuals is useful since it converts their equivalence to a topological question. The diagrammatic version of virtuals embodies the stabilization in the detour moves. We shall rely on the diagrammatic approach here.

A **virtual** *braid is defined similarly. The group of virtual braids on n strings is denoted by VB_n. Given a birack structure on X we can use Kamada's presentation of the virtual braid group, VB_n, to get a representation of VB_n to $P_n(X)$ as follows.*

Let generators of VB_n be $\sigma_1,\ \sigma_2,\ \ldots,\sigma_{n-1}$ where σ_i corresponds to the positive real crossing of the i-th and $(i+1)$-th string and $\tau_1,\ \tau_2,\ \ldots,\ \tau_{n-1}$ where τ_i corresponds to the virtual crossing of the i-th and $(i+1)$-th string. The following relations hold:

i) *Braid relations*

$$\sigma_i\sigma_j = \sigma_j\sigma_i, \qquad |i-j|>1$$
$$\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$$

ii) *Permutation group relations*

$${\tau_i}^2 = 1$$
$$\tau_i\tau_j = \tau_j\tau_i, \qquad |i-j|>1$$
$$\tau_i\tau_{i+1}\tau_i = \tau_{i+1}\tau_i\tau_{i+1}$$

iii) *Mixed relations*

$$\sigma_i\tau_j = \tau_j\sigma_i, \qquad |i-j|>1$$
$$\sigma_i\tau_{i+1}\tau_i = \tau_{i+1}\tau_i\sigma_{i+1}$$

See, **[KK]**. At this stage the reader may care to compare these relations with the relations for the braid-permutation group, see **[FRR]**. The braid-permutation group is a quotient of the virtual braid group.

Let S_i denote the element of $P_n(X)$ given by $S_i = id_{i-1}\times S\times id_{n-i-1}$. Then the fundamental relation becomes $S_1S_2S_1 = S_2S_1S_2$. Similarly let $T_i = id_{i-1}\times T\times id_{n-i-1}$ where T is the twist.

3.18.2 Lemma *Let the switch S define a birack structure on X. Then there is a representation $\varphi : VB_n \to P_n(X)$ given by $\varphi(\sigma_i) = S_i,\ \varphi(\tau_i) = T_i$.* □

Example: Let β be the braid on three strings given by $\sigma_2\sigma_1\tau_2\sigma_1^{-1}\sigma_2^{-1}\tau_1$. Then for any rack $S(a,b) = (b,a^b)$ the induced permutation in $P_3(X)$ is the identity. But for the Alexander birack $S(a,b) = (\mu b, \lambda a + (1-\lambda\mu)b)$ the induced permutation is $(a,b,c)\to(a+\kappa, b-\kappa, c)$ where $\kappa = (\lambda^{-1}-\mu)(\mu-1)c$.

3.19. Labelling Diagrams

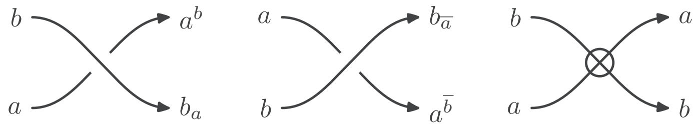
Let the edges of a crossing in a diagram be arranged diagonally and called geographically *NW, SW, NE* and *SE*. Assume that initially the crossing is oriented and the edges oriented towards the crossing from left to right ie west to east. The **input** edges, oriented towards the crossing, are in the west and the edges oriented away from the crossing, the **output** edges, are in the east. Let S be a switch on X and let a and b be labellings from X of the input edges with a labelling SW and b labelling NW. For a positive crossing, a will be the label of the undercrossing input and b the label of the overcrossing input. Then we label the undercrossing output NE by a^b just as in the case of the rack, but the overcrossing output SE is labeled b_a.

We usually read a^b as – the undercrossing line a is acted upon by the overcrossing line b to produce the output a^b. In the same way, we can read b_a as – the overcossing line b is operated on by the undercrossing line a to produce the output b_a.

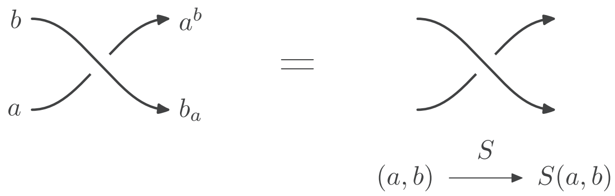
The labels for a negative crossing are similar but with an overline placed on the letters. Thus in the case of the negative crossing, we would write $a^{\overline{b}}$ and $b_{\overline{a}}$, respectively.

For a virtual crossing the labellings carry across the strings.

The following figure shows the labelling for the three kind of crossings.



If we represent a virtual braid by a horizontal diagram oriented from left to right then we can use these labelling rules to describe the representation of the virtual braid group into the group of permutations of X^n where n is the number of strings. Specifically, if $(x_1, x_2, \ldots, x_n)$ is in X^n then label the bottom string on the extreme left with x_1, the next string by x_2 and so on. Now move across to the right labelling as you go using the rules above. If the resulting labelling on the extreme right is $(x'_1, x'_2, \ldots, x'_n)$ then the correspondence $(x_1, x_2, \ldots, x_n) \to (x'_1, x'_2, \ldots, x'_n)$ is the induced permutation. Locally we can think of the output labels from a crossing as being the action of S, $\overline{S}$ or T.



Note that the labelling is extended by the algebraic, left to right ordering and not the functional or right to left ordering.

This permutation is an invariant of the braid and not just of the diagram which represents it because of the switch properties, namely 1) S is a bijection and 2) S satisfies the Yang-Baxter equation. These represent particular types of Reidemeister moves of kind II and III respectively. That is the particular type where the orientations of the strings are all in one direction.

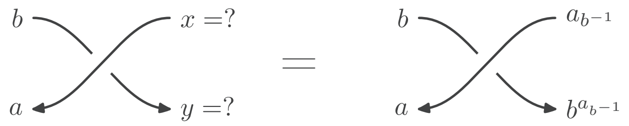
Conjecture A *For a suitable switch (the free switch?) the representation of VB_n is faithful*

The reader is reminded that the equivalent statement is true if the switch is a rack switch and VB_n replaced by B_n, see for example **[FV]**.

If we want the switch to give an invariant of virtual links then we must be prepared to encounter Reidemeister moves in which the string orientations are not all in one direction. In order to extend the labelling in this case the switch must define a birack.

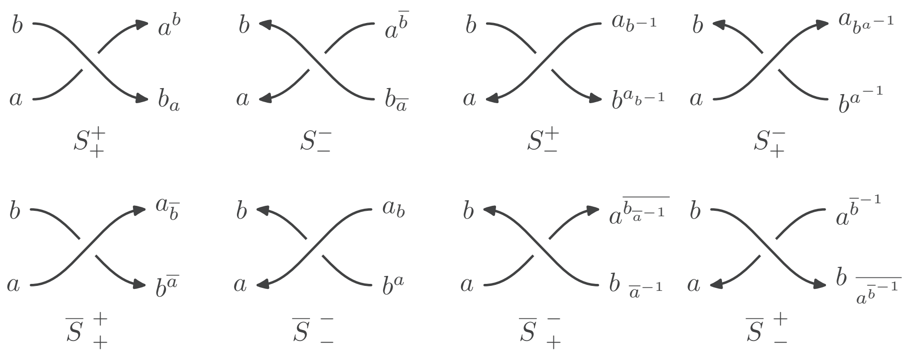
For example if we are given one of the input edges and one of the output edges and the switch defines a birack, then we can compute the values that should be attached to the two edges remaining. Consider the example below where the crossing is positive. The elements a and b are known and x and y are to be determined. Then $x_b = a$ can be solved as $x = a_{b^{-1}}$ and from this it follows that $y = b^x = b^{a_{b^{-1}}}$. This coincides with our previous definition for

$S^+_-(a,b)$.



$$(a,b) \xrightarrow{S^+_-} S^+_-(a,b)$$

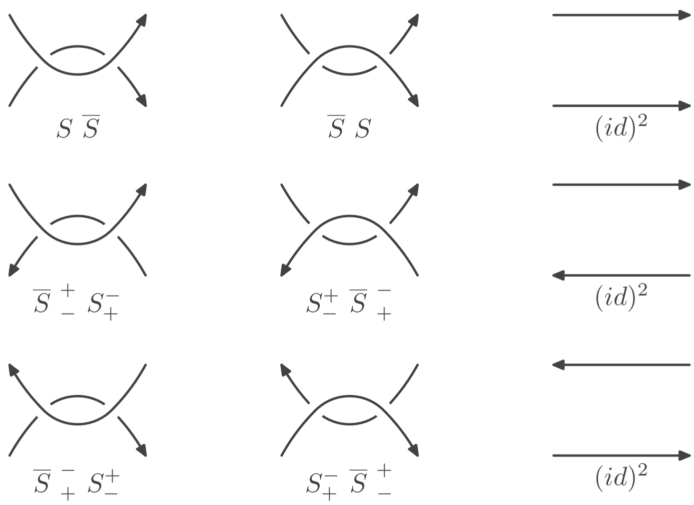
The same can be done for every possibility of orientation and ordering in each crossing and we get all eight posibilities in the following way.



The above diagrams now explain the notation for indices and suffixes. For example S^+_- means a positive entry (arrow from left to right) at the top and a negative entry (arrow from right to left) at the bottom of a positive crossing. The notation $\overline{S}\,^+_-$ means the same but for a negative crossing. Note that there is no difficulty extending the various labellings of any virtual crossing however oriented.

We are now ready to define a **labelling** of the diagram as an attachment of an element of the birack to each semi-arc (or more prosaically a function from the set of semi-arcs to the birack X) so that the conditions given above for each crossing are satisfied.

