INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

NOTE TO USERS

This reproduction is the best copy available.

UMI

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A

Loops, Waves, and an "Algebra" for Heegaard Splittings

Joel Zablow

A dissertation submitted to the Graduate Faculty in Mathematics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1999

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

UMI Number: 9946235

Copyright 1999 by
Zablow, Joel

All rights reserved.

UMI Microform 9946235
Copyright 1999, by UMI Company. All rights reserved.

This microform edition is protected against unauthorized
copying under Title 17, United States Code.

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

©1999
Joel Zablow
All Rights reserved

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

This manuscript has been read and accepted by the Graduate Faculty in Mathematics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

9/22/99
Date

Martin Bendersky
Chair of Examining Committee

9/23/99
Date

Executive Officer

Martin Bendersky

Joseph Roitberg

Alphonse Vasquez

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Loops, Waves, and an "Algebra" for Heegaard Splittings

by

Joel Zablow

Advisor: Prof. Martin Bendersky

In this thesis, I consider an algebraic structure involving isotopy classes simple closed curves (circles) on a closed orientable surface F. We may consider$F$ to be the boundary of a handlebody H of genus g. I relate this to Heegaard splittings, (H_1, H_2, F), of closed orientable 3-manifolds, M. The operations of this structure are used in the proof of topological results. In particular, I prove results about reducibility of splittings, characterization of loops on F bounding disks in H (i = 1,2), and properties of iterated connected sums. I also give a presentation of a subgroup of MCG(F) which preserves the number of waves of a given splitting, and use this to reprove a result of Homma, Ochiai, and Takahashi. A number of algebraic relations in the structure are detailed and some questions are asked about further connections between the algebra and the topology.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Acknowledgements

In the course of thinking about and writing this thesis, I have received help and support from any number of people. I would thus like to thank my advisor, Prof. Martin Bendersky, for having put up with this project, and with me, for so long. I would especially thank him for having allowed himself to be dragooned into an enterprise so far afield of his immediate interests, and for allowing me the time and the lattitude to pursue this. His patience, careful criticism, and willingness to countenance all of this new material, were amazing.

I would further like to thank Prof. Joan Birman of Columbia most gratefully, for providing me with a sounding board for all sorts of (sometimes nebulous) ideas, and for her comments, criticism, and recommendations of directions in which to pursue these ideas. She generously allowed me to bend her ear and spend her time (sometimes needlessly), when I was in no way her 'responsibility', even amidst a host of her own students. I appreciate the support she showed me, and belief that I could indeed carry this out.

I would further like to thank Profs. Alex Heller and Al Vasquez for their openness in allowing me to drop in and discuss a number of these notions, no matter how ill-formed at the time, and their willingness to listen to much of this foreign material. I thank them for their patience with me, and for attempting to elicit clarity, where perhaps initially, none existed, and for their comments and criticisms, which were virtually all warranted, and on the nose.

I would like also to extend thanks to Prof. Dennis Sullivan for having sat through my description of the nearly all of the contents of this work, and for having forced me to clarify the statements and the proofs, above all in my own mind. In particular, this led to fixing a mistake in a proof of one of the main theorems, and streamlining and injecting rigor into the statements of a number of others. Also, in this regard, I must extend many thanks to Prof. Bryan Clair, for suggestions in the streamlining of the proof of Thm. 4.1) and for his willingness to listen to me blather on about all of this, ad infinitum.

I appreciate greatly the time and effort that Prof. Joseph Dodziuk has spent in helping me learn TEX and turning all of this into the physical object which is this thesis. This, amid all of the complications that have come about due to my attempts to put in an inordinate number of pictures, in color no less. His patience and help are most gratefully acknowledged.

I finally must thank my family and friends for having put up with this, and me, for so godawfully long, and for their assurance that it would indeed come to pass. Thank you all.

This thesis is dedicated to the memory of Nan Schmitz, who delighted in art, nature, and the ability to get things done.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Contents

0 Introduction **1**

1 Background **3**

- 1.1 Heegaard Splittings 3
- 1.2 The Mapping Class Group of a Closed Orientable Surface F 7

2 Some Algebra: Quandles **9**

- 2.1 Quandle Definitions and Proposition. 9
- 2.2 Connected Sums, $\#_\gamma$ 13
- 2.3 Further Algebraic Operations and Characteristics 13

3 Some Relations and Consequences in the Extended Quandle **18**

4 Disks and Spheres **29**

5 Waves **34**

- 5.1 Reducibility of Splittings 34
- 5.2 Further Results on Waves, and Relating Waves to a Subgroup of $MCG(F)$ 40

6 Dessert: Quirks, Oddities, and some Post-Prandial Speechifying **54**

7 References **63**

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

0 Introduction

In this thesis I consider an algebraic structure for simple closed curves on the boundary F, of a handlebody H, of genus g, and relate this to Heegaard splittings of closed orientable 3-manifolds, M. Chapter 1) gives some background information regarding Heegaard splittings, and the mapping class group $MCG(F)$. In Chapter 2) I give a definition of a "quandle" as described in [Kauffman], and show that the collection P, of isotopy classes of simple closed curves, (circles) on F forms a quandle, with the action of MCG(F) via Dehn twist on the circles of P. I then introduce some further operations on P, including connected sum of circles, and another operation ν_γ, which turns out to be an inverse of the connected sum operation, when it is defined. The structure consisting of the collection of isotopy classes P, together with the operations Dehn twist, connected sum, and the others, I call the extended quandle. The remaining portions of this section, and of the beginning of the next section, are devoted to giving some algebraic relations among the operations in the extended quandle.

The last segment of Chapter 3) gives a theorem, Thm. 3.1) which relates an algebraic condition for iterated connected sums of circles on F, to the topological imbedding of these circles in F. This then, is the beginning of some topological ramifications of the algebraic structure. In the next chapter, I look more deeply at some of the other properties of the connected sum operation and prove Thm. 4.1), which characterizes circles on $F = \partial H$ which bound disks in H, as connected sums of "meridian" circles on F. Some algebraic corollaries to a theorem of [Casson, Gordon] follow from this.

In Chapter 5) the idea of a "wave" for a Heegaard splitting is introduced. Existence of waves, for a splitting, was shown in Ch. 3) to essentially correspond to the ability to find inverses of connected sums of circles on F. A theorem of [Starr] also gives a correspondence between waves and compression disks. Using these, we prove Thm. 5.1) which says that if a splitting has a collection of waves which kill the genus of a subsurface, then the splitting is *reducible*, i.e. there is a 2-sphere which meets the splitting surface F in a single essential circle. A partial converse is also proven. The remainder of the chapter is given over to further exploration of waves, and their effects on splittings. I look at a subgroup of $MCG(F)$ which preserves the number of waves associated to a given splitting, and Thm. 5.3) characterizes and gives a presentation for this subgroup. This subgroup and its properties are then used to reprove a result of [Homma, Ochiai, Takahashi]; Among genus 2 splittings of S^3, all but the standard splitting, admit waves.

The final chapter examines some further algebraic structures that arise from consideration of the extended quandle, and draws some parallels (perhaps questionable) between these and some

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

known algebraic and topological structures. Also, a number of further questions and avenues of inquiry are proposed.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1 Background

1.1 Heegaard Splittings

We will use Heegaard splittings to construct closed orientable 3-manifolds from pairs of homeomorphic handlebodies of genus g. In particular, let H_1 and H_2 be two copies of a handlebody of genus g. Let $\partial H_1 = \partial H_2 = F$, a closed surface of genus g. Consider a collection of g meridian disks in H_2, m_i, as shown in Fig. 1.1, along with their boundary circles. (waist curves w_i are also shown)

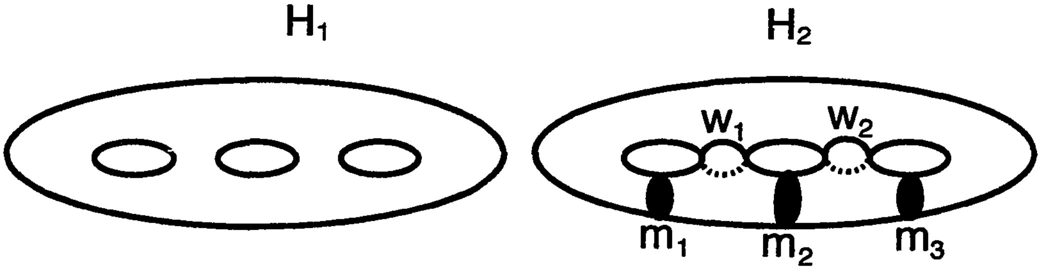
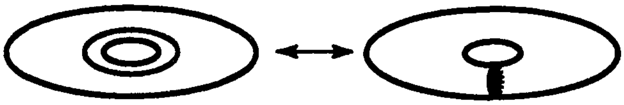
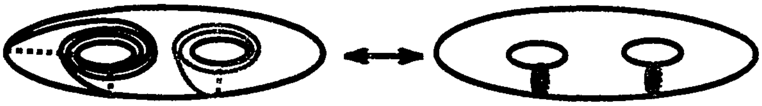


Fig. 1.1)

We form a Heegaard splitting of genus g for a (closed orientable) manifold M, using the handlebodies H_1 and H_2 and the surface F, by glueing H_1 and H_2 along F via an orientation reversing homeomorphism. Thus we form $M = H_1 \bigcup_F H_2$ so that the meridian circles of H_2 are glued to g disjoint, simple, closed, nonseparating curves on $F = \partial H_1$, called characteristic curves. Fig. 1.2 gives some examples.



yields S^3



yields L(3,2) # (L(3,1)

Fig. 1.2)

To view this construction in another way, take the handlebody H_1 with g simple closed curves on it's boundary F, and glue g 2-handles to these curves, identifying the boundaries of the core

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

disks to the curves. Finally attach 3-handles to the boundary of the resulting object, capping off any boundary spheres. A 3-manifold is completely determined by specifying a splitting surface M, the g characteristic curves, corresponding to disks in H_2, and g simple closed curves corresponding to disks in H_1. For the moment we will take these latter curves to be the g meridians of H_1.

Definition 1.1 *The Heegaard genus of a 3-manifold is the minimum genus of the splitting surface F for M over all possible splittings for M.*

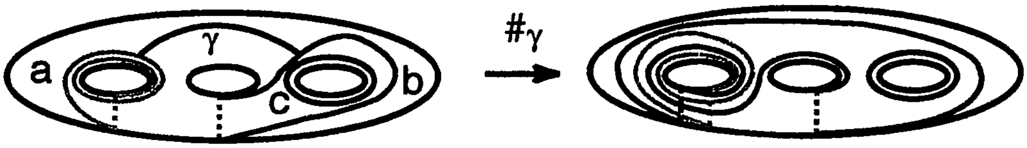
For example, S^3 is the only 3-manifold with Heegaard genus 0, and is created by glueing two copies of B^3 along their boundary spheres. The lens spaces $L(p,q)$ all have splittings with splitting surface $F \cong T^2$, the torus, and hence have Heegaard genus 1.