Using the results proved earlier for biracks we can now show that any labelling of a diagram by a birack can be extended in a unique way to a labelling of a diagram obtained by one Reidemeister move of type II or III. The extension after a type II move is a consequence of the bijective properties of the S operators and is illustrated below.



To extend the labelling after a type III move we need Yang-Baxter type equations for the more general S operators.

Eight cases are possible with different orientations and two are illustrated below.

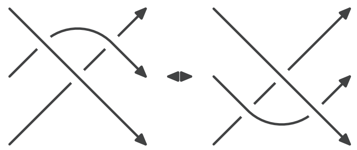


Fig A

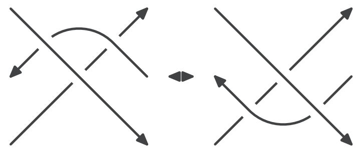
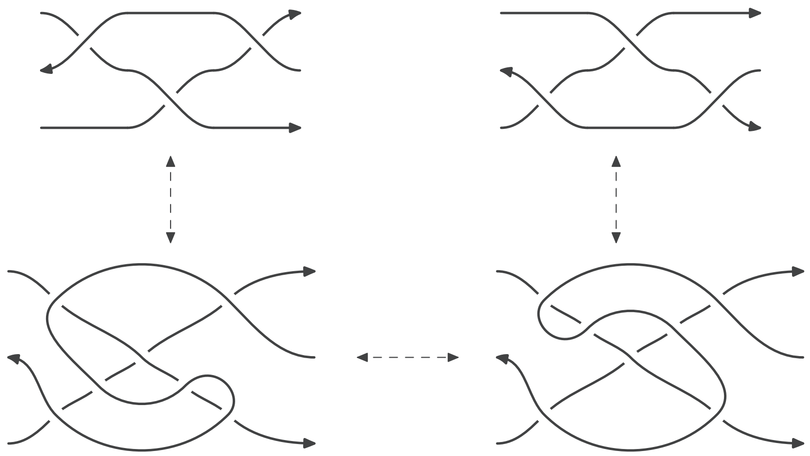


Fig B

Note that any extension of the labelling in Fig A will follow from the original Yang Baxter property. To prove the rest algebraically we would need to verify identities such as $(id\times S_+^-)(S_+^-\times id)(id\times S)=(S\times id)(id\times S_+^-)(S_+^-\times id)$. This is possible but we can use geometry to avoid the algebraic calculations.

The following diagram shows how the Reidemeister move of type II can be utilised so that the move of type III has orientations all in one direction.



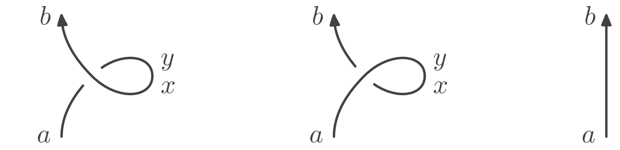
There now remains to consider the case of moves involving the virtual moves. Algebraically we replace the S operators by the twist T, labelling a virtual crossing. For the virtual Reidemeister moves the algebraic analogues are $T(a,b)=(b,a)$, $T^2=\mathrm{id}^2$ and $(id\times T)(T\times id)(id\times T)=(T\times id)(id\times T)(T\times id)$. For the mixed Reidemeister move the algebraic analogue is $(id\times T)(S\times id)(id\times T)=(T\times id)(id\times S)(T\times id)$. All these equations are easily verified.

The next theorem sums up the previous discussion.

3.19.1 Theorem *Let diagrams D_1 and D_2 of virtual links be equivalent by a series of Reidemeister moves of type II and III and those indicated by B and C in figure 1, (a regular homotopy). Then any labelling of D_1 by elements of a birack X defines a unique labelling of D_2 using these moves.* □

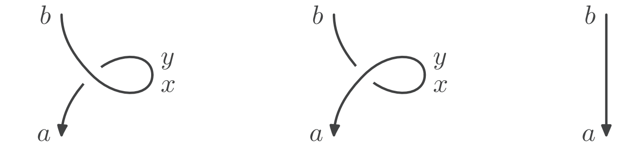
It follows that if $\mathrm{lab}_X(D)$ denotes the set of labellings by X of the virtual diagram D then the moves above induce a bijection between $\mathrm{lab}_X(D_1)$ and $\mathrm{lab}_X(D_2)$ provided X is a birack and D_1 and D_2 are equivalent under these moves.

In order for a labelling to extend after a Reidemeister move of type I the labelling set will have to be a biquandle. For the Reidemeister move of type I the various cases are illustrated below with labellings which force the labelling set to be a biquandle.



$\overline{S}\,{}^{-}_{+} = (S^{+}_{-})^{-1}$ $\qquad S^{-}_{+} \qquad$ id

and



$\overline{S}\,{}^{+}_{-} = (S^{-}_{+})^{-1}$ $\qquad S^{+}_{-} \qquad$ id

So in all the diagrams $a = b$ if and only if $x = y$. This is equivalent to the operators S^{+}_{-} and S^{-}_{+} preserving the diagonal and this is a defining condition for a biquandle.

These results are summed up in the following theorem.

3.19.2 Theorem *With the notation above let the diagrams D_1 and D_2 represent the same virtual link (so they are related by a series of moves indicated by A, B and C in figure 1). Let D_1 be labelled by elements of a biquandle X. Then there is a uniquely defined labelling of D_2 induced by the series of these moves. In particular $\mathrm{lab}_X(D_1)$ and $\mathrm{lab}_X(D_2)$ are in bijective correspondence.* □

A similar theorem can be given for labellings by a weak biquandle. However the labelling defined by a sequence of Reidemeister moves only exists. It may not be unique.

(Note: the labellings by a rack of a classical link can be given a canonical topological definition, see [**FR**]. We do not know of such a definition for the labellings of a virtual link.)

From the above we see that the set of labellings of a virtual link by elements of a biquandle is an invariant. Of particular interest is the set of labellings by the Alexander biquandle. If the semi-arcs are labelled by $(2n)$ generators then the Alexander biquandle conditions on the n vertices define $2n$ relations giving a square presentation of a $\mathbb{Z}[\lambda, \lambda^{-1}, \mu, \mu^{-1}]$ module. (Here as elsewhere we assume that the diagram has no free floating circles.) As in the classical case the minors of the presentation matrix define a sequence of determinants $\Delta_0, \Delta_1, \cdots$. But contrary to the classical case the top determinant is not necessarily zero. We define the zeroth **Alexander polynomial** of a virtual link L to be $\Delta = \Delta_0$ (modulo units), see [**BF, SAW**]. For a classical link $\Delta = 0$.

We would like to define the *fundamental biquandle* of a virtual link by proceeding as in the classical case with the fundamental rack. That is, let the arcs of a representative diagram be the set of generators and the crossings as a set of relations in some sort of presentation.

But that is to suppose that we have a good idea of the notion of a *free biquandle*. Let us use this section as a forum for discussion.

It is unusual that an algebra would have axioms asserting the existence of fixed points with respect to operations involving its own elements. We plan to take up the study of this aspect of biquandles in a separate publication. For now it is worth remarking that a slight change in the axiomatic structure allows an easy definition of the free biquandle. The idea is this: Suppose that one has an axiom that states the existence of an x such that $x = a \,\underline{\ \mathrm{x}\ }\rfloor$

($x = a_x$) for each a. Then we change the statement of the axiom by adding a new operation (unary in this case) to the algebra, call it $\mathrm{Fix}(a)$ such that $\mathrm{Fix}(a) = a_{\mathrm{Fix(a)}}$. Existence of the fixed point follows from this property of the new operation, and we can describe the free biquandle on a set by taking all finite biquandle expressions in the elememts of the set, modulo these revised axioms for the biquandle.

Let a be an element of a weak biquandle S. Then there exists an x in S satisfying the equation $x = a_x$ and $a = x^a$. Note that if S is a (strong) biquandle the x is uniquely defined by $x = a^{a^{-1}}$. In this condition the second equation is a normal relation involving the elements x and a, but the first equation is an existence statement about x satisfying the equation $x = a_x$. Consider the simplest instance of this situation. Let $BQ = (a|)$ denote the "free biquandle" generated by the single element a. In the usual case of universal algebras the "free object" on a single generator is constructed by taking all finite algebraic expressions involving the formalism of the algebra and the generating element, subject to the natural equivalence relations that ensue for this algebra (this depends on the axioms which usually are expressed as relations, not as existence statements). But here we are asked to know that there is an x satisfying the equation above. One solution for x is the infinitary expression

$$x = a_{a_{a_{a\dots}}}$$

since this formally satisfies the equation $x = a_x$. It is not clear to us how to add infinitary expressions in a controlled way to obtain an adequate definition of a free biquandle.

If we label the semi-arcs of a diagram D by the generators of a free biquandle and impose the relations forced by the crossings then we have a presentation of the **fundamental biquandle**, $B(L)$ of the virtual link L defined by D. The set of labellings of L by X is then

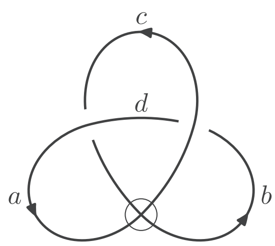
$$\mathrm{lab}_X(L) = \hom(B(L), X)$$

We do not have a canonical definition of $B(L)$ and in general it is a fairly mysterious object, but we hope to return to the associated algebra and considerations of free biquandles in a later paper.

Conjecture B *The fundamental biquandle $B(L)$ is a complete invariant of virtual links up to mirror image.* The mirror image of a (virtual) link is obtained from the diagram by interchanging positive and negative real crossings and reversing the orientation of the knot. In opposition to classical knots and links there is a further symmetry defined by a reflection in a line of the plane. The motive for this conjecture is to mimic the result in the real case and because our calculations so far bare this out.

3.20. Examples and Calculations

1.-Consider the **virtual trefoil** as shown in the figure. If we label as indicated then the fundamental biquandle has a presentation with 3 generators a, b, c and relations $c = a_b,\ a = d_c,\ b = c^d,\ d = b^a$.



We now consider labellings from the Alexander birack. If we eliminate $c = a_b = \mu a$ and $d = b^a = \lambda b + (1 - \lambda\mu)a$ we arrive at the following equations.

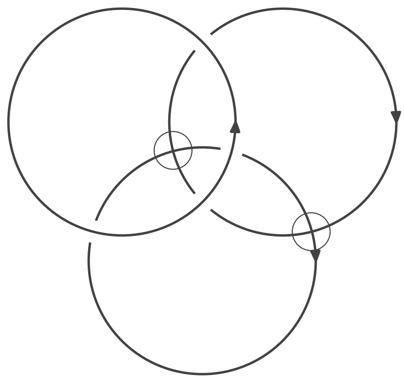
$$(\mu - \lambda\mu^2 - 1)a + \lambda\mu b = 0$$
$$(\lambda^2\mu^2 - \lambda\mu + 1)a + (\lambda - \lambda^2\mu - 1)b = 0$$

The determinant of these equations is

$$\Delta = (\lambda - 1)(\lambda\mu - 1)(\mu - 1).$$

The fundamental quandle (and hence group) is trivial.

2.-Another virtual link with trivial fundamental rack is the closure of the braid given at the end of section 5. The resulting three component link is pictured below.



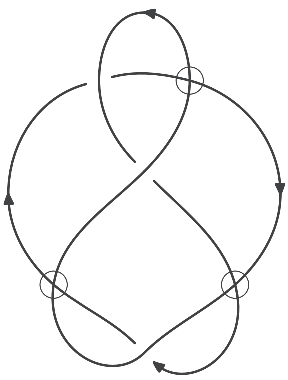
In order to make the equations on the labellings work, we need $\kappa = 0$ so

$$\Delta = (\lambda\mu - 1)(\mu - 1).$$

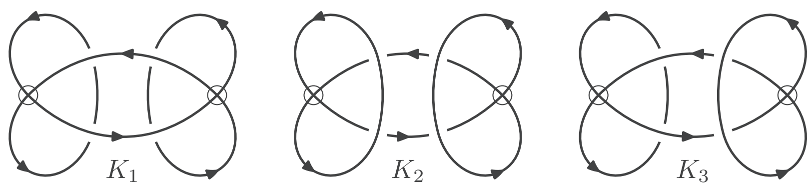
One could also argue that the link is non-trivial since the component without virtual crossings encircles one of the virtual crossings of the other two.

3.-The following virtual knot is interesting in having a trivial Jones-polynomial as well as a trivial fundamental rack. In this case

$$\Delta = (\mu - 1)(\lambda^2(\mu + 1) - \lambda(\mu + 1)(\mu^{-1} + 1) - \mu).$$



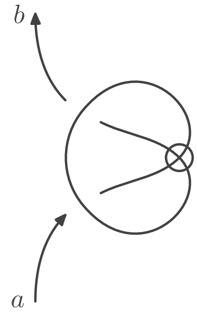
4.-The Kishino knots K_1, K_2 and K_3 are illustrated below. All are ways of forming the connected sum of two unknots. K_1 and K_2 are mirror images and K_3 is amphichæral. Both have trivial racks and Jones polynomial. The Alexander polynomial Δ is zero in all three cases. On the other hand for K_1, Δ_1 is $1 + \mu - \lambda\mu$ and for K_2, Δ_1 is $1 + \lambda - \lambda\mu$. Since these are neither units nor associates in the ring, K_1, K_2 are non-trivial and non amphichæral. The Alexander polynomial Δ_1 of K_3 is 1.



Since the Alexander invariants do not show that K_3 is non-trivial new methods are needed. This is done in the next chapter using the Budapest biquandle. This argument may be paraphrased as follows. The semi-arcs are labelled by quaternion variables and the crossing changes by the Budapest switch. This defines a module which is shown to be non-trivial by considering the ideal generated by codimension 2 Study determinants.

The non-triviality has also been detected by the three string parallel, see **[KS]**.

The knot K_3 is made up of tangles such as the one illustrated in the tangle figure. We now show that this is non-trivial by showing that the output, b, may be different from the input, a. All labels are quaternions and the Budapest switch is used. An easy calculation shows that $3(a-b)=0$ so if we consider the quotient $\mathbb{Z}_3[i,j,k]$ then indeed a need not equal b.



Tangle Figure

4. QUATERNION ALGEBRAS and INVARIANTS of VIRTUAL KNOTS and LINKS

4.21. Introduction

In this chapter we show how invariants of virtual knots can be found using non-commuting rings. In particular generalised quaternions and 2×2 matrices can be used. In the previous section commuting rings were used to find Alexander type polynomials. The ideas can be transferred by careful consideration of determinants in the non-commuting case.

It can be shown, see below, that if we consider the seven equations considered earlier then all we need to do is to find solutions of the equation

$$A^{-1}B^{-1}AB - B^{-1}AB = BA^{-1}B^{-1}A - A$$

in which A, B are elements of a suitable algebra say the quaternions or 2×2 matrices. This later case can be divided into two cases, the elliptic case considered here and the hyperbolic case see [].

Note that the above equation can also be written

$$[B,(A-1)(A,B)] = 0$$

where $[X,Y]=XY-YX$ is one type of commutator and $(X,Y)=X^{-1}Y^{-1}XY$ is another.

Consider the algebra with the following presentation

$$\mathcal{F} = \{A, B \mid A^{-1}B^{-1}AB - B^{-1}AB = BA^{-1}B^{-1}A - A\}.$$

We will call this the **fundamental** algebra and the single relation will be called the fundamental relation or equation. The aim is to find all representations of this algebra into quaternion algebras.

Let K be a virtual knot or link. We show that if A, B are two elements of a ring, R, satisfying the fundamental relation, then there is a left R-module $\mathcal{M}$, dependant on A, B which is an invariant of K. From this module more tractible invariants such as polynomials can be defined.

As an application, we define two 4-variable polynomials of virtual knots and links. In addition, we give a proof that invariants defined in this manner extend the Alexander polynomial for classical knots and links.

4.22. The fundamental equation

Given a set X let S be an endomorphism of X^2. In the previous chapter or the paper **[FJK]**, such an S is called a **switch** if

1 S is invertible and

2 the set theoretic Yang-Baxter equation $(S \times id)(id \times S)(S \times id) = (id \times S)(S \times id)(id \times S)$ is satisfied.

Given a switch S there is a representation of the braid group B_n into the group of permutations of X^n defined by

$$\sigma_i \longrightarrow (id)^{i-1} \times S \times (id)^{n-i-1}$$

where σ_i are the standard generators. Denote this representation by $\rho = \rho(S, n)$.

In this paper we will only be interested in the following linear switches. So let R be an associative but not necessarily commutative ring and let X be a left R-module. Suppose

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

where the matrix entries A, B, C, D are elements of R. The 3×3 matrices of the Yang-Baxter equation are

$$S \times id = \begin{pmatrix} A & B & 0 \\ C & D & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad id \times S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & A & B \\ 0 & C & D \end{pmatrix}$$

The representation ρ is now into $n \times n$ matrices with entries from R.

Let us consider methods to find such switches S. It is not difficult to see that the following 7 equations considered in the last chapter are necessary and sufficient conditions for an invertible S to be a switch,

$$\begin{array}{rl} 1: A = A^2 + BAC & 2: [B, A] = BAD \\ 3: [C, D] = CDA & 4: D = D^2 + CDB \\ 5: [A, C] = DAC & 6: [D, B] = ADB \\ 7: [C, B] = & ADA - DAD \end{array}$$

where $[X, Y] = XY - YX$.

If $A, A-1, B$ are invertible, A, B do not commute and satisfy the fundamental equation

$$A^{-1}B^{-1}AB - BA^{-1}B^{-1}A = B^{-1}AB - A$$

and

$$C = A^{-1}B^{-1}A(1 - A), \ D = 1 - A^{-1}B^{-1}AB$$

then S will be a switch. We will call this the **non-commuting** switch. A special case of this is the matrix with quaternion entries

$$S = \begin{pmatrix} 1+i & j \\ -j & 1+i \end{pmatrix}$$

called the **Budapest** switch.

If $S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is a switch then so is $S(t) = \begin{pmatrix} A & tB \\ t^{-1}C & D \end{pmatrix}$ where t is a commuting variable. We say that $S(t)$ is S **augmented** by t.

In **[BuF]** and **[BF]** the following results can be found.

4.22.1 Theorem *Suppose R is a division ring. Then any switch of the above form is either the Alexander or the non-commuting switch.* □

The representation of the braid group induced by any non-commuting switch looks complicated but is in fact equivalent to the Burau representation. This has been pointed out previously by Dehornoy, see **[De]**. The following lemma gives an explicit proof.

4.22.2 Lemma *Let $S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ be a non-commuting switch and let $S' = \begin{pmatrix} 0 & 1 \\ Q & 1-Q \end{pmatrix}$ be the Burau switch where $Q = (1 - A)(1 - D)$. Let M be the $n \times n$ matrix*

$$M = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ A & B & 0 & \cdots & 0 \\ A & BA & B^2 & \cdots & 0 \\ & \vdots & & \vdots & 0 \\ A & BA & B^2A & \cdots & B^{n-1} \end{pmatrix}$$

In words: the rows of M, after the first, start with A and then the previous row multiplied on the left by B. Clearly M is invertible. Then $\rho(S, n) = M^{-1}\rho(S', n)M$.