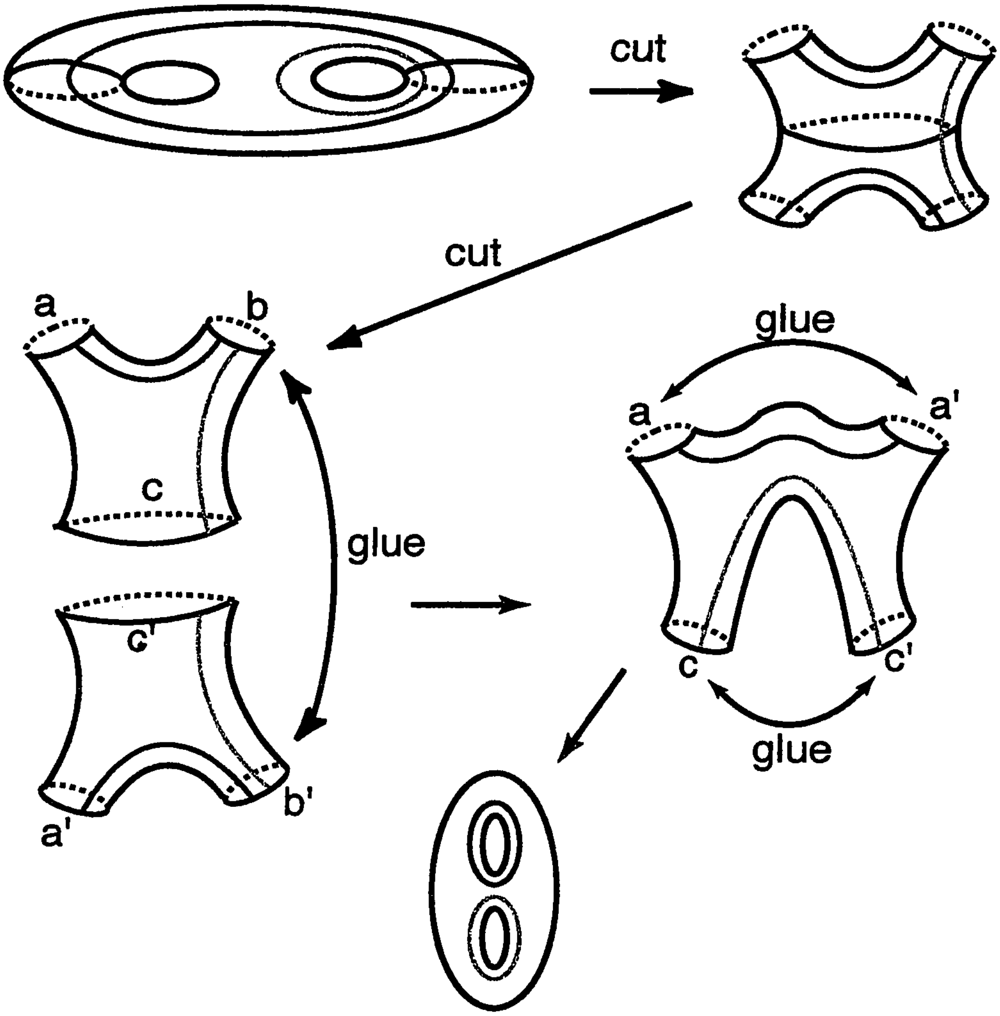
One of the difficulties arising out of the representation of 3-manifolds via Heegaard splittings is that a given manifold M will have an infinite number of representations through surfaces of different genera. Even if we restrict attention to surfaces of a particular genus, we have many different options for the characteristic curves of the splitting. The curves of one splitting of M, on F, may not be in the same isotopy class (on F) as the curves of another splitting of M. It is clear though, that manifolds arising from choosing one or another curve in a given isotopy class, are homeomorphic. As we see however, there is generally a massive nonuniqueness of representation.

In the mid '30's , Reidemeister and Singer developed " moves" that allowed passage between one Heegaard splitting of a manifold M and another, i.e. moves that preserve the homeomorphism type of the underlying manifold, while perhaps changing the characteristic curves and possibly the genus of the splitting. The moves are as follows, and some are illustrated in Fig. 1.3. See also [Singer].

1. change orientation of a characteristic curve
2. take a connected sum along a path γ (disjoint from all other curves), which joins any pair of characteristic curves a and b, taking as the resulting pair, one of the original circles and the connected sum, $a \#_\gamma b$.
3. cut F along meridian loops so that the result is a sphere with 2g holes. Remember which boundary circles on the sphere arose from cutting a given meridian loop. Let a and b be two such boundary components, arising from cutting two such meridians. Find a circle c on the sphere with holes such that a and b lie in the interior of a disk bounded by c. Cut the sphère along c. Reglue either a and its mate a' or b and its mate b'. We again have a sphere with 2g holes. Now reglue corresponding mated pairs of circles.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

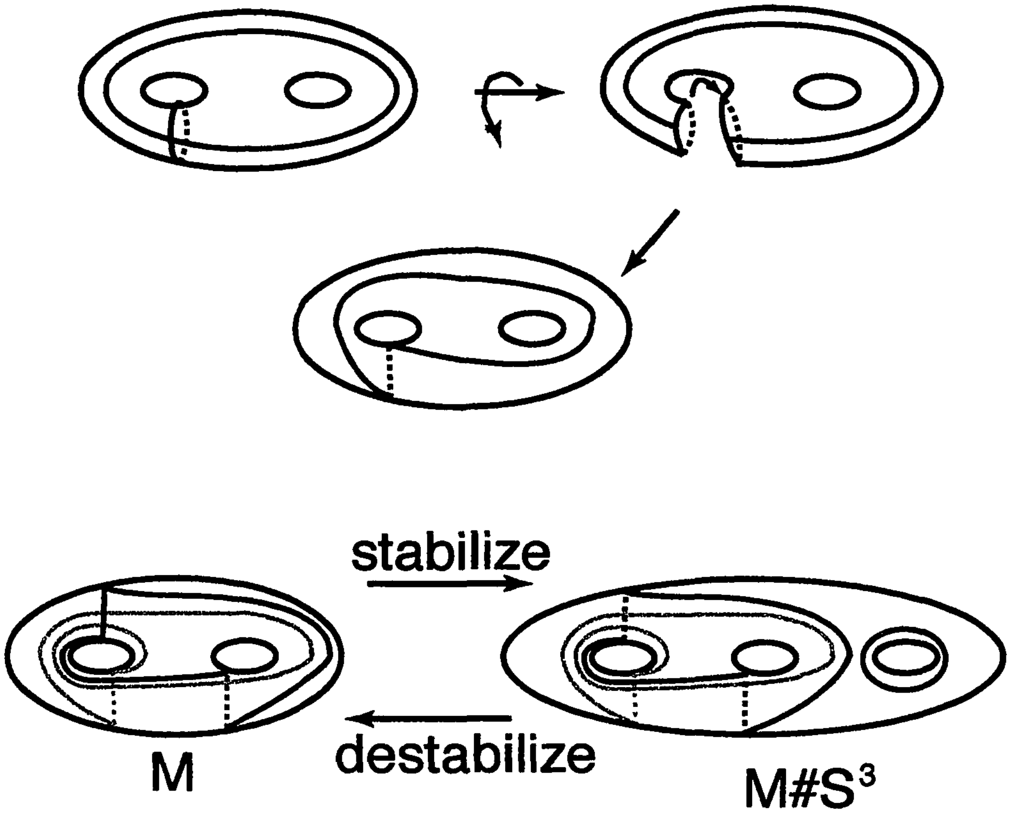




Figs. 1.3.2, 1.3.3)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4. cut F along a meridian circle and twist one of the resulting boundary components through 2π, then reglue.
5. add a new handle to F, increasing the genus by 1, and take a characteristic curve associated to this new handle to be the longitudinal curve around the new hole.



Figs. 1.3.4, 1.3.5)

The fifth move described is called stabilization, and it corresponds to taking a connected sum $M\#S^3$. I shall discuss this further later on.

There is the following theorem due to Reidmeister and Singer.

Theorem 1.1 *(R-S) For a given manifold M, any two Heegaard splittings F and F' are equivalent via moves 1-5.*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In other words, it is possible to get from any splitting , given by a surface F and its characteristic curves, to any other splitting F' and its curves, by applying the above moves and their inverses to F. The process may involve going through splittings of higher or lower genus. (use of move 5)

1.2 The Mapping Class Group of a Closed Orientable Surface F

The mapping class group of a surface F is the group of isotopy classes of orientation preserving self-homeomorphisms of F, modulo isotopy to the identity. It is generated by Dehn twists around the loops on F, shown in Fig. 1.4a), where a Dehn twist is a local twist, restricted to an annular neighborhood in F as seen in Fig. 1.4b). Here, we see the effect on the loop a of a twist along the circle b. This may be thought of as cutting the surface along the circle which we twist about, holding one of the resulting boundary components fixed, and reattaching the other, after rotating it through 2π.

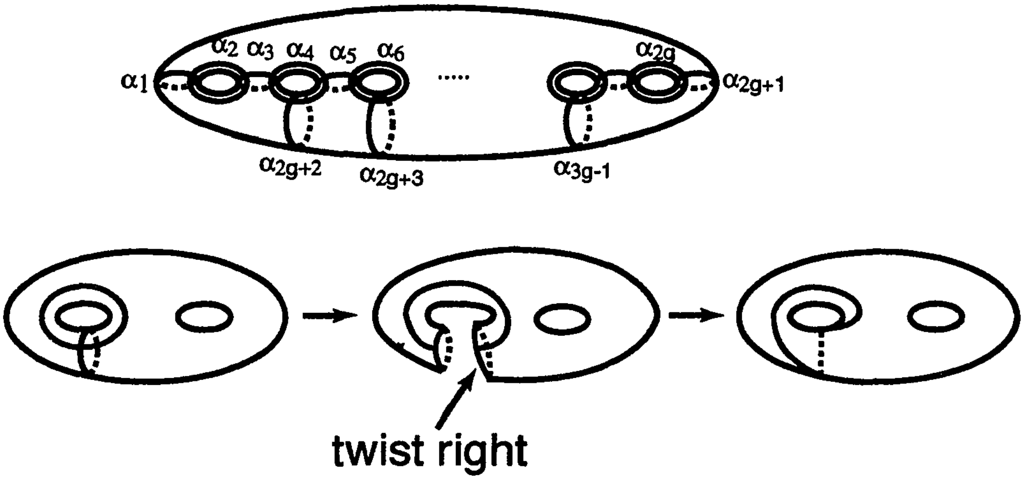


Fig. 1.4)

Any orientation preserving self-homeomorphism of F may be expressed as a composition of Dehn twists. The mapping class group of F, henceforth $MCG(F)$, plays a key role in the study of 3-manifolds and their Heegaard splittings. In particular, let us consider a surface with the meridians of handles highlighted as shown in Fig. 1.5 to be our initial surface.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

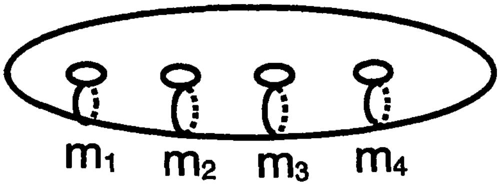


Fig. 1.5)

Any Heegaard splitting is a configuration of g disjoint, nonseparating, essential loops on the surface, and may be obtained by applying elements of $MCG(F)$ to an initial configuration. Thus, we may also consider the sequence of homeomorphisms which lead to a desired configuration of circles, as representing the Heegaard splitting. So a splitting may be specified by giving an element or product of generators in $MCG(F)$. This leads to a couple of essentially equivalent ways of viewing the totality of Heegaard splittings of manifolds, which possess splittings of genus g. Namely, we can think of the collection of splittings of genus g (for arbitrary manifolds) as a groupoid, where each configuration (each diagram) is an object, and the morphisms are elements of $MCG(F)$. In a similar fashion, we may take the collection of splittings of genus g to be Cayley graph of $MCG(F)$, where the vertex representing the identity element is the initial configuration above, (which happens to correspond to a connected sum of copies of $S^2 \times S^1$), and all other vertices (splittings) are reached by successive applications of the generators of $MCG(F)$ to this initial configuration. [Thurston and Hatcher] describe a simplicial complex that is similar. For a surface of genus g, we associate to the isotopy class of a loop on the surface, a vertex of the complex. By abuse of notation, allow a loop to represent its isotopy class. Then any n-tuple of disjoint loops forms an (n-1)-simplex and so a splitting of genus g, given by g loops as above, is represented by a simplex of maximum dimension. There is an obvious action of $MCG(F)$ on the complex. Since splittings of a given manifold M are highly nonunique, it is important to have some idea of which elements of $MCG(F)$ preserve the homeomorphism class of M. Borrowing notation and reference from [Birman], let $\mathcal{F}_g < MCG(F)$ be the subgroup of mapping classes whose representatives extend to H_1. Birman quotes a result of MacMillan which says these are precisely the elements $\phi \in MCG(F)$ such that for the induced map on homotopy, $\phi_*(N) = N$, where $N = << \textit{meridian loops} >> \triangleleft \pi_1(F)$. (Here <<>> denotes normal closure.) In particular, twists around meridians m_i and waist curves w_i represent elements of $\mathcal{F}_g$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2 Some Algebra: Quandles

2.1 Quandle Definitions and Proposition.

In what follows, I will be discussing an algebraic object associated to the collection of imbedded simple closed curves on a closed orientable surface F, of genus g. Actually, the algebraic object will be defined for isotopy classes of such curves, but by abusive notation, the elements wil be called variously, circles or loops, i.e. an isotopy class will be identified with its representative.