Proof A calculation shows that $M\rho(S, n) = \rho(S', n)M$. In this calculation the fundamental relation is used. For example Q commutes with B. So to prove that

$$B^iA^2 + B^{i+1}C = QB^i + (1 - Q)B^iA$$

we need to show that

$$A^2 + BC = Q + (1 - Q)A$$

which follows from the fundamental relation. □

However, if we extend the representation to the **virtual braid group**, defined below, then we get a representation which is **not** equivalent to the Burau.

The virtual braid group, VB_n **[KK]**, has generators $\sigma_i, \ i = 1, \ldots, n - 1$ and braid group relations

i)

$$\sigma_i\sigma_j = \sigma_j\sigma_i, \qquad |i - j| > 1$$

$$\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$$

In addition there are generators $\tau_i,\ i = 1, \ldots, n-1$ and permutation group relations

ii)
$$\begin{aligned} {\tau_i}^2 &= 1 \\ \tau_i\tau_j &= \tau_j\tau_i, \qquad |i-j| > 1 \\ \tau_i\tau_{i+1}\tau_i &= \tau_{i+1}\tau_i\tau_{i+1} \end{aligned}$$

together with mixed relations

iii)
$$\begin{aligned} \sigma_i\tau_j &= \tau_j\sigma_i, \qquad |i-j| > 1 \\ \sigma_i\tau_{i+1}\tau_i &= \tau_{i+1}\tau_i\sigma_{i+1} \end{aligned}$$

We can extend the representation $\rho(S, n)$ by sending the generator τ_i to $(id)^{i-1} \times T \times (id)^{n-i-1}$ where $T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, that is, Burau with unit variable.

Consider now the element $\beta = \sigma_2\sigma_1\tau_2\sigma_1^{-1}\sigma_2^{-1}\tau_1$ in VB_3. If $S = \begin{pmatrix} 0 & B \\ C & 1-BC \end{pmatrix}$, the Alexander switch, then

$$\rho(\beta) = \begin{pmatrix} 1 & 0 & (C^{-1}-B)(B-1) \\ 0 & 1 & (C^{-1}-B)(1-B) \\ 0 & 0 & 1 \end{pmatrix}$$

So if $B \neq 1$ this representation can not be equivalent to the Burau representation which has $B = 1$.

We can now ask the following question: *Let $K \subset B_n$ be the kernel of the Burau representation. Let $\overline{K}$ denote the normal closure of K in VB_n. If β is a virtual braid and $\rho(\beta) = 1$ for all switches S, is it true that β lies in $\overline{K}$?*

4.23. Quaternion Algebras

It is clear from the previous section that it is important to find solutions to the fundamental equation. The main result in the next section is a sufficient condition for two generalised quaternions to satisfy the fundamental equation. Except for 2×2 matrices, this condition is also necessary.

In this section we describe the necessary algebra. The results which are already in the literature are mainly presented without proof. For more details see [**L**] or [**MR**].

Let F be a field of characteristic not equal to 2. Pick two non-zero elements $\lambda,\ \mu$ in F. Let $(\frac{\lambda,\ \mu}{F})$ denote the algebra of dimension 4 over F with basis $\{1, i, j, k\}$ and relations $i^2 = \lambda,\ j^2 = \mu,\ ij = -ji = k$. The multiplication table is given by

$$\begin{array}{c c} & \begin{array}{ccc} i & j & k \end{array} \\ \begin{array}{c} i \\ j \\ k \end{array} & \begin{pmatrix} \lambda & k & \lambda j \\ -k & \mu & -\mu i \\ -\lambda j & \mu i & -\lambda\mu \end{pmatrix} \end{array}.$$

Throughout the paper a general quaternion algebra will be denoted by $\mathcal{Q}$. Elements of $\mathcal{Q}$ are called (generalized) quaternions. The field F is called the **underlying** field and the elements $\lambda\ \mu$ the **parameters** of the algebra. We will denote quaternions by capital roman letters such as $A, B, \ldots$ and (if pure) by bold face lower case, $\mathbf{a}, \mathbf{b}, \ldots$. Field elements, (scalars) will be denoted by lower case roman letters such as $a, b, \ldots$ and lower case greek letters such as $\alpha, \beta, \ldots$.

The classical quaternions are $(\frac{-1,\ -1}{\mathbb{R}})$. The algebra of 2×2 matrices with entries in F is $M_2(F) = (\frac{-1,\ 1}{F})$.

Conjugation, Norm and Trace

Let $A = a_0 + a_1 i + a_2 j + a_3 k$ be a quaternion where $a_0, a_1, a_2, a_3 \in F$. The coordinate a_0 is called the **scalar** part of A and the 3-vector $\mathbf{a} = \mathbf{a_1 i} + \mathbf{a_2 j} + \mathbf{a_3 k}$ is called the **pure** part of A. Evidently $A = a_0 + \mathbf{a}$ is the sum of its scalar and pure parts and is pure if its scalar part is zero and is a scalar if its pure part is zero.

The **conjugate** of A is $\overline{A} = a_0 - \mathbf{a}$, the **norm** of A is $N(A) = A\overline{A}$ and the **trace** of A is $\mathrm{tr}(A) = A + \overline{A}$.

Conjugation is an anti-isomorphism of order 2. That is it satisfies

$$\overline{A+B} = \overline{A} + \overline{B}, \quad \overline{AB} = \overline{B}\,\overline{A}, \quad \overline{aA} = a\overline{A}, \quad \overline{\overline{A}} = A.$$

Also $\overline{A} = A$ if and only if A is a scalar and $\overline{A} = -A$ if and only if A is pure.

The norm is a scalar satisfying $N(AB) = N(A)N(B)$. We will denote the set of values of the norm function by $\mathcal{N}$. It is a multiplicatively closed subset of F and $\mathcal{N}^* = \mathcal{N} - \{0\}$ is a multiplicative subgroup of F^*. An element A has an inverse if and only if $N(A) \neq 0$ in which case $A^{-1} = N(A)^{-1}\overline{A}$.

The trace of a quaternion is twice its scalar part.

Multiplying Quaternions

Let A, B be two quaternions.There is a bilinear form given by

$$A \cdot B = \frac{1}{2}(A\overline{B} + B\overline{A}) = \frac{1}{2}(\overline{A}B + \overline{B}A) = \frac{1}{2}\mathrm{tr}(A\overline{B}).$$

In terms of coordinates this is

$$A \cdot B = a_0 b_0 - \lambda a_1 b_1 - \mu a_2 b_2 + \lambda\mu a_3 b_3.$$

Since λ and μ are non-zero this is a non-degenerate form. The corresponding quadratic form is

$$N(A) = a_0^2 - \lambda a_1^2 - \mu a_2^2 + \lambda\mu a_3^2.$$

Let $\mathbf{a}, \mathbf{b}$ be pure quaternions. Then

$$\mathbf{ab} = -\mathbf{a} \cdot \mathbf{b} + \mathbf{a} \times \mathbf{b}$$

where

$$\mathbf{a} \cdot \mathbf{b} = -\lambda \mathbf{a_1 b_1} - \mu \mathbf{a_2 b_2} + \lambda\mu \mathbf{a_3 b_3}$$

is the restriction of the bilinear form to the pure quaternions and $\mathbf{a} \times \mathbf{b}$ is the **cross product** defined symbolically by

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} -\mu i & -\lambda j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

The cross product has the usual rules of bilinearity and skew symmetry. The triple cross product expansion

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{c} \cdot \mathbf{a})\mathbf{b} - (\mathbf{b} \cdot \mathbf{a})\mathbf{c}$$

is easily verified. The **scalar triple product** is

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \lambda\mu \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

from which all the usual rules (except volume) can be deduced.

Dependancy Criteria

In this subsection we will consider conditions for sets of quaternions to be linearly dependant or otherwise. A non-zero element, A, of $\mathcal{Q}$ is called **isotropic** if $N(A) = 0$ and **anisotropic** otherwise. So only non-zero anisotropic elements have inverses. We note the following theorem.

4.23.1 Theorem *The following statements about a quaternion algebra $\mathcal{Q}$ are equivalent.*

1. *$\mathcal{Q}$ contains an isotropic element.*

2. *$\mathcal{Q}$ is the sum of two hyperbolic planes.*

3. *$\mathcal{Q}$ is not a division algebra.*

4. *$\mathcal{Q}$ is isomorphic to $M_2(F)$ as an algebra.*

Proof *See* [**L**] p 58.

We will call a quaternion algebra above **hyperbolic**. Otherwise it is called **anisotropic**. The classic quaternions are anisotropic: 2×2 matrices are hyperbolic.

4.23.2 Lemma *A pair of pure quaternions $\mathbf{a}, \mathbf{b}$ is linearly dependant if and only if $\mathbf{a} \times \mathbf{b} = \mathbf{0}$.*

Proof The proof is clear one way using the antisymmetry of the cross product. Conversely suppose $\mathbf{a} \times \mathbf{b} = \mathbf{0}$. Then $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{b} \cdot \mathbf{c})\mathbf{a} = \mathbf{0}$. This can be made into a linear dependancy by a suitable choice of $\mathbf{c}$, for example if $\mathbf{a} \cdot \mathbf{c} \neq \mathbf{0}$. $\square$

As a corollary we have the following

4.23.3 Lemma *Two quaternions commute if and only their pure parts are linearly dependant.* $\square$

Now we look for conditions for the triple of pure quaternions, $\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}$, to be linearly dependant. The required condition is given by the following lemma.

4.23.4 Lemma *The pure quaternions $\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}$, are linearly dependant if and only if*

$$N(\mathbf{a})\mathbf{N}(\mathbf{b}) = (\mathbf{a} \cdot \mathbf{b})^{\mathbf{2}}.$$

This is equivalent to the equations

$$N(\mathbf{a} \times \mathbf{b}) = -\mu(\mathbf{a_2 b_3} - \mathbf{a_3 b_2})^{\mathbf{2}} - \lambda(\mathbf{a_1 b_3} - \mathbf{a_3 b_1})^{\mathbf{2}} + (\mathbf{a_1 b_2} - \mathbf{a_2 b_1})^{\mathbf{2}} = \mathbf{0},$$

ie $\mathbf{a} \times \mathbf{b}$ is isotropic or zero.

Proof Three 3-dimensional vectors are linearly dependant if and only if the determinant they form by rows is zero. In the case of pure quaternions this means the scalar triple product is zero

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \lambda\mu \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \mathbf{0}.$$

Replacing $\mathbf{c}$ with $\mathbf{a} \times \mathbf{b}$ and expanding out using the triple cross product formula gives the first equation. Using the expansion formula

$$N(\mathbf{a} \times \mathbf{b}) = \mathbf{N}(\mathbf{a})\mathbf{N}(\mathbf{b}) - (\mathbf{a} \cdot \mathbf{b})^{\mathbf{2}}$$

gives the second formula. $\square$

We have the following corollaries.

4.23.5 Lemma *If* $\mathbf{a}, \mathbf{b}$ *are linearly independant pure quaternions and* $\mathbf{a} \times \mathbf{b}$ *is anisotropic, then the triple* $\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}$, *is linearly independant.*

4.23.6 Lemma $\mathcal{Q}$ *is isomorphic to* $M_2(F)$ *as an algebra if and only if there are linearly independant pure quaternions* $\mathbf{a}, \mathbf{b}$ *such that the triple* $\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}$, *is linearly dependant.*

2×2 matrices

We will interpret all the previous results in terms of 2×2 *matrices,* $M_2(F) = \left(\frac{-1,\ 1}{F}\right)$. *This is the only quaternion algebra with zero divisors.*

The generators of $\left(\frac{-1,\ 1}{F}\right)$ *are, together with the identity, the Pauli matrices*

$$i = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\ j = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ k = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

By an abuse of notation we will often confuse the scalar matrix $\begin{pmatrix} \nu & 0 \\ 0 & \nu \end{pmatrix}$ *with the corresponding field element* ν.

A general matrix can be written uniquely as

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \frac{1}{2}\big[(\alpha+\delta) + (\beta-\gamma)i + (\beta+\gamma)j + (\alpha-\delta)k\big]$$

Conversely

$$A = a_0 + a_1 i + a_2 j + a_3 k = \begin{pmatrix} a_0 + a_3 & a_2 + a_1 \\ a_2 - a_1 & a_0 - a_3 \end{pmatrix}$$

Conjugation is

$$\overline{A} = \mathrm{adj}A = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix} = \begin{pmatrix} a_0 - a_3 & -a_2 - a_1 \\ a_1 - a_2 & a_0 + a_3 \end{pmatrix}$$

and norm is

$$N(A) = \det A = \alpha\delta - \beta\gamma = a_0^2 + a_1^2 - a_2^2 - a_3^2$$

The scalar part of A *is* $a_0 = \mathrm{tr}A/2 = (\alpha+\delta)/2$ *and the pure part is*

$$\begin{pmatrix} a_3 & a_2 + a_1 \\ a_2 - a_1 & -a_3 \end{pmatrix} = \begin{pmatrix} (\alpha-\delta)/2 & \beta \\ \gamma & (\delta-\alpha)/2 \end{pmatrix}$$

Multiplying Matrices

4.23.7 Lemma *Suppose* $A, B \in M_2(F) = \left(\frac{-1,\ 1}{F}\right)$. *Then* $AB = AB$.

The statement is deliberately provocative. It says that multiplying A, B as matrices and as quaternions is the same. This can be checked directly. □

The above lemma allows quick checking of formulæ so if

$$A = \begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{pmatrix} \text{ and } B = \begin{pmatrix} \beta_1 & \beta_2 \\ \beta_3 & \beta_4 \end{pmatrix} \text{ then } A \cdot B = \frac{1}{2}(\alpha_1\beta_4 - \alpha_2\beta_3 - \alpha_3\beta_2 + \alpha_4\beta_1)$$

If $\mathbf{a}=\begin{pmatrix}\alpha_1 & \alpha_2\\ \alpha_3 & -\alpha_1\end{pmatrix}$ and $\mathbf{b}=\begin{pmatrix}\beta_1 & \beta_2\\ \beta_3 & -\beta_1\end{pmatrix}$ are pure then $\mathbf{a}\cdot\mathbf{b}=-\alpha_1\beta_1-(\alpha_2\beta_3+\alpha_3\beta_2)/2$ and

$$\mathbf{a}\times\mathbf{b}=\begin{pmatrix}(\alpha_2\beta_3-\alpha_3\beta_2)/2 & \alpha_1\beta_2-\alpha_2\beta_1\\ \alpha_3\beta_1-\alpha_1\beta_3 & (\alpha_3\beta_2-\alpha_2\beta_3)/2\end{pmatrix}$$

4.24. Solving the Fundamental Equation

In this section we give a sufficient condition for two generalised quaternions to satisfy the fundamental equation. Except for 2×2 matrices, this condition is also necessary.

Let $A=a_0+\mathbf{a}$ and $B=b_0+\mathbf{b}$ be quaternions. We will need the following easily checked lemma.

4.24.1 Lemma *Conjugation by multiplication is* $B^{-1}AB=N(B)^{-1}\overline{B}AB$ *where*

$$\overline{B}AB=a_0(b_0^2+N(\mathbf{b}))+(\mathbf{b_0^2}-\mathbf{N}(\mathbf{b}))\mathbf{a}+\mathbf{2}(\mathbf{a}\cdot\mathbf{b})\mathbf{b}+\mathbf{2b_0}(\mathbf{a}\times\mathbf{b}).$$

The two commutators are

$$[A,B]=AB-BA=2\mathbf{a}\times\mathbf{b}$$

and $(A,B)=A^{-1}B^{-1}AB=N(A)^{-1}N(B)^{-1}\overline{A}\,\overline{B}AB$ *where*

$$\begin{aligned}\overline{A}\,\overline{B}AB=\;&a_0^2b_0^2+b_0^2N(\mathbf{a})+\mathbf{a_0^2}\mathbf{N}(\mathbf{b})-\mathbf{N}(\mathbf{a})\mathbf{N}(\mathbf{b})+\mathbf{2}(\mathbf{a}\cdot\mathbf{b})^{\mathbf{2}}\\ &-2(b_0(\mathbf{a}\cdot\mathbf{b})+\mathbf{a_0}\mathbf{N}(\mathbf{b}))\mathbf{a}+\mathbf{2}(\mathbf{a_0}(\mathbf{a}\cdot\mathbf{b})+\mathbf{b_0}\mathbf{N}(\mathbf{a}))\mathbf{b}+\mathbf{2}(\mathbf{a_0}\mathbf{b_0}-\mathbf{a}\cdot\mathbf{b})\mathbf{a}\times\mathbf{b}\end{aligned}.$$

□

We will call a quaternion, A, **balanced** if $\mathrm{tr}(A)=N(A)\neq 0$. A quaternion $A=a_0+a_1i+a_2j+a_3k$ in a quaternion algebra with parameters λ,μ is balanced if it lies on the quadric 3-fold

$$(a_0-1)^2-\lambda a_1^2-\mu a_2^2+\lambda\mu a_3^2=1,\quad a_0\neq 0.$$

A balanced classical quaternion A lies on the 3-sphere, centre 1 and radius 1. Note that if A is balanced then $N(A-1)=1$. A pair of invertible non-commuting quaternions, A,B, will be called **matching** if A is balanced and $A\cdot B=0$.

4.24.2 Theorem *If the quaternion algebra is anisotropic then a necessary and sufficient condition for the non-commuting, invertible quaternions* A,B *to be solutions of the fundamental equation is that they are a matching pair. Otherwise the condition is only sufficient.*

Proof The proof formally follows **[BuF]**. In terms of quaternions the equation is

$$\overline{A}\,\overline{B}AB-B\overline{A}\,\overline{B}A=N(A)\overline{B}AB-N(A)N(B)A.$$

Using the formulæ and notation developed above, the left hand side is

$$-4a_0N(\mathbf{b})\mathbf{a}+\mathbf{4a_0}(\mathbf{a}\cdot\mathbf{b})\mathbf{b}-\mathbf{4}(\mathbf{a}\cdot\mathbf{b})\mathbf{a}\times\mathbf{b}$$

whereas the right hand side is

$$-2(a_0^2+N(\mathbf{a}))\mathbf{N}(\mathbf{b})\mathbf{a}+\mathbf{2}(\mathbf{a_0^2}+\mathbf{N}(\mathbf{a}))(\mathbf{a}\cdot\mathbf{b})\mathbf{b}+\mathbf{2b_0}(\mathbf{a_0^2}+\mathbf{N}(\mathbf{a}))\mathbf{a}\times\mathbf{b}$$

Considering half the difference of the two sides we arrive at

$$\mathbf{c}=(\mathbf{tr}(\mathbf{A})-\mathbf{N}(\mathbf{A}))\mathbf{N}(\mathbf{b})\mathbf{a}+(\mathbf{N}(\mathbf{A})-\mathbf{tr}(\mathbf{A}))(\mathbf{a}\cdot\mathbf{b})\mathbf{b}+(\mathbf{b_0}(\mathbf{N}(\mathbf{A})-\mathbf{tr}(\mathbf{A}))+\mathbf{2A}\cdot\mathbf{B})\mathbf{a}\times\mathbf{b}$$

So the two sides are equal if $\mathbf{c}=\mathbf{0}$.