Let F be a closed orientable surface. Let

$$P = \{isotopy\,classes\,of\,unoriented\,,\,imbedded\,circles\,on\,F\}.$$

Let $MCG(F)$ be the mapping class group of F. MCG(F) acts on P, so there is a natural map $\mathbf{A} : MCG(F) \longrightarrow Aut(P)$. This action shall be written as a right action of $MCG(F)$ on P. We will show that $\{P\,,MCG(F)\}$ can be given an algebraic structure which captures many important geometric properties of the action of $MCG(F)$ and consequently, of Heegaard splittings.

Definition 2.1 *A* <u>*quandle*</u> *consists of a group G acting on the right on a set S, together with a pair of maps $\overline{}\rceil\,,\,\lceil\overline{}$: $S \rightarrow G$ for which the following relations hold:*

Q1) $a\overline{a}\rceil = a\lceil\overline{a} = a \qquad\qquad \forall\, a \in S$

Q2) $a\lceil\overline{b}\,\overline{b}\rceil = a\overline{b}\rceil\lceil\overline{b} = a \qquad\qquad \forall\, a,b \in S$

Q3) $x\lceil\overline{b}\,\overline{a}\rceil\overline{b}\rceil = \overline{a\overline{b}\rceil}\rceil \qquad\qquad \forall\, a,b,x \in S$

These relations, in the given order, are analogs of the Reidemeister moves I, II, III, for knots and links. The quandle encodes how the action of the group on the set, and the "actions" of the set back on the group are intertwined. The definitions and notions of a quandle used here are taken from [Kauffman]. He also gives examples of quandles, one of which is the quandle associated to a knot K and $\pi_1(S^3 - K)$, where the set S is the set of based lassos around strands of the knot, and the group is $\pi_1(S^3 - K)$. Another example is given by the action of $\pi_1(X)$ on $\pi_2(X)$, for a space X, where S consists of representative 2-spheres. I believe that quandles are also known as crossed G-sets for a group G acting on a set.

Define the maps $\overline{}\rceil$, $\lceil\overline{}$ such that

$$\overline{}\rceil, \lceil\overline{} : P \rightarrow MCG(F),$$

where

$$a \xmapsto{\rceil} \overline{a}\rceil$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and $\overline{a}\rceil$ is the right Dehn twist around the circle $a \in P$, while similarly,

$$a \xmapsto{\ulcorner} \lceil\overline{a}$$

and $\lceil\overline{a}$ is the left Dehn twist around the circle $a \in P$. With these definitions,

Proposition 2.1 *The quintuple* $\{P, MCG(F), \mathbf{A}, \overline{}\rceil, \lceil\overline{}\}$ *has the structure of a quandle.*

The proof will be given below. Note that if x is oriented, $x\overline{a}\rceil$ and $x\lceil\overline{a}$ inherit the orientation of x, since Dehn twists affect x only locally, where a is inserted.

I will give two proofs of the proposition, the second of which will make use of further properties of the quandle. I doing so, it will be convenient at times, to make use of another notation for the actions of $\overline{}\rceil$ and $\lceil\overline{}$, which will present these operations in a manner emphasizing a binary quality to the operation, so that for $x\overline{a}\rceil$, for instance, the roles of a as operator and x as operand are not visually distinguished as they are in the above notation. This suggests a symmetry of the roles of the operator and operand circles. However, this symmetry is only fully realized in certain special situations, as we shall see later. Nonetheless, with this caveat, the notation is harmless and shall be used where needed. Hence, let

$$x *_R b = x\overline{b}\rceil$$

and let

$$x *_L b = x\lceil\overline{b}$$

Pf. of Prop 2.1) Taking $\overline{a}\rceil$ and $\lceil\overline{b}$ to be Dehn twists gives the following interpretations of the three relations above:

Q1) Dehn twisting (right or left) around a circle a leaves a invariant as a set.

This is clear since we may isotope a copy of a off of itself, twist about it, and this has no effect on the original copy.

Q2) twisting left and twisting right about a circle a are inverses of one another. This is obvious.

Q3) If the result of a right twist around b on a circle a, is the circle c, then to perform a right twist around c;

1. take c to a via left twist on b $\qquad \lceil\overline{b}$
2. twist right, around a $\qquad \overline{a}\rceil$
3. take the result back to c via right twist on b $\qquad \overline{b}\rceil$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In other words, conjugation takes a twist around one circle a to a twist around its image c. (Image under some other twist $\overline{b}\rceil$) ///

Alternatively we may prove Q3) in a way that exhibits certain other important quandle properties. We need the following lemma.

Lemma 2.1 *$\urcorner$ and $\ulcorner$ distribute over $*_R$ and $*_L$. i.e. $(x *_R a)\overline{b}\rceil = x\overline{b}\rceil *_R a\overline{b}\rceil$, for a, $b \in P$.*

Pf. Let $a, b, x \in P$. Let $|x \cap a| = L$, $|x \cap b| = M$, and $|a \cap b| = N$, where L, N, and M are minimal with respect to the isotopy classes of a, b, and x. We may always think of "pulling the loops tight" and assume the intersections are transverse. Then performing Dehn twists may be thought of as inserting a requisite number of copies of the operating circle into the operand circle, with deference to the sense, L or R, of the twist.

Thus $x\overline{a}\rceil$ = (x with L copies of a inserted into it) See Fig. 2.1) for an example.

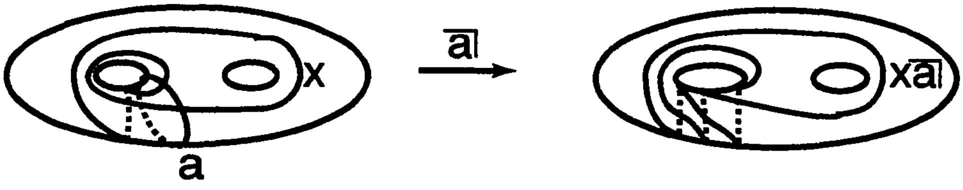


Fig. 2.1)

We have $(x *_R a)\overline{b}\rceil = x\overline{a}\rceil\overline{b}\rceil =$ x with L copies of a inserted into it and in each such a component, insert N copies of b, and additionally insert M copies of b in the x component.

So we have x with M copies of b in it, L copies of a in it, each of which has N copies of b in it.

Now consider $x\overline{b}\rceil *_R a\overline{b}\rceil$.

Here, $x\overline{b}\rceil$ is x with M copies of b in it.
$a\overline{b}\rceil$ is a with N copies of b in it.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and the total collection, corresponding to $*_R$, consists of

x with M copies of b in it
L copies of a in it (from $*_R$)
each with N copies of b in it.

Had we worked with $*_L$ and $\rceil$ or $\lceil$, similar numerology would have occurred, the only difference being in the sense, L or R, of the inserted piece. In the calculation above, both for

$$(x *_R a)\overline{b}\rceil$$

and for

$$x\overline{b}\rceil *_R a\overline{b}\rceil,$$

since the inserted circles are placed in the receiving circles in the same place, positionwise ("reading" along the receiving circle) with the same sense, L or R, we then have

$$(x *_R a)\overline{b}\rceil = x\overline{b}\rceil *_R a\overline{b}\rceil.$$

Again, a similar argument holds for any combination of $*_R$, $*_L$, and $\rceil$ or $\lceil$.

///

So the operation of Dehn twisting distributes over itself.

Alternate Pf. of Proposition 2.1 As we saw above, the fact that Q1) and Q2) hold for circles and Dehn twists is clear. So we consider an alternate proof of Q3). We need to see that

$$x\lceil\overline{b}\ \overline{a}\rceil\overline{b}\rceil = \overline{a\overline{b}\rceil}\rceil.$$

But

$$\begin{aligned} x\lceil\overline{b}\ \overline{a}\rceil\overline{b}\rceil &= (x\lceil\overline{b} *_R a)\overline{b}\rceil && \text{by notation} \\ &= (x\lceil\overline{b}\ \overline{b}\rceil) *_R (a\overline{b}\rceil) && \text{by Lemma 2.2 (distributivity)} \\ &= x *_R (a\overline{b}\rceil) && \text{by Q2) (inverses)} \\ &= x\ \overline{a\overline{b}\rceil}\rceil && \text{by notation.} \end{aligned}$$

Again, the argument holds for all possible combinations of $*_R$, $*_L$, $\rceil$ and $\lceil$.
Thus $\{P, MCG(F), \mathbf{A}, \rceil, \lceil\}$ fulfills all the conditions for quandlehood. ///

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The quandle structure on the collection of circles on the surface F is part of a larger (quasi) algebraic structure. In this context, the quandle may be thought of as the "multiplicative" part of the overall structure. The quotes, " ", on multiplicative, are meant to suggest that to a certain extent, what we consider multiplication on P, will depend on the context. There are further operations that may be imposed on P, to form the "extended quandle".

2.2 Connected Sums, $\#_\gamma$

Consider a family of binary operations $\#_\gamma$ on the circles in P. These operations will be connected sums of circles a, $b \in P$ where $a \cap b = \emptyset$, and where the path γ along which we take the connected sum, is imbedded, and meets each of a and b in exactly one point. It is clear that the (isotopy class of the) resulting circle is independent of representative of the isotopy class of γ which is used to form the connected sum. So these connected sums are parametrized by, and depend only on the isotopy classes of paths γ on the surface F. Notice also that the path γ is taken relative to the ending circles a and b, rather than to specific endpoints on those circles, since we may slide the endpoints around the circles, and the resulting connected sums will be in the same isotopy class, because we may "unwind" the path by rotating the end circles as necessary. See Fig. 2.2)

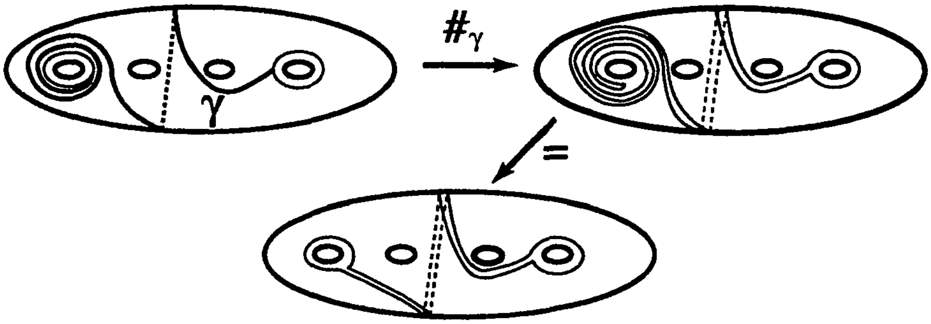


Fig. 2.2)

2.3 Further Algebraic Operations and Characteristics

Let $\mathbf{t}$ denote the isotopy class of a homotopically trivial circle on F, and let t be a representative. Then

$$a\overline{t}\rceil = a\lceil\overline{t} = a \qquad \forall t \in \mathbf{t} \; and \; \forall a \in P,$$

since any such t may be isotoped off a, and hence has no effect on a. Clearly also

$$a \#_\gamma t = a$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

for all paths γ connecting a and t. Thus t plays the role of the unit element with respect to both $\overline{\ \ }\rceil$ and $\lceil\overline{\ \ }$ and the connected sum $\#$.

We saw that $a\lceil\overline{b}\ \overline{b}\rceil = a\overline{b}\rceil\lceil\overline{b} = a \;\; \forall a, b \in P$ by Q2). Now if a is a circle with a given orientation, then let a^{-1} denote the same circle with the opposite orientation. If we form a tubular neighborhood of a, i.e. the cobordism $a \times [-1, 1]$, then the two boundary copies of a inherit opposite orientations. We may form the connected sum of one copy of a with the oppositely oriented a^{-1}, along a fiber γ, (of the product neighborhood) . We see that $a \#_\gamma a^{-1} = t$. Note that the short path γ chosen for this connected sum is crucial, for if we were to use another path, the result would not be the homotopically trivial circle. As a matter of fact, we may form a separating circle by summing a nonseparating circle a, and its inverse a^{-1}, along another path. In general, when forming connected sums of circles with orientation, we will assume that the orientations of the circles are compatible. See Fig. 2.3)

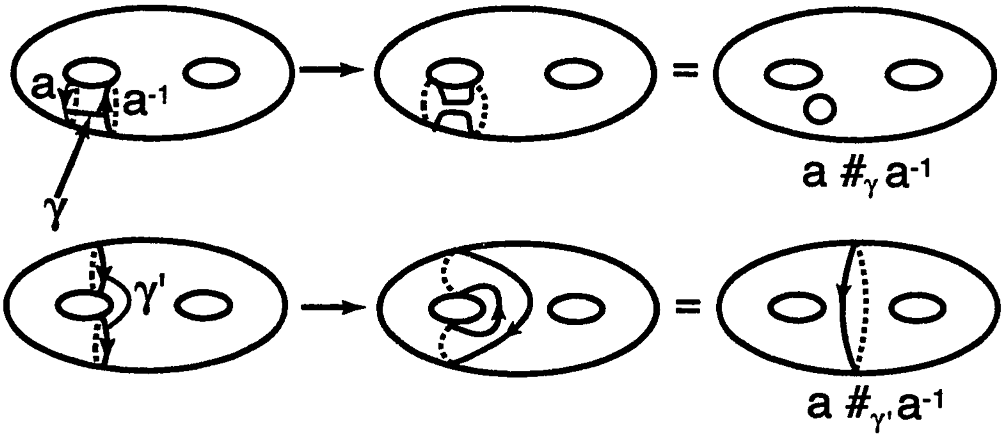


Fig. 2.3)

According to Lickorish, the set of circles $\{\alpha_i \,|\, i = 1...3g-1\}$ (see Fig. 1.4a)) generates $MCG(F)$ as a group, via Dehn twists about these circles. It is clear from the foregoing that any circle in P may be obtained from the operations of Dehn twisting about the α_i, and the taking of connected sums, as described. Since even "basic" separating circles may be obtained as above, and all others are homeomorphic images of these, the extended quandle is thus generated . We saw above (Lemma 2.1) that Dehn twists distribute over each other. There is also a "distributivity" of Dehn twists with respect to connected sums. More specifically,if a , b , $c \in P$, $a \cap b = \emptyset$, and γ is a path as above, then

$$(a \#_\gamma b)\overline{c}\rceil = a\overline{c}\rceil \#_{\gamma\overline{c}\rceil} b\overline{c}\rceil$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Here, $\gamma\overline{c}\rceil$ is the image of γ under the homeomorphism $\overline{c}\rceil$. The equation above makes sense precisely because $\overline{c}\rceil$ is a homeomorphism of all of F, and the right hand side merely records its effects on the individual summands and on the connecting path, separately. An example is given below in Fig. 2.4).

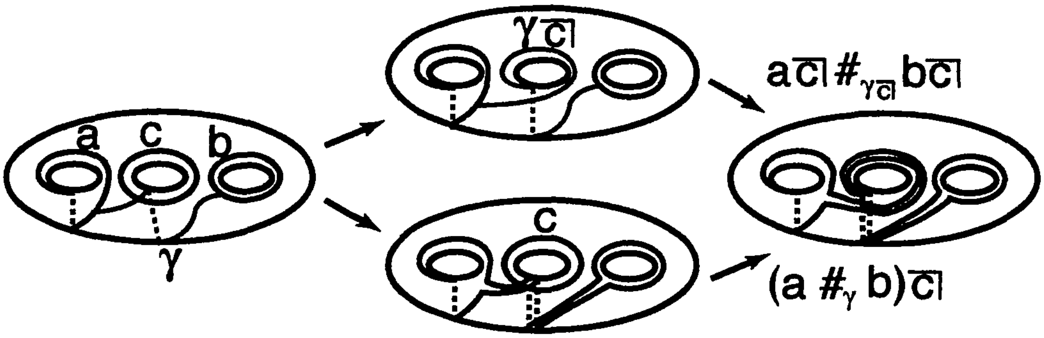


Fig. 2.4)

Now, let K be a commutative ring with unit, or a field, and form the free module $P_K = K[P]$ over K. So typical elements of P_K are formal sums of isotopy classes of loops on F with coefficients in K. Consider an operation $\Delta : P \longrightarrow P \otimes P$ defined on a loop $a \in P$ by taking a pair of disjoint copies of a, with the same orientation, so that $\Delta(a) = a \otimes a$. Define an operation; $\mathcal{A} : P \longrightarrow P$, so that $\mathcal{A}(a) = a^{-1}$ where a^{-1} is the loop a on F with the opposite orientation. We also insist that $\mathcal{A} : *_R \leftrightarrow *_L$ and that it be an antimorphism. Then, $\mathcal{A}$ inverts elements in $MCG(F)$ and thus affects the quandle appropriately. We also consider another operation similar to Δ, called ν_γ, where γ is an arc in F with both of its endpoints on the same side of $a \in P$ and $\nu_\gamma : P \longrightarrow P \otimes P$ is defined by letting $\nu_\gamma(a)$ be the pair of components of the boundary of a regular neighborhood of $a \cup \gamma$, neither of which is isotopic to the original loop a. Sometimes, I will refer to this operation as "self-connected- sum". Notice that if the path γ, with endpoints on $a \in P$, is homotopically trivial rel. endpoints, then one of the resulting circles is in the class of the trivial loop t on F, and the other is isotopic to the original loop $a \in P$. Fig. 2.5 shows examples of applications of these operations.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

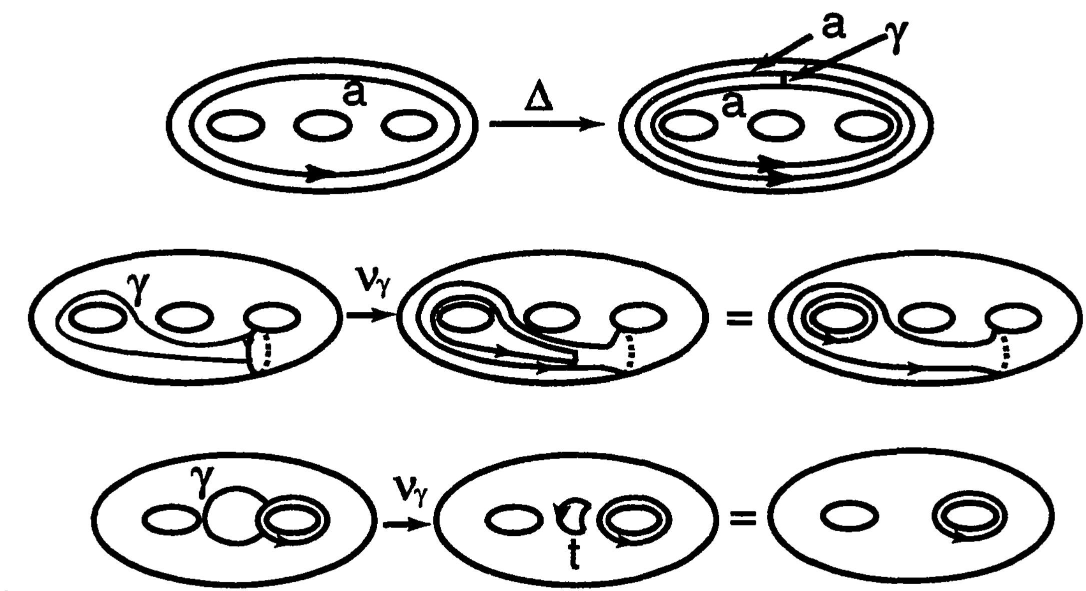


Fig. 2.5)

If we choose a particular path γ along which to form a connected sum of elements $a\,,\,b \in P$ we may think of the connected sum as a “multiplication” on P, so that $\#_\gamma\,(a\otimes b)\to a\,\#_\gamma\, b$. We then have the following interactions of the above operations. Let γ be the (isotopically unique) short path used to define the parallel copy of $a \in P$. Then

$$\#_\gamma\,(I_P\otimes\mathcal{A})\,\Delta\,(a) \;=\; \#_\gamma\,(\mathcal{A}\otimes I_P)\,\Delta\,(a) \;=\; \mathbf{t}$$

where I_P is the identity morphism in P. Diagramatically,

$$\begin{tikzcd}
& P\otimes P \arrow[r, "I_P\otimes\mathcal{A}"'] & P\otimes P \arrow[dr, "\#_\gamma"'] & \\
P \arrow[ur, "\Delta"'] \arrow[dr, "\Delta"] & & & P \\
& P\otimes P \arrow[r, "\mathcal{A}\otimes I_P"] & P\otimes P \arrow[ur, "\#_\gamma"] &
\end{tikzcd}$$

Similarly, if we again specify a particular path γ along which to form the self-connected-sum , we have

$$\#_\tau\,(I_P\otimes\mathcal{A})\,\nu_\gamma\,(a) \;=\; \#_\tau\,(\mathcal{A}\otimes I_P\,)\,\nu_\gamma\,(a) \;=\; a$$

Here, τ is the (isotopically unique)“shortest path” joining the two components of $\nu_\gamma\,(a)$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Diagramatically, we have:

$$\begin{tikzcd}
 & P\otimes P \arrow[r, "I_P\otimes A"'] & P\otimes P \arrow[dr, "\#_\gamma"] & \\
P \arrow[ur, "\nu_\gamma"'] \arrow[dr, "\nu_\gamma"] & & & P \\
 & P\otimes P \arrow[r, "A\otimes I_P"] & P\otimes P \arrow[ur, "\#_\gamma"'] &
\end{tikzcd}$$

Fig. 2.6) gives examples of the above relations.