Note that we are assuming A,B do not commute. So $\mathbf{a}\times\mathbf{b}\neq\mathbf{0}$. If A,B is a matching pair then $\mathbf{c}=\mathbf{0}$. Conversely if $\mathbf{a},\mathbf{b},\mathbf{a}\times\mathbf{b}$ are linearly independent then their coefficients will be zero and this implies that A,B is a matching pair. The bilinear form is definite unless the algebra is $M_2(F)$ and so, except for this case, $\mathbf{a}, \mathbf{b}, \mathbf{a}\times\mathbf{b}$ will be linearly independent. $\square$

Note: The non-definite case is considered in **[BuF2]**

4.25. The general matching pair

In this section we will use the results of the previous section to describe the most general matching pair of matrices. That is A, B are 2×2 matrices with entries in some field and satisfying

1. $\mathrm{tr}(A) = \det(A)$ $(\neq 0)$

2. $A\mathrm{adj}(B) + B\mathrm{adj}(A) = 0$

Note that condition 2. can also be written $\mathrm{tr}(AB^{-1}) = 0$

Since B can be multiplied by any non-zero scalar we may assume temporarily that

3. $\det(B) = 1$.

We can conjugate the matrices A, B to simplify matters. Consider the following two cases:

Case 1, A has two distinct eigenvalues and is diagonal.

$$A = \begin{pmatrix} a & 0 \\ 0 & a/(a-1) \end{pmatrix} \text{ and } B = \begin{pmatrix} b & c \\ (b^2+a-1)/c(1-a) & b/(1-a) \end{pmatrix}$$

where a, b, c are general and $c, 1-a$ must be invertible (ie non zero).

Inverses are given by

$$A^{-1} = \frac{1}{a}\begin{pmatrix} 1 & 0 \\ 0 & a-1 \end{pmatrix}$$

so a must also be invertible and

$$B^{-1} = \mathrm{adj}(B) = \begin{pmatrix} b/(1-a) & -c \\ -(b^2+a-1)/c(1-a) & b \end{pmatrix}.$$

The 4×4 matrix S is $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ where

$$C = A^{-1}B^{-1}A - A^{-1}B^{-1}A^2 = \begin{pmatrix} b & c/(1-a)^2 \\ (1-a)(b^2+a-1)/c & b/(1-a) \end{pmatrix}$$

and

$$D = 1 - A^{-1}B^{-1}AB = \begin{pmatrix} (2-3a+ab^2+a^2-2b^2)/(1-a)^2 & (a-2)bc/(1-a)^2 \\ (a-2)b(b^2+a-1)/c(1-a) & (2-3a+ab^2+a^2-2b^2)/(1-a) \end{pmatrix}$$

Call this switch $\mathbf{E_2}$

Case 2 A has one eigenvalue and is lower triangular

$$A = \begin{pmatrix} 2 & 0 \\ x & 2 \end{pmatrix}.$$

Then B has the form,

$$B = \begin{pmatrix} y & z \\ (xyz-2y^2-2)/2z & (xz-2y)/2 \end{pmatrix}$$

where x, y, z are general and $2, z$ must be invertible.

Inverses are given by

$$A^{-1} = \frac{1}{4}\begin{pmatrix} 2 & 0 \\ -x & 2 \end{pmatrix}.$$

and

$$B^{-1} = \mathrm{adj}(B) = \begin{pmatrix} (xz-2y)/2 & -z \\ (-xyz+2y^2+2)/2z & y \end{pmatrix}$$

The 4×4 matrix S is $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ where

$$C = A^{-1}B^{-1}A - A^{-1}B^{-1}A^2 = \begin{pmatrix} y+xz & z \\ -(x^2z^2+3xyz+2y^2+2)/2z & -xz/2-y \end{pmatrix}$$

and

$$D = 1 - A^{-1}B^{-1}AB = \begin{pmatrix} xyz/2 & xz^2/2 \\ -x(2y^2+xyz-2)/4 & -xz(xz+2y)/4 \end{pmatrix}$$

Call this switch $\mathbf{E_1}$

If we now multiply B by the scalar t, ie the switch is augmented by t we have two possible switches: one dependant on four variables a, b, c, t (E_2) and one dependant on four variables x, y, z, t (E_1).

4.26. Determinants over Quaternion Algebras

In order to define workable invariants we consider in this section a determinental function on the matrices in $M_n(\mathcal{Q})$. That is $n \times n$ matrices with entries in a quaternion algebra $\mathcal{Q}$. For background reading see [**As**]. In fact the invariants defined later can also be defined for any solutions of the fundamental equation over a ring with a determinant function satisfying the rules listed below.

If R is a commutative ring let $\det : M_n(R) \to R$ denote the usual determinant. The classic quaternions, $\mathbb{H}$, may be embedded as a subalgebra of $M_2(\mathbb{C})$ and determinants taken in the usual way. Our aim is to generalize this.

Suppose $\mathcal{Q}$ has underlying field F and parameters λ, μ. Let $\overline{F}$ denote the algebraic closure of F. Embed i, j, k in $M_2(\overline{F})$ by

$$i = \begin{pmatrix} 0 & \sqrt{-\lambda} \\ -\sqrt{-\lambda} & 0 \end{pmatrix},\ j = \begin{pmatrix} 0 & \sqrt{\mu} \\ \sqrt{\mu} & 0 \end{pmatrix},\ k = \begin{pmatrix} \sqrt{-\lambda\mu} & 0 \\ 0 & -\sqrt{-\lambda\mu} \end{pmatrix}.$$

Define $d : M_n(\mathcal{Q}) \to \overline{F}$ as the composition of the embedding $M_n(\mathcal{Q}) \subset M_{2n}(\overline{F})$ with $\det$.

Alternatively the determinant function may be calculated by induction on the size of the matrices. The value $d(A) = N(A)$ starts the induction. Consider a matrix in $M_n(\mathcal{Q})$. This may be reduced to diagonal form, by multiplying on the left and the right by elementary matrices having unit determinant, (see below). Suppose this matrix has diagonal elements $d_1, \ldots, d_n$. Define the determinant as $d = N(d_1)\cdots N(d_n)$. So the determinant function takes values in $\mathcal{N}$. For $M_2(F)$ this subset is the whole of F: for classic quaternions it is the non-negative reals.

The determinant function satisfies the rules

0. $d(M) = 0$ if and only if M is singular, moreover $d(MN) = d(M)d(N)$. It follows that $d(1) = 1$.

1. d is unaltered by a permutation of the rows (columns).

2. If a row (column) is multiplied on the left (right) by a unit then d is multiplied by d of that unit.

3. $d(M)$ is unaltered by adding a left multiple of a row to another row or a right multiple of a column to another column.

4. $d\begin{pmatrix} x & \mathbf{u} \\ \mathbf{0} & M \end{pmatrix} = N(x)d(M)$ where $\mathbf{u}$ is any row vector and $\mathbf{0}$ is a zero column vector.

4'. $d\begin{pmatrix} x & \mathbf{0} \\ \mathbf{v} & M \end{pmatrix} = N(x)d(M)$ where $\mathbf{v}$ is any column vector and $\mathbf{0}$ is a zero row vector.

5. $d(M^*) = \partial(M)$ where $M^* = \overline{M^T}$ denotes the Hermitian conjugate.

6. if the entries in M all commute then $d(M) = \det^2(M)$.

An **elementary** matrix of type 1 is a permutation matrix. An elementary matrix of type 2 is the identity matrix with one diagonal entry replaced by a unit and an elementary matrix of type 3 is a square matrix with zero entries except for 1's down the diagonal and one other entry off diagonal.

The properties i. above for $i = 1, 2, 3$ follow from multiplying M on the right or left by an elementary matrix of type i.

The matrix S can be written as a product of elementary matrices

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ C & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1-A^{-1} \end{pmatrix}\begin{pmatrix} 1 & A^{-1}B \\ 0 & 1 \end{pmatrix}.$$

Note that $1 - A^{-1}$ is invertible.

Hence $d(S) = d(A)d(1 - A^{-1}) = d(A - 1) = 1$. Therefore the representation of VB_n, induced by such an S, is into $SL(F, 2n)$.

4.27. Virtual knots and the invariant knot modules

We saw in the last section that given a biquandle we can label the semi-arcs with generators and define a presentation using the switch on each classical crossing. Given an associative ring R, a 2×2 matrix S with entries in R and a virtual link diagram $\mathcal{D}$, we define a presentation of an R-module which depends only on the link class of the diagram provided S is a switch. This construction also works for classical knots and links but is only the Alexander module in disguise. The generators are the semi-arcs of $\mathcal{D}$, that is the portion of the diagram bounded by two adjacent classical crossings. There are 2 relations for each classical crossing.

Suppose the diagram $\mathcal{D}$ has n classical crossings. Then there are $2n$ semi-arcs labelled $a, b, \ldots$. These will be the generators of the module. Let the edges of a positive real crossing in a diagram be arranged diagonally and called geographically *NW, SW, NE* and *SE*. Assume that initially the crossing is oriented and the edges oriented towards the crossing from left to right ie west to east. The **input** edges, oriented towards the crossing, are in the west and the edges oriented away from the crossing, the **output** edges, are in the east. Let a and b be the labels of the input edges with a labelling SW and b labelling NW. For a positive crossing, a will be the label of the undercrossing input and b the label of the overcrossing input. Suppose now that

$$S(a, b)^T = (c, d)^T \text{ where } S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

Then we label the undercrossing output NE by d and we label the overcrossing output SE by c.

For a negative crossing the direction of labelling is reversed. So a labels SE, b labels NE, c labels SW and d labels NW.

Finally for a virtual crossing the labellings carry across the strings. This corresponds to the twist function $T(a,b) = (b,a)$.

The knot diagram therefore gives rise to a presentation of an R-module with $2n$ generators and $2n$ relations. Note that in all cases B, C are invertible since the identity switch is uninteresting.