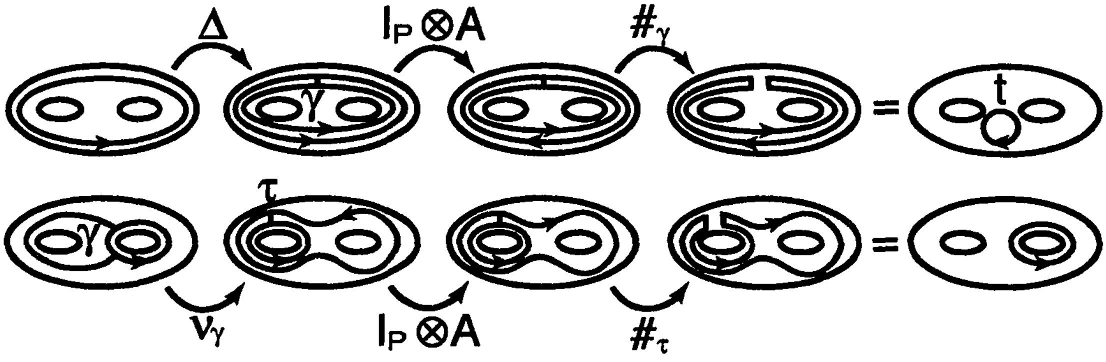


Fig. 2.6)

Δ also has the property that it behaves like a morphism with respect to $\#_\gamma$, i.e. for any $a\,,\,b \in P$, for which $\#_\gamma$ is defined,

$$\Delta(a\#_\gamma b) = \Delta(a)\#_\gamma\Delta(b).$$

Here, the notation "γ" on the right side of the equation implies taking the connected sums respectively of the "outer" and "inner" pairs of circles formed by the operation Δ on a and b. We sum along the path γ, to connect the "outer" pair. Then we sum along a parallel copy of γ, which has been extended by the unique "fibers" of the Δ construction (see the beginning of section 2.3), for each of a and b, to connect the "inner" pair. See Fig. 2.7)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

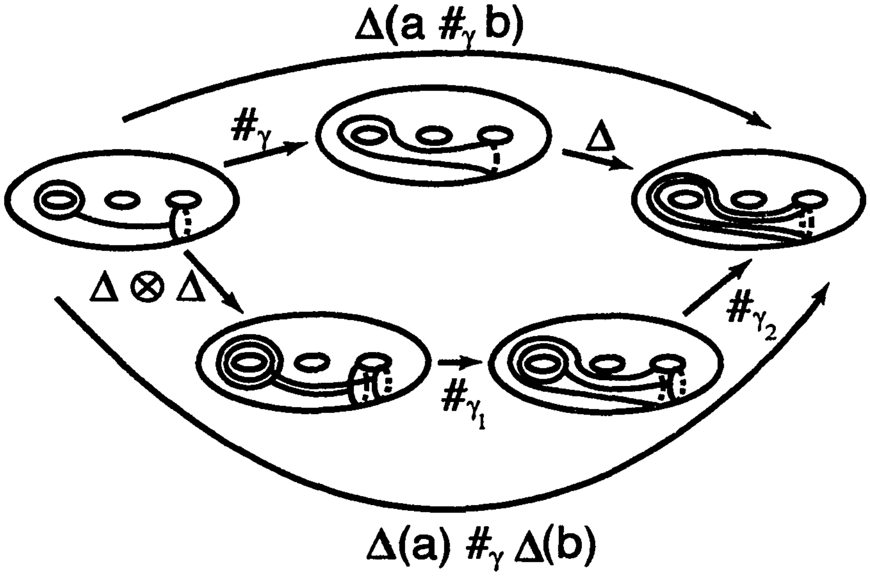


Fig. 2.7)

With the operations $\#_\gamma$ (as "multiplication"), $\mathcal{A}$ (as antipode), and Δ (as comultiplication), defined above, P comes very close to being an involutary Hopf algebra, or some generalization thereof. However, we lack uniqueness in the definition of the "multiplication" $\#_\gamma$, since there is no unique path between nonintersecting elements $a\,,\, b \in P$, along which to form a connected sum. Notice however, that the issue of uniqueness of connected sum did not arise in the relations above. There, a unique "shortest" path exists and is well defined. If some notion of canonical path between nonintersecting elements $a\,,\, b \in P$ could be found, then the structure just defined would reduce to a Hopf algebra, assuming the use of this path for the multiplication.

3 Some Relations and Consequences in the Extended Quandle

We now look at some consequences that arise from the definitions above, among them, certain basic relations. The following are basic facts regarding Dehn twists.

Lemma 3.1 *Let $a\,,\, b \in P$ and suppose a and b are representatives of distinct isotopy classes. Then $a\,\overline{b}\rceil = b\lceil\overline{a}$ iff $|a \cap b| = 1$.*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Pf. Since $|a \cap b| = 1$, a and b must be nonseparating. Then as we have seen above, $a\overline{b}\rceil$ is a copy of a with a copy of b inserted in it, where a turns to the right when it encounters b; similarly $b\lceil\overline{a}$ is a copy of b with a inserted in it, such that b turns to the left when it encounters a, and the resulting loops are the same. See Fig. 3.1 for an example.

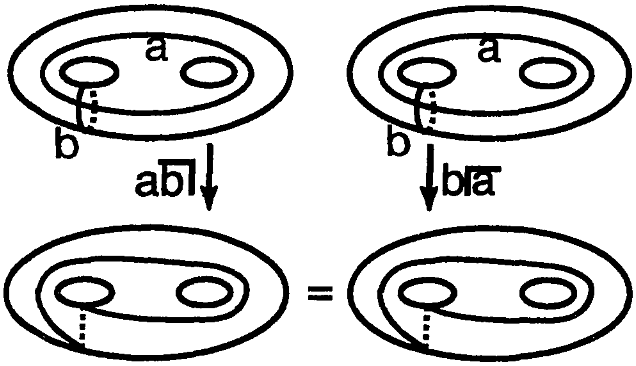


Fig. 3.1)

So if $|a \cap b| = 1$, then $a\overline{b}\rceil = b\lceil\overline{a}$.

Now suppose that $a\overline{b}\rceil = b\lceil\overline{a}$, and $|a \cap b| \neq 1$. Let $\{a\}$, $\{b\}$ denote the homology classes of a and b respectively. If $\{a\} = \{b\}$, then a and b cobound a subsurface of of genus ≥ 1, since we have assumed a and b are in different isotopy classes. But then $a \cap b = \emptyset$, and so $a\overline{b}\rceil = a$ and $b\lceil\overline{a} = b \Rightarrow a = b$, since we assumed $a\overline{b}\rceil = b\lceil\overline{a}$. This contradicts the assumption that a and b were in distinct isotopy classes. Hence we may take $\{a\} \neq \{b\}$. So suppose now that $|a \cap b| \geq 2$. Then

$a\overline{b}\rceil = a$ with n copies of b inserted in it, so
homologically it looks like $a + nb$

while,

$b\lceil\overline{a} = b$ with n copies of a inserted in it, so
homologically it looks like $b + na$

If we now consider a homology basis for F in which

$$\{a\} = (\alpha_1, \ldots, \alpha_{2g+2}) \; and \; \{b\} = (\beta_1, \ldots, \beta_{2g+2})$$

we would have

$$\alpha_i + n\beta_i = \beta_i + n\alpha_i$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

where α_i , $\beta_i \in \mathbf{Z}$, $i = 1 ... 2g-2$

$$\begin{aligned} \Rightarrow &\quad (n-1)\beta_i = (n-1)\alpha_i \\ \Rightarrow &\quad \beta_i = \alpha_i \qquad \forall i \qquad (1) \\ \Rightarrow &\quad \{\alpha\} = \{\beta\} \end{aligned}$$

which contradicts our initial assumption that $\{a\} \neq \{b\}$.

So $a\overline{b}\rceil = b\lceil\overline{a} \;\Rightarrow\; |a \cap b| = 1$. ///

We also have the following small lemma, which is similar.

Lemma 3.2 *If a , $b \in P$ and they are in different isotopy classes, then $a\overline{b}\rceil = a$ and $b\lceil\overline{a} = b$ iff $|a \cap b| = 0$.*

Pf. We may assume again that a and b are nonseparating, and that $|a \cap b|$ is minimal in the isotopy classes of a and b. Then as mentioned above, if $|a \cap b| = 0$, $\overline{b}\rceil$ does not affect a, and $\lceil\overline{a}$ does not affect b. Conversely, if $|a \cap b| = n \neq 0$, then $a\overline{b}\rceil$ yields a with n copies of b inserted into it, which is not isotopic to a. Similarly for $b\lceil\overline{a}$, hence the result.

We will use these, and the other quandle relations to prove a slight generalization of the "braid relation" in $MCG(F)$. According to Lickorish, $MCG(F)$ is generated by the Dehn twists about the 3g-1 curves shown on F in Fig. 1.4a). In $MCG(F)$ there exists a relation $\alpha_i\,\alpha_{i+1}\,\alpha_i = \alpha_{i+1}\,\alpha_i\,\alpha_{i+1}$, where by abuse of notation, α_i is taken to mean the Dehn twist about that curve. Notice that as shown, $|\alpha_i \cap \alpha_{i+1}| = 1$ for all i=1...3g-1. More generally,

Proposition 3.1 *Suppose a , $b \in P$ and $|a \cap b| = 1$. Then $\overline{a}\rceil\overline{b}\rceil\overline{a}\rceil = \overline{b}\rceil\overline{a}\rceil\overline{b}\rceil$*

Pf.

$$\begin{aligned} \lceil\overline{b}\;\overline{a}\rceil\overline{b}\rceil &= \overline{a\overline{b}\rceil}\rceil && \text{by Q3)} \\ &= \overline{b\lceil\overline{a}}\rceil && \text{by Lemma 3.1 above, since } |a \cap b| = 1 \\ &= \overline{a}\rceil\overline{b}\rceil\lceil\overline{a} && \text{by Q3) variant,} \end{aligned}$$

$$\begin{aligned} \text{so} &\qquad \lceil\overline{b}\;\overline{a}\rceil\overline{b}\rceil = \overline{a}\rceil\overline{b}\rceil\lceil\overline{a} \\ \text{thus by applying } \overline{a}\rceil, &\qquad \lceil\overline{b}\;\overline{a}\rceil\overline{b}\rceil\overline{a}\rceil = \overline{a}\rceil\overline{b}\rceil \\ \text{and by applying } \overline{b}\rceil, &\qquad \overline{a}\rceil\overline{b}\rceil\overline{a}\rceil = \overline{b}\rceil\overline{a}\rceil\overline{b}\rceil \end{aligned}$$

so in particular $\alpha_i\,\alpha_{i+1}\,\alpha_i = \alpha_{i+1}\,\alpha_i\,\alpha_{i+1}$. ///

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

While on this subject, it is also worth noting a connection between the statement of Lemma 3.1a) and the behavior of the "antipode" $\mathcal{A}$. If $|a \cap b| = 1$, then

$$\begin{aligned}
\mathcal{A}(a\overline{b}\rceil) &= \mathcal{A}(a *_R b) && \text{by notation} \\
&= \mathcal{A}(b) *_L \mathcal{A}(a) && \text{by definition of A} \\
&= b^{-1} *_L a^{-1} && \\
&= b^{-1}\lceil\overline{a^{-1}} && \text{by notation} \\
&= b\lceil\overline{a} && \text{as circles, since } \overline{}\rceil\,,\ \lceil\overline{} \text{ do not notice orientations.}
\end{aligned}$$

So $\mathcal{A}$ acts as the identity on circles of the form $a\overline{b}\rceil$ where a, b are as above. There also exists another relationship between $\#$ and the comultiplication-like ν. Recall that P_K is the free vector space over the field K, generated by elements of P, and that $\otimes$ is defined.

Proposition 3.2 *Let a and b be disjoint nonseparating circles in P_K. Choose a path γ joining them. Then there exists a canonical path τ (up to isotopy) such that*

$$\nu_\tau \#_\gamma (a \otimes b) = (a \otimes b).$$

Conversely, given a circle c and a path τ joining c to itself (intersecting c on the same side of c at both endpoints), there exists a canonical path γ (up to isotopy) such that

$$\#_\gamma \nu_\tau (c) = c,$$

so $\#_\gamma$ and ν_τ are "inverses" of one another.

Pf. We may consider the formation of $\#_\gamma (a \otimes b) = a \#_\gamma b$ as being accomplished by "tubing together" a and b along γ. Then we may choose the path τ to be the short path (unique up to isotopy) running from one side of the "tube" along γ to the other, which crosses γ once, transversely. Then the result of applying ν_τ is $a \otimes b (= a \cup b)$ again. The converse holds similarly. See Fig. 3.2) ///

We shall see instances of this and of related behavior, when we look at situations involving "waves" and connected sums later on.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

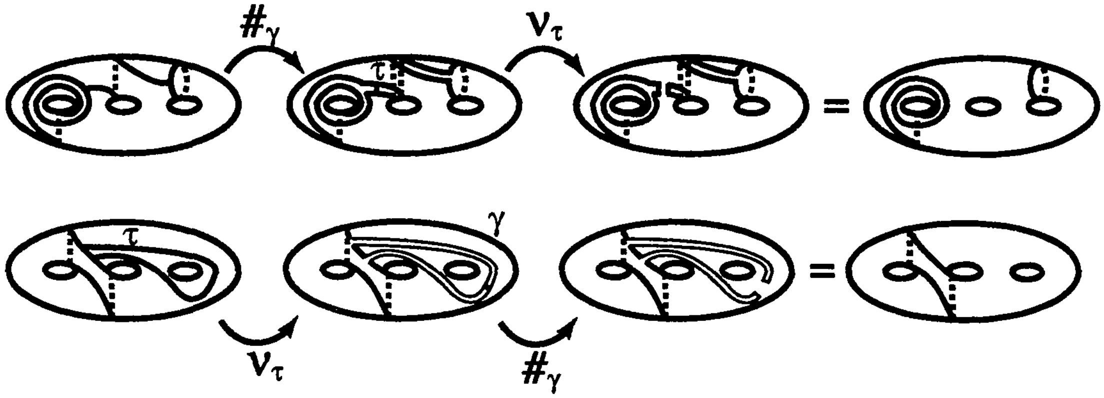


Fig. 3.2)

There are also certain relatively simple relations that arise from the interaction of the operations of Dehn twist and connected sum, when the intersection numbers of the twisting circles with the twistees are small (0 or 1). We have already seen one interaction in the form of the "distributivity" of Dehn twists with respect to connected sum. Here, we will look at simple instances of twisting about a circle that is a connected sum, and the effect on other circles. Throughout the following discussion, suppose that a, b, $c \in P$ and that $|b \cap c| = 0$, so that $b \#_\gamma c$ is defined, for some path γ joining b and c. Furthermore, suppose that γ is chosen so that it may be homotoped off a if necessary. The situation will be 'graded' by intersection numbers.

Case 1)

Assume $|a \cap b| = |a \cap c| = 0$ (again we are taking $|b \cap c| = 0$).

Then

$$|(b \#_\gamma c) \cap a| = 0 \text{ so } a\,\big|\overline{(b \#_\gamma c)} = a\,\overline{(b \#_\gamma c)}\big| = a.$$

See Fig. 3.3)

Case 2)

Assume $|a \cap b| = 1$ and $|a \cap c| = 0$

Here, $a\,\big|\overline{(b \#_\gamma c)}$ inserts a copy of $b \#_\gamma c$ into a point $x = a \cap b$. However since $|a \cap c| = 0$, and γ does not meet a, this is the same as inserting a copy of b into a at x, and then forming the connected sum with c. Hence,

$$a\,\big|\overline{(b \#_\gamma c)} = a\,\big|\overline{b}\ \#_\gamma c$$

and similarly for the twist $a\,\overline{(b \#_\gamma c)}\big|$. Again see Fig. 3.3) for an example.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

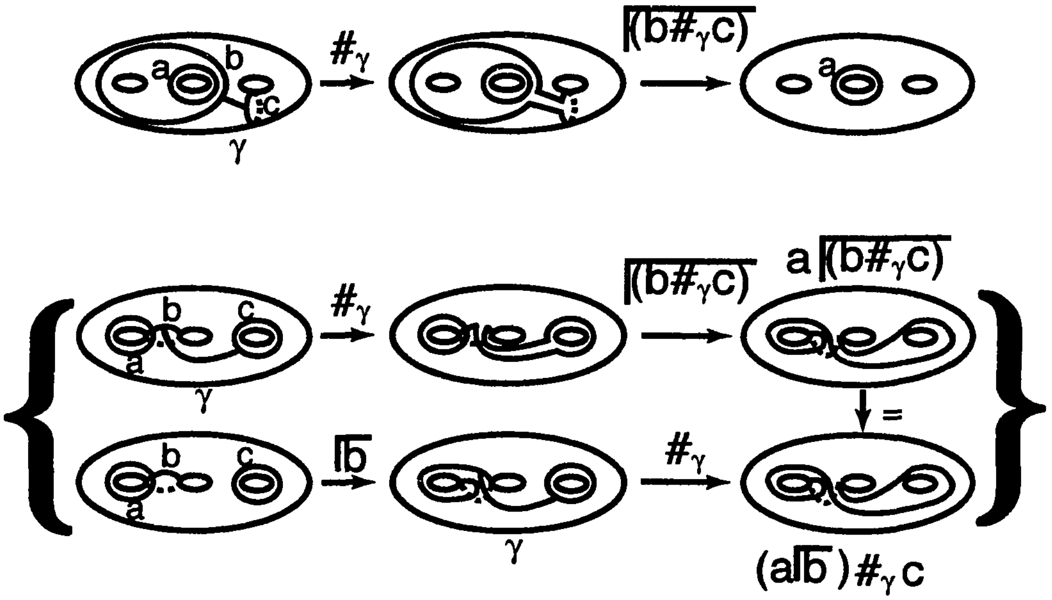


Fig. 3.3)

We may allow one of the intersection numbers, say $|a \cap b|$, to increase arbitrarily, while holding the other one at 0, and the situation is still manageable in that a simple description of the result may be given.

Before the next case, let us introduce some new terminology and notation. In particular, let $\Delta_n : P \longrightarrow P^{\otimes n}$ be defined by

$$\Delta_n = (\underbrace{I \otimes I \otimes \cdots \otimes I}_{n-2} \otimes \Delta)(\underbrace{I \otimes I \otimes \cdots \otimes I}_{n-3} \otimes \Delta) \ldots (I \otimes I \otimes \Delta)(I \otimes \Delta)\Delta.$$

Case 3)

Let $|a \cap b| = n > 1$, and let $|a \cap c| = 0$. Now $a\overline{|(b\#_\gamma c)}$ inserts n copies of $b\#_\gamma c$ into a. But consider $a\overline{|b}$. This is a with n copies of b inserted in it. Consecutive (adjacent) "strands" of $a\overline{|b}$ come from different copies of b, i.e., from copies arising from distinct adjacent points of $a \cap b$. Now form connected sums with c, repeatedly, along paths isotopic to γ, for each of the n strands of the "b-portion" of $a\overline{|b}$. Then each of the n copies of b is now connect-summed with a copy of c, and the result is the same as $a\overline{|(b\#_\gamma c)}$. So in general, if $|a \cap b| = n > 1$ and $|a \cap c| = 0$, and if (the isotopy class of) γ joining b and c is chosen, then

$$a\overline{|(b\#_\gamma c)} = (\ldots((a\overline{|b})\#c)\#c\ldots)\#c)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

where $\# c$ occurs n times in the right hand side of the above equation and the repeated appearance of γ each time has been suppressed. Using the notation developed above, we might write this alternatively as

$$a \lceil\overline{(b \#_\gamma c)} = (a\lceil\overline{b}\,)\#_{n\gamma}\, \Delta_n(c)$$

where $\#_{n\gamma}$ means repeated summing along n parallel copies of γ. Similarly for $a\,\overline{(b \#_\gamma c)}\rceil$. See Fig. 3.4) for an example.

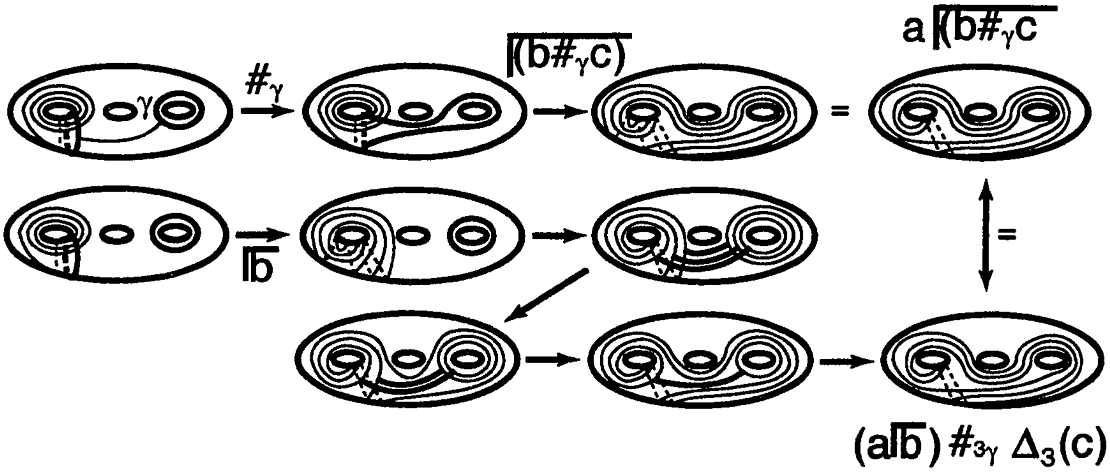


Fig. 3.4)

In the previous three cases, we held $|a \cap c| = 0$, while $|a \cap b|$ was allowed to vary. If we simultaneously were to allow $|a \cap c|$ to vary, the situation would become far more involved, depending to a much greater extent on the path γ chosen, and a general characterization as above seems much less likely. See Fig. 3.5) for a simple comparative example.