4.27.1 Theorem *The module defined above for any diagram D is invariant under the Reidemeister moves, and hence is a knot invariant, if S is a switch.* □

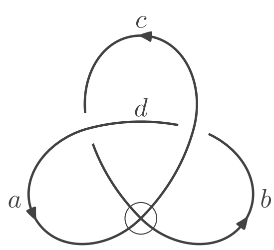
4.28. Determinant Invariants

The Determinant Δ_0

Given a module with a square presentation the obvious invariant of the module is the determinant, if it can be defined. This will be the case if the ring is represented by matrices with commuting entries, for example the ring of generalised quaternions where each quaternion is represented by a 2×2 matrix with complex entries. In this case if d denotes the determinant and $M\mathbf{x} = \mathbf{0}$ is the presentation let $\Delta_0 = d(M)$. Since the module depends on the switch S we illustrate this dependency by $\Delta_0 = \Delta_0(S)$.

A close look at how the presentation of the module changes under the Reidemeister moves shows that Δ_0 is invariant up to multiplication by $d(B)$ or $d(C)$. Typically $d(B)$ is denoted by the variable t and $d(C)$ is t^{-1}. If we take the switch to be E_1 (E_2) defined in section 5 then Δ_0 is a polynomial p_1 (p_2) in the four variables x, y, z, t (a, b, c, t). We can normalise these polynomial so that as a polynomial in t it has a non-zero constant term and only positive powers of t.

Let us illustrate the previous discussion by calculating invariants for the **virtual trefoil** as shown in the figure.



If we label as indicated then the module has a presentation with 4 generators a, b, c, d and relations $c = Ab + Ba$, $a = Ac + Bd$, $b = Cc + Dd$, $d = Cb + Da$. Restricting to the E_1 case gives the polynomial p_1 equal to

$$\frac{64 + 4x^3tz^3 + 4x^3t^3z^3 + 128t^2 - 64xtz - x^4t^2z^4 - 64xt^3z + 8x^2t^2z^2 - 4x^2z^2 + 64t^4 - 4x^2t^4z^2}{16}$$

For the E_2 case we get the polynomial p_2 equal to

$$-\frac{\begin{array}{c} -4 + 16\,a - 40\,tba^3 - 4\,t^4a^4 + 16\,t^4a^3 - 24\,t^4a^2 + 16\,t^4a - 36\,tb^3a^2 + 72\,a^2tb + 4\,b^2 + 13\,a^2t^4b^2 \\ +40\,tb^3a + 4\,t^4b^2 - 4\,t^4 - 24\,a^2 - 12\,b^2a + 13\,a^2b^2 + 16\,a^3 - 12\,t^4b^2a + 16\,tb - 6\,t^4b^2a^3 \\ +t^4b^2a^4 - 56\,atb - 16\,tb^3 + b^2a^4 - 6\,b^2a^3 + 8\,ba^4t - 2\,b^3a^4t + 14\,b^3a^3t - 4\,a^4 - 8\,t^2a^4 + 32\,t^2a^3 \\ -48\,t^2a^2 + 32\,t^2a - 8\,t^2 + a^4t^2b^4 - 2\,a^4t^2b^2 - 8\,a^3t^2b^4 - 8\,t^2b^2 + 16\,t^2b^4 + 16\,t^3b - 16\,t^3b^3- \\ 26\,a^2t^2b^2 + 12\,a^3t^2b^2 + 24\,a^2t^2b^4 + 8\,a^4t^3b - 40\,a^3t^3b + 72\,a^2t^3b + 24\,at^2b^2 - 32\,at^2b^4 \\ -56\,at^3b - 2\,t^3b^3a^4 + 14\,t^3b^3a^3 - 36\,t^3b^3a^2 + 40\,t^3b^3a \end{array}}{t^2\,(a-1)^4}$$

Note that the fundamental quandle (and hence group) as defined by the Wirtinger presentation is trivial.

The Determinant Δ_1

For many knots and links, including the classical, the determinant Δ_0 is zero.

For example, as we have seen earlier any switch S with entries in the ring R defines a representation of the virtual braid group VB_n into the group of invertible $n \times n$ matrices with entries in R by sending the standard generator σ_i to $S_i = (id)^{i-1} \times S \times (id)^{n-i-1}$ and the generator τ_i to $T_i = (id)^{i-1} \times T \times (id)^{n-i-1}$. This representation is denoted by $\rho = \rho(S, n)$. For classical braids this representation is equivalent to the Burau representation and so we would expect the closure of a classical braid to have Δ_0 zero. We now confirm this by looking at the fixed points of S_i both on the left and right.

4.28.1 Lemma *Let $P = A^{-1}B^{-1}A$ and $Q = B^{-1}(1 - A)$. Then*

$$(P^{n-1}, \dots, P, 1)S_i = (P^{n-1}, \dots, P, 1) \qquad (*)$$

and

$$S_i(1, Q, \dots, Q^{n-1})^T = (1, Q, \dots, Q^{n-1})^T \qquad (**).$$

Proof We need only check that

$$(P, 1)S = (P, 1) \text{ and } S(1, Q)^T = (1, Q)^T.$$

□

Therefore the following lemma gives a necessary condition for the knot or link to be classical.

4.28.2 Theorem *For all classical knots and links $\Delta_0 = 0$.*

Proof Since Δ_0 is an invariant of the module we can assume that the diagram from which it is defined is the closure of a braid. However from $*$ (or $**$) there is a linear relationship amongst the rows (columns), so $\Delta_0 = 0$. □

The Kishino knot is illustrated below.

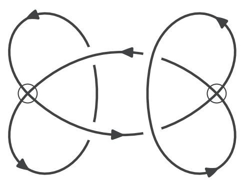


fig: The Kishino Knot

This is one of the ways of forming the connected sum of two unknots. The invariant Δ_0 is zero.

It is clear that we need an invariant in this case. Let M be the $2n \times 2n$ presentation matrix. Let $M_1, M_2, \dots, M_{(2n)^2}$ be the submatrices obtained by deleting a row and a column. Let $d_1, d_2, \dots, d_{(2n)^2}$ be the determinants. These all lie in a ring of polynomials with coefficients in a field. Therefore the determinants have an hcf, call it Δ_1, which is well defined up to multiplication by a unit. Now look at what happens to this construction under a Reidemeister I move. The hcf, Δ_1, is multiplied or divided by d of a unit.

A calculation with the Alexander switch shows that for the Kishino knot, Δ_1 is 1. But we can show that it is non-trivial by using the Budapest switch augmented by t. Then $\Delta_1 = 2 + 5t^2 + 2t^4$ see [**BuF**].

For a classical knot or link the invariant Δ_1 is not just the Alexander polynomial in disguised form but is independant of the deleted rows or columns chosen, up to multiplication by a power of t.

4.28.3 Theorem *Let $\mathcal{D}$ be a diagram of a classical knot or link and let M be the presentation matrix associated with $\mathcal{D}$. Consider $d(M_{ij})$ where M_{ij} is obtained from M by deleting the ith row and the jth column. Then $d(M_{ij})$ is independant of i, j up to multiplication by a power of t and is equal to Δ_1.*

Proof Assume initially that $\mathcal{D}$ is the closure of a braid.

Write

$$M = (\, C_1 \quad C_2 \quad \dots \quad C_{2n} \,)$$

in terms of its columns and let

$$M_{ij} = (\, C_1^0 \quad C_2^0 \quad \dots \quad C_{2n}^0 \,)$$

where each column has its ith component removed and C_j^0 does not appear in the list. From $**$,

$$C_j^0 = -C_1^0 Q^{(1-j)} - \cdots - C_{2n}^0 Q^{2n-j}$$

and C_j^0 does not appear on the right hand side of the equation.

So by column operations which do not change the value of the determinant we can change any column to C_j^0. Now note that the value of the determinant is unchanged by interchanging two columns. A similar argument works for the rows.

A general diagram is obtained from $\mathcal{D}$ by a sequence of Reidemeister moves. A glance at the change of M under the Reidemeister moves shows that Δ_1 is invariant up to multiplication by a power of t . $\square$

Let us now return to the Kishino knot. This is the closure of the braid

$$\tau_2(\sigma_1\sigma_2\sigma_1)\tau_2(\sigma_1\sigma_2\sigma_1)^{-1}$$

Suppose the representation of this as a 3×3 matrix, using the Budapest switch augmented by t, is M. Then the representation matrix of the module is $M - id$. The nine codimension 1 subdeterminants are

$$p = 2t^{-2} + 5 + 2t^2,\ (4 \text{ times})\ q = (2 + 2t^2)p(t^{-1}),\ (\text{twice})\ q(t^{-1}), p^2$$

This not only shows that it is non-trivial but that it cannot be classical.

5. Flat Knots

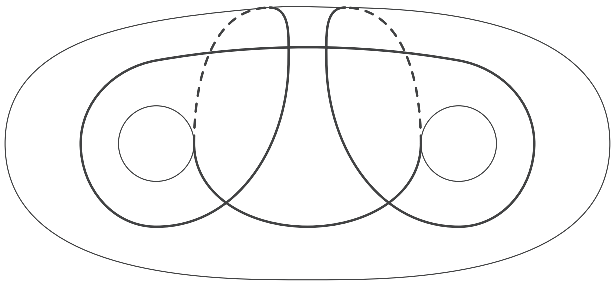
In this chapter we see how the previous methods can be applied to flat knots. That is virtual knots in which the classical crossings have no distinction between over and under arcs. Flat knots may be thought of as immersed curves on a surface modulo cobordism.

It is a consequence of the condition on classical crossings that for flat knots the switch must satisfy $S = S^{-1}$ or $S^2 = 1$. This naturally leads to the Weyl algebra, sometimes called the Heisenberg algebra of the harmonic oscillator. This has generators u, v satisfying the relation $uv - vu = 1$.