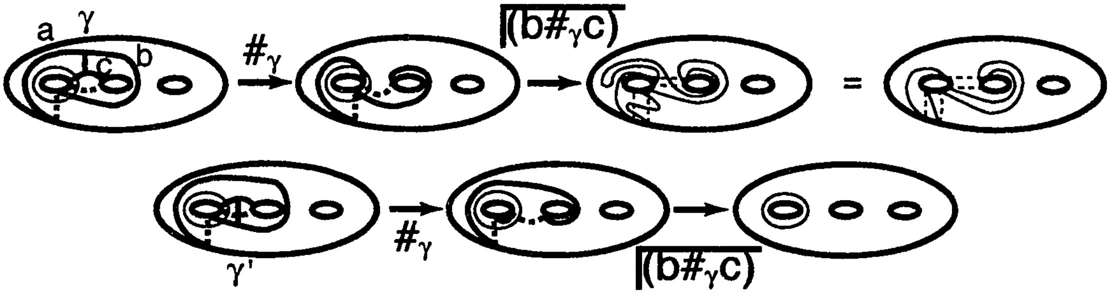


Fig. 3.5)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Note in the cases 1) and 2) of Fig. 3.5), $|a \cap b| = 1$ and $|a \cap c| = 1$, but different γ and γ' were used. $\overline{|(b \#_{\gamma} c)}$ acts nontrivially in 1), while $\overline{|(b \#_{\gamma'} c}$) acts trivially in 2), so characterization depends on more than just $|a \cap b|$ and $|a \cap c|$.

The next result examines the behavior of certain iterated connected sums. It is somewhat of interest in that it concludes that a strictly algebraic result is equivalent to a strictly topological one. We consider three mutually disjoint circles a, b, $c \in P$, an let γ_{ij} be paths joining appropriate pairs of circles, where $i, j \in \{a, b, c\}$, and $i \neq j$. We also assume that each such γ_{ij} may be homotoped so as to meet the i-th and j-th circles in a single point apiece, and that it may be homotoped to be disjoint from other circles and paths. To simplify notation, I will write $i \#_{\gamma_{ij}} j$ simply as γ_{ij}, where i and j are as above. So for instance, a 2-fold connected sum might look like

$$(i \#_{\gamma_{ij}} j) \#_{\gamma_{ik}} k \doteq \gamma_{ij}\gamma_{ik},$$

where $i, j, k \in \{a, b, c\}$ and $i \neq j \neq k$. This is meant to imply that first we connect i to j via γ_{ij}, and then we take the result and connect its i-portion to k via γ_{ik}. There are six such 2-fold sums possible from the above description. Due to the commutativity of $\#$, pairs of them are equivalent (isotopic). For instance $\gamma_{ab}\gamma_{bc} = \gamma_{bc}\gamma_{ab}$. See Fig. 3.6a) It is also worth mentioning a slightly subtle point here. If we happen to choose a pair of paths that approach a given circle from opposite sides, then it is not sufficient to merely specify their isotopy classes in order to determine the resulting 2-fold connected sum. We must also specify the positions of their endpoints with respect to one another on the circle. However this situation will not arise in the following theorem, as as we will show.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

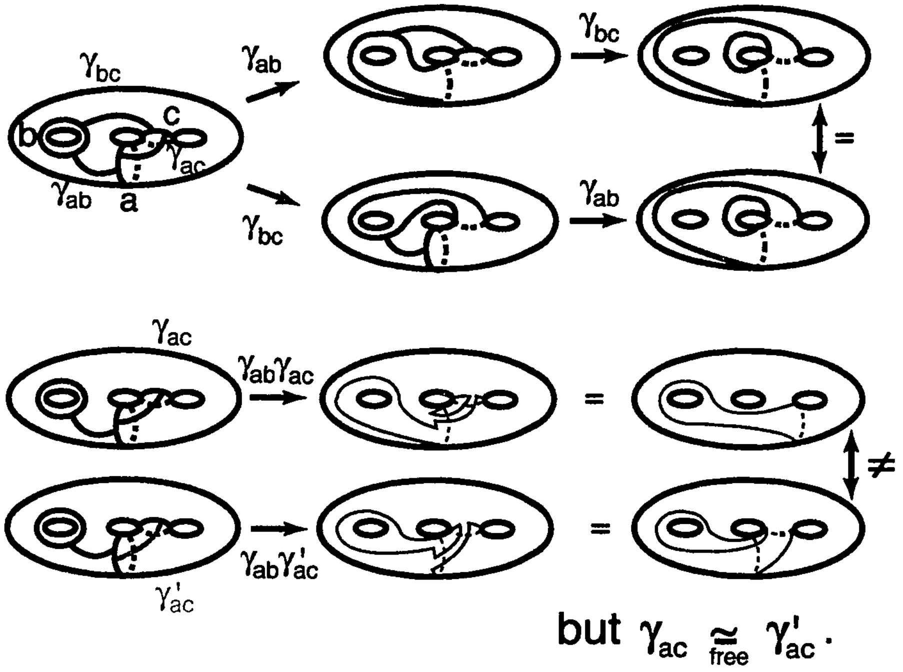


Fig. 3.6)

Notice that in Fig. 3.6b), the paths chosen on each side of the circle in each case represent the same isotopy class, but their relative placements on the circle give rise to different results.

Theorem 3.1 *Let* a , b , $c \in P$ *be mutually disjoint circles on* F, *and let* γ_{ab} , γ_{bc}, *and* γ_{ac} *be three fixed paths joining pairs of them, as above. Then the three possibly distinct 2-fold sums are in fact equivalent (isotopic) to one another, iff a circuit formed by the union of the paths and portions of the circles a, b, c bounds a disk in* F.

Pf. Suppose the union of the paths and portions of the circles does bound a disk in F. Any 2-fold connected sum $\gamma_{ij}\gamma_{ik}$ may be thought of as the component of the boundary of the regular neighborhood N of the union

$$i \cup j \cup k \cup \gamma_{ij} \cup \gamma_{ik}$$

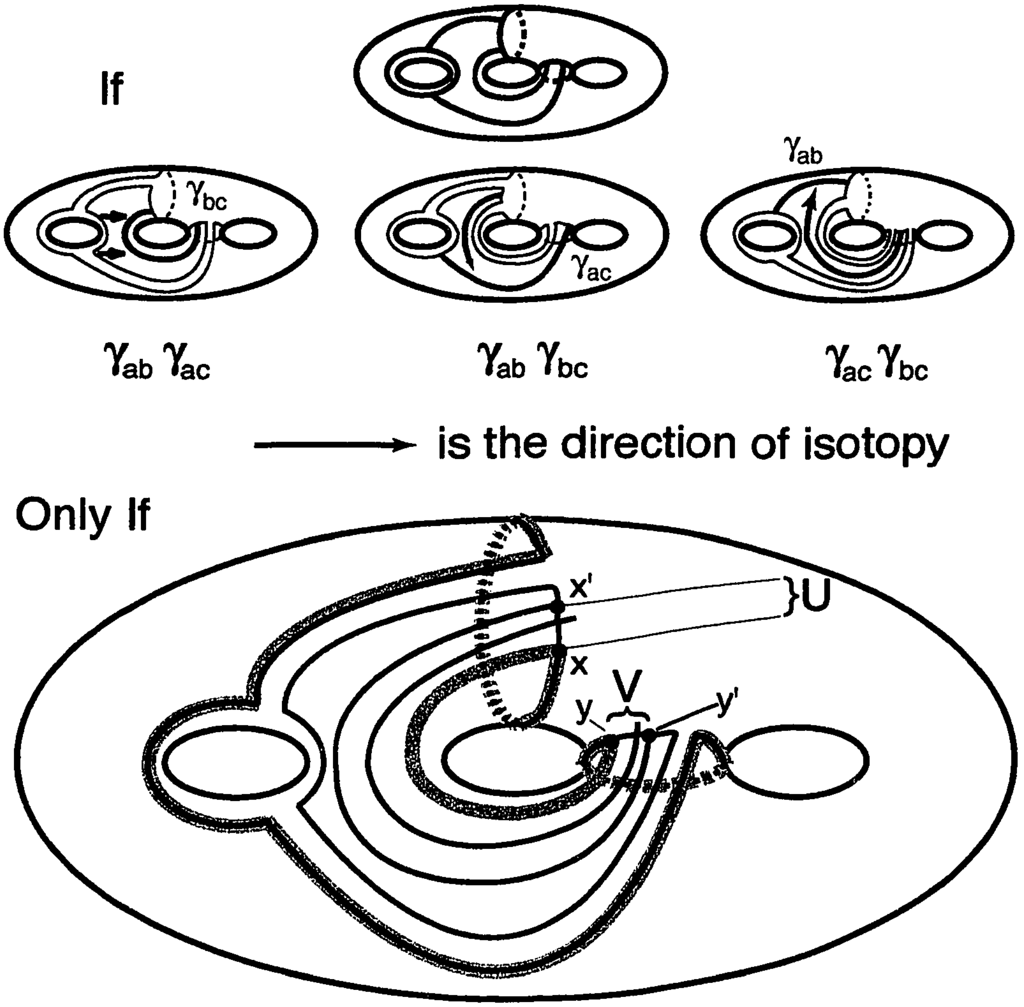
which is not isotopic to any of i , j , k. I will use the term "tube arc" to refer to a piece of the boundary of N arising from one of the "tubed" paths γ. Similarly, a "circle arc" will denote a portion of the boundary of $N(i \cup j \cup k \cup \gamma_{ij} \cup \gamma_{ik})$ arising from one of the circles. The portion of any such 2-fold sum which abuts the disk (as per assumption) is comprised of 3 circle arcs and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2 tube arcs. Each tube arc meets two circle arcs in one point apiece, and the totality of the 5 arcs mentioned (abutting the disk) is itself an arc. Call this arc d. The only arc not represented in d is the tube arc arising from γ_{jk}, the unused path in the connected sum. It also abuts the disk. Hence, we may isotope d to γ_{jk} across the disk. But we may do this for any of the three 2-fold connected sums. Since in each case, d was isotoped to the tube arc that was "missing" in that sum but present in the others, and since a given sum coincides with another sum away from the arcs which are missing from each, the results of the isotopies in all cases are the same. See Fig. 3.7).

Conversely, suppose the three sums $\gamma_{ab}\gamma_{ac}$, $\gamma_{ab}\gamma_{bc}$, and $\gamma_{ac}\gamma_{bc}$ are isotopic. Again, we may consider each sum as being obtained by tubing together the circles a, b, and c via tubes $\gamma_{ij} \times S^0$. So each sum is again a collection of tube arcs and circle arcs, and a given connected sum $\gamma_{ij}\gamma_{jk}$ can again be represented by a fixed component of the boundary of the regular neighborhood $N(i \cup j \cup k \cup \gamma_{ij} \cup \gamma_{ik})$. Since the three sums $\gamma_{ab}\gamma_{ac}$, $\gamma_{ab}\gamma_{bc}$, and $\gamma_{ac}\gamma_{bc}$ are isotopic, there exists a representative, d, in their common isotopy class, which coincides with each of the fixed sums to as great an extent as possible, but does not intersect each of the other tubed circles corresponding to the other sums. See Fig. 3.7). So for the fixed sum $\gamma_{ij}\gamma_{jk}$, let $x \in$ (circle arc corresponding to i) and $y \in$ (circle arc corresponding to k) be the points at which $\gamma_{ij}\gamma_{jk}$ and d cease to coincide. They lie on one side of the path γ_{ik}. Choose points x' and y' on the circle arcs corresponding to i and k respectively, which lie on the other side of γ_{ik}. Connect x' and y' by a path γ'_{ik} parallel to γ_{ik}. Let U and V be the small neighborhoods in the circle arcs of $\gamma_{ij}\gamma_{jk}$ bounded by x, x' and y, y' respectively. Let $W = d \cup U \cup V$. Then $\gamma_{ij}\gamma_{jk} - W$ is a connected curve consisting of tube arcs and circle arcs of $\gamma_{ij}\gamma_{jk}$ with endpoints x' and y'. So $(\gamma_{ij}\gamma_{jk} - W) \cup \gamma'_{ik}$ consists of a pair of (otherwise nonintersecting) curves with common endpoints, x' and y', so it is a simple closed curve. $\gamma_{ij}\gamma_{jk}$ fails to coincide with d precisely on $\gamma_{ij}\gamma_{jk} - W$. But $\gamma_{ij}\gamma_{jk}$ is isotopic to d. So in particular, $\gamma_{ij}\gamma_{jk} - W$ is isotopic to γ'_{ik}. Thus $(\gamma_{ij}\gamma_{jk} - W) \cup \gamma'_{ik}$ bounds a disk in F. But it is isotopic to a union of paths and portions of a, b, and c. /// See Fig. 3.7).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



■ $=\gamma_{ij}\gamma_{jk}$ $(=\gamma_{ab}\gamma_{ac}$ in this ex.)