To indicate the sort of theorem this allows, consider the following.

5.28.1 Theorem *Let W be any algebra with invertible elements u, v satisfying $uv - vu = 1$. Put $A = v^{-1}u^{-1}$, $B = u$, $C = uvu^{-1}v^{-1}u^{-1}vuv^{-1}u^{-1}$, $D = -u^{-1}v^{-1}$ and $S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Then the $W-$module defined in the usual way is an invariant of flat knots.*

For example consider the flat Kishino knot, defined as an immersed curve on the closed orientable surface of genus 2 and illustrated below.



Using the representation of the Weyl algebra given by

$$u = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad v = \begin{pmatrix} y & 0 & 0 \\ 1 & y & 0 \\ 0 & 2 & y \end{pmatrix}$$

with underlying ring $\mathbb{Z}_3[y]$ we find

$$\Delta_0 = 0, \quad \Delta_1 = 2 + 2y.$$

Since $\Delta_1 \neq 1$, the Kishino flat knot is indeed non-trivial.

References

[Azc] H.Azcan, *Spherical representations of link quandles*, Ph.D. thesis, Sussex University 1991
[BF] A. Bartholomew and Roger Fenn. Quaternionic Invariants of Virtual Knots and Links, JKTR
[Bir] J.Birman, *Braid links and mapping class groups*, Annals of Math. Studies #82, Princeton 1975
[Bir2] J.Birman, *New Points of View in Knot Theory*, to appear in Bull. A.M.S.
[Br] E.Brieskorn, *Automorphic sets and singularities*, Contemporary maths., 78 (1988) 45-115
[BuF] S. Budden and Roger Fenn. The equation

$$[b, (a-1)(a, b)] = 0$$

and virtual knots and links,
[BuF2] S. Budden and Roger Fenn.Quaternion Algebras and Invariants of Virtual Knots and Links II: The Hyperbolic Case to appear in *JKTR*
[BZ] G.Burde and H.Zieschang, *Knots*, Walter de Gruyter Studies in Maths. #5, 1985
[CS] J. S. Carter, D. Jelsovsky, Seiichi Kamada, Laurel Langford, Masahico Saito. Quandle Cohomology and State-sum Invariants of Knotted Curves and Surfaces
http://arxiv.org/abs/math.GT/9903135
[CW] J.C.Conway and G.C.Wraith, correspondence, 1959
[Cox] H.S.M.Coxeter, *Regular complex polytopes*, C.U.P. 1974
[D] P.Dehornoy, *Free distributive groupoids*, J. Pure and Applied Algebra, 16 (1989) 123-146
[D] P. Dehornoy. Braids and Self-Distributivity. Progress in Mathematics no. 192 Birkhäuser, (2000).
[Dev] D.Devine, *Alexander invariants of links and the fundamental rack*, M.Phil. thesis, Sussex University 1991

[Dr] V. Drinfeld, On some Unsolved Problems in Quantum Group Theory, Quantum Groups, Lectures Notes in Maths. 1510, Springer 1-8 (1990)

[ESS] P. Etingof, T. Schedler and A. Soloviev, Set-Theoretic Solutions to the Quantum Yang-Baxter Equations. Duke Math Journal 100 no. 2 169-209 (1999)

[EGS] P. Etingof, R. Guralnik and A. Soloviev, Indecomposable set-theoretical solutions to the Quantum Yang-Baxter Equation on a set with prime number of elements. J of Algebra 242 (2001) 709-719

[FR] R.Fenn and C.Rourke, *The Braid–Permutation Group*, Topology, 18 (1979) 1-15

[FR2] R.Fenn and C.Rourke, *On Kirby's calculus of links*, Topology, 18 (1979) 1-15

[FRB1] R.Fenn, C.Rourke and B.Sanderson, *An introduction to species and the rack space*, Proceedings of the Topology Conference, Erzurum, 1992

[FBR2] R.Fenn, C.Rourke and B.Sanderson, *Trunks and classifying spaces*, Applied Categorical Structures, Volume 3 (1995) pages 321-356

[FR] **R.Fenn, C.Rourke**, *Racks and Links in Codimension Two*, Journal of Knot Theory and its Ramifications, Vol. 1 No.4 (1992) 343–406

[FRR] **R.Fenn, R.Rimanyi, C.Rourke** *Some Remarks on the Braid–Permutation Group* to appear in the Proceedings of the Erzurum Conference, Turkey (1992)

[Hem] J. Hempel, *3–manifolds*, Ann. of Math. Studies #86 Princeton 1976

[Hum] J.E.Humphries, *Reflection groups and Coxeter groups*, Cambridge Studies in Advanced Math. #29, C.U.P. 1990

[[**Hum2**]] **S.P.Humphries**, *On Weakly Distinguished Bases and Free Generating Sets of Free Groups*, Quart. J. Math. Oxford 920, 36 (1985), 215–219

[Je] T.Jech, *A short proof of the irreflexivity conjecture*, preprint, Penn. State University

[LYZ] Jiang-Hua Lu, Min Yan, and Yong-Chang Zhu, On the set-theoretical Yang-Baxter equation, Duke Math. J. 104 1-18 (2001)

[LYZ2] Jiang-Hua Lu, Min Yan, and Yong-Chang Zhu, Quasi-triangular structures on Hopf algebras, Contemp. Math. 267, 339-356 (2000)

[Jo] V.Jones, *Polynomial invariants of knots via von Neumann algebras*, Bull. Amer. Math. Soc. 12 (1985) 103-111

[J] D.Joyce, *A classifying invariant of knots; the knot quandle*, J. Pure Appl. Alg., 23 (1982) 37-65 **[KK]** N. Kamada and S. Kamada, Abstract Link Diagrams and Virtual Knots,
JKTR 12, 93-109 (2003).

[Ku] G. Kuperberg, What is a Virtual Link? Algebraic and Geometric Topology 587-591 (2003)

[K] L.Kauffman. Virtual Knot Theory, European J. Comb. Vol 20, 663-690, (1999)

[K2] Louis H. Kauffman. Detecting virtual knots. Atti Sem. Mat. Fix. Univ. Modena, Supplemento al Vol. IL, 241-282 (2001).

[K] L.H.Kauffman, *Knot-crystals – classical knot theory in modern guise*, Knots and Physics, World Sci.

[KR] Louis H. Kauffman and David E. Radford. Bioriented quantum algebras and a generalized Alexander polynomial for virtual links. in "Diagrammatic Morphisms and Applications (San Fransico, CA 2000), 113-140. contemp. Math. 318, AMS, Providence, RI (2003).

[Kel] A.Kelly, *Groups from link diagrams*, Ph.D. thesis, Warwick University, 1991

[Kh] M.Khovanov, *New generalisations of braids and links*, preprint

[KS] T. Kishino and S. Satoh, A note on classical polynomials, preprint (2001)

[Ko] **K.H.Ko**, *Pseudo-conjugations*, Bull. Korean Math. Soc. 25 (1988), No 2, 247–251

[19] **R.Kirby**, *A calculus for framed links in* S^3, Invent. Math., 45 (1978) 35-56

[Kr] B.Krüger, *Automorphe Mengen und die artinschen Zopfgruppen.* Dissertation, Rheinische Friedrich-Wilhelms-Universität Bonn, 1989 Bonn Mathematische Schriften, 207

[Lam] S.Lambropoulou, *Markov's theorem for general 3–manifolds and applications*, Ph.D. thesis, Warwick University, 1992

[La] R.Laver, *The left distributive law and the freeness of an algebra of elementary embeddings*, Advances in Math.,

91 (1992) 209-231
[LS] R.C.Lyndon and P.E.Schupp, *Combinatorial group theory*, Springer-Verlag, Ergebnisse series #89, 1977
[Mat] S. Matveev, *Distributive groupoids in knot theory*, Math. USSR Sbornik, 47 (1984) 73-83
[O] A. Odesskii, Set-theoretical solutions to the Yang-Baxter Relation from factorization of matrix polynomials and θ-functions. http://arxiv.org/abs/math/0205051
[RT] N. Reshetikhin and V. Turaev, *Invariants of 3–manifolds via polynomials and quantum groups*, preprint
[Ros] A.W.Roscoe, *Maintaining consistency in distributed databases*, preprint, Oxford University Computing Laboratory
[Ryd] H.Ryder, *The structure of racks*, Ph.D. thesis, Warwick University, 1992
[SAW] J. Sawollek, On Alexander-Conway polynomials for virtual knots and links, (arXiv:math.GT/9912173 21Dec 1999.) preprint
[Scott] D. Scott. Continuous Lattices, Toposes, Algebraic Geometry and Logic, Lecture Notes in Mathematics Vol. 274, Springer-Verlag, (1972)
[SW] D.S. Silver and S.G. Williams, Alexander Groups and Virtual Links, to appear in JKTR.
[S] Alexandre Soloviev, Non-unitary set-theoretical solutions to the Quantum Yang-Baxter Equation, math.QA/0003194
[Stan]D. Stanovsky, On axioms of biquandles, J. Knot Th. Ramifications 15/7 (2006). 931933.
[Tak] M. Takasaki, *Abstraction of symmetric transformations,* Tohoku Math. J. 40 (1943), 145-207
[W]M. Wada. Twisted Alexander Polynomial for finitely presentable group. Topology 33 No. 2. 241-256 (1994)
[Wal] F. Waldhausen, *Gruppen mit Zentrum und 3-dimensionale Mannigfaltigkeiten*, Topology 6 (1967) 505-517
[Wh] J.H.C.Whitehead, *On adding relations to homotopy groups*, Annals of Math. (2) 42 (1941) 409-428
[Win] S. K. Winker, *Quandles, Knot Invariants, and the n-fold Branched Cover*, Doctoral Thesis, Chicago Illinois, 1984