■ $=\gamma'_{ik}$

▨ = d

■ $=\gamma_{ik}$ $(=\gamma_{bc}$ in this ex.)

Fig. 3.7)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Corollary 3.1 *If one of the summands j has γ_{ij} and γ_{jk} meeting j from different sides, then no pair of the 2-fold sums are isotopic.*

Pf. If γ_{ij} and γ_{jk} approach j from different sides, yet are connected to i and k, the union

$$(\gamma_{ij} \cup \gamma_{jk} \cup \gamma_{ik}) \bigcup (\text{circle arcs})$$

must go around or through a hole, so the corresponding collection of tube arcs and circle arcs does not bound a disk in F. See Fig. 3.8)

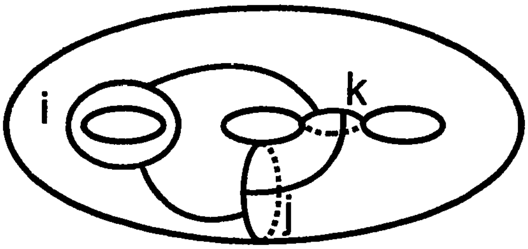


Fig. 3.8)

4 Disks and Spheres

In this section, we will look at some results involving interactions of disks and spheres within the context of Heegaard splittings. We start off with a result concerning the meridian disks of a handlebody, which seems rather natural, but whose proof is perhaps slightly more involved than might have been anticipated. It will be used in later discussions as well, and interacts in useful and important ways with the actions of $MCG(F)$ and the operations of the extended quandle.

Definition 4.1 *A loop is said to be* <u>*essential*</u> *in a surface F if is homotopically nontrivial in F.*

Theorem 4.1 *Let c be an essential imbedded loop in a surface F, where $F = \partial H$ and H is a handlebody of genus g. Then c bounds an imbedded disk in H iff c is, up to isotopy, a connected sum of meridian loops in F.*

Pf. If $c = a \#_\gamma b$ where a and b are meridians and γ is a path connecting them, then certainly a and b bound disks in H, and the "tube" which follows the path γ, along which we form the connected sum, may be pushed slightly into H, yielding a disk bounded by $a \#_\gamma b$. We may repeat this process for further connected sums.

Conversely, suppose that c is as stated in the hypotheses, and that $1 \neq [c] \in \pi_1(F)$, and that c bounds an imbedded disk in H (by Dehn's lemma). Suppose furthermore that c is nonseparating

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(i.e. homologous to 0) in F. We will assume that c is in general position with respect to all meridian curves m_i and waist curves w_i of F, and that it is " pulled tight", so that there are a minimal number of intersections with the m_i's and w_i's.

Claim: There exists some m_i or w_i (for brevity call it v) in F such that D may be pushed into F in a small neighborhood of v, and an arc α will then coincide with a subarc β of v.

Pf. of Claim Let us hold the case where $c = w_i$ in abeyance for a while.We shall see later that in this instance, c is clearly a connected sum of meridian curves m_i. So assuming that c is not a waist curve, if $|v \cap c| = 0$ for all v as above, then c does not intersect any m_i or w_i, and c is then homotopically trivial in F. Otherwise, let E be the disk in H bounded by v. If we put D into general position with respect to E, then D intersects E in a collection of arcs (with endpoints on $\partial E = v$) and circles. See Fig. 4.1a)

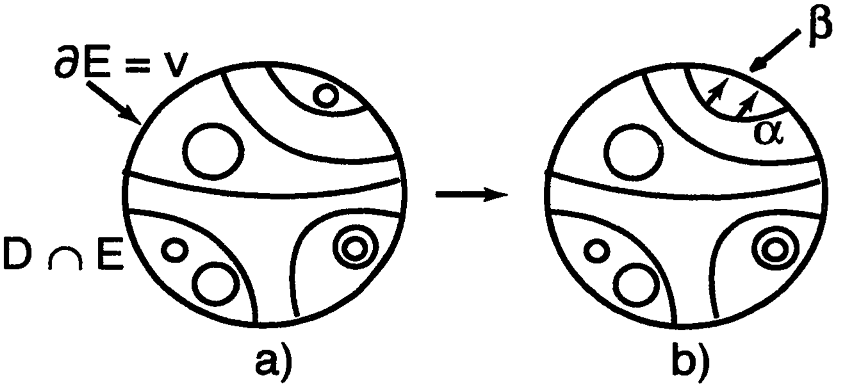


Fig. 4.1)

Suppose α is an innermost arc of intersection in E, and that there are no circles of intersection lying in the subdisk bounded by the arc $\beta \subset v$ and α. Then we can push α into F so that it coincides with β. On the other hand, if there are circles of $D \cap E$ lying in the subdisk of E bounded by α and β, pick an innermost one, Ω. It could not occur as an intersection of a "tube" portion of D and E, for this would imply D had positive genus. Thus Ω bounds a subdisk of D as well as a subdisk of E, so we get a copy of S^2 in H. H is a handlebody, so its universal cover is a thick tree T. But $0 = \pi_2(T) \cong \pi_2(H)$, so the S^2 bounds a ball, hence we may push the subdisk (bubble) of D through E, and eliminate this intersection. Continue doing this until no more circular intersection components lie in the disk region bounded by α and β. Now as before, the arc α may be pushed into F to coincide with β. This proves the claim. See Fig. 4.1b).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

But now this arc α, together with portions of $\partial D = c$, bound a pair of subdisks of D. See Fig. 4.2a). Cut along α, and perform surgery on the resulting curves, along α, to get a new pair of closed curves c' and c'', each of which bound disks. Clearly, performing the connected sum $c' \#_\gamma c''$ yields c, where γ is the short arc joining the images α' and α'' of α, in c' and c'' respectively. See Fig. 4.2b).

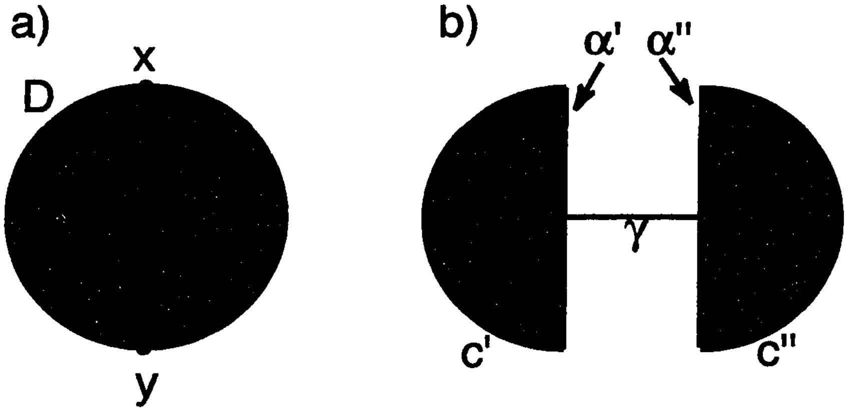


Fig. 4.2)

We now use the claim to complete the proof of the theorem.

We proceed by induction on the number of waist curves w_i, or meridian curves m_i, which c crosses. By assumption, c bounds a disk in H. The collection of $m_i's$ and $w_i's$ decompose F into pairs of pants, and the disks they bound cut H into balls, B^3. If c does not cross any w_i or m_i, then it lies entirely in one pair of pants. It bounds a disk in the corresponding portion of H, and is essential in F. But any such simple closed curve in a pair of pants must be parallel to the boundary circles of the pair of pants. In our case, the boundary circles are either meridians m_i, in which case c itself is a meridian, and we are done, trivially, or the boundary circle is a waist curve w_i. But a waist curve w_i is clearly a connected sum of a pair of adjacent meridian curves, and again we are done. See Fig. 4.3)

Fig. 4.3)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Now for the general case. Assume that if c crosses $\leq k-1$ meridian or waist curves, then it is a connected sum of meridian loops. Now suppose it crosses k such meridian or waist curves. By the claim above, there exists an arc α such that performing surgery on c, along α, yields two imbedded closed curves c' and c'', bounding disks in H. Each of these crosses fewer than k m_i's or w_i's, and so the inductive hypothesis holds, and they themselves are connected sums of the m_i's. But then as above, $c = c' \#_\gamma c''$, where γ is the short arc joining the images α' and α'' of α, in c' and c'' respectively, thus c is a connected sum of meridian curves, and we are done.///

We may use this to yield "algebraic" interpretations, in the language of the extended quandle, of a deep theorem of Casson and Gordon regarding the reducibility of Heegaard splittings, and of a statement about stabilizations of Heegaard splittings. We consider the latter first.

In chapter 1, a stabilized Heegaard splitting of a manifold M was described as one in which an extra handle had been added to each handlebody of the splitting, in such a way that the resulting 3-manifold was a connected sum of a manifold homeomorphic to M, and S^3. The splitting surface F is then a connected sum of a surface of genus g, and a torus. [Scharlemann] gives another characterization of a splitting which is stabilized. Namely, a splitting is stabilized iff there exist a pair of imbedded disks, D_1 and D_2, with $D_1 \subset H_1$ and $D_2 \subset H_2$ such that $|D_1 \cap D_2| = 1$. These may be thought of as meridian disks of the solid tori which are attached to H_1 and H_2 respectively, in the formation of the stabilized splitting, whose boundaries are joined to create the copy of S^3. (just the standard genus 1 splitting of S^3)

Lemma 4.1 *Let a and b be essential loops in F such that $b = x \#_\gamma y$, where x and y are also essential. If $|a \cap b| = 1$, then a may be isotoped so that either $|a \cap x| = 1$ or $|a \cap y| = 1$.*

Pf. Assume $|a \cap x| \neq 1$ or $|a \cap y| \neq 1$. Then a intersects b along the "tube" which follows γ. Slide the small segment of a, which contains the intersection point, to either x or y, along the tube wall.///

Now consider a Heegaard splitting of a 3-manifold M, with handlebodies H_1 and H_2 glued along their common boundary, F. Let $\mathcal{D}_1 = \{m_i : 1 \leq i \leq g\}$ be a set of meridian curves bounding disks D_i in H_1. Similarly, let $\mathcal{D}_2 = \{n_j : 1 \leq j \leq g\}$ be a set of meridian curves bounding disks D_j in H_2, such that each $\partial D_j = n_j$ is glued to the characteristic curve C_j on H_1.

Definition 4.2 *An imbedded disk D in a handlebody H is boundary reducing in H if ∂D lies in $\partial H = F$ and is essential in F.*

So the meridian disks D_i and D_j are boundary reducing in H_1 and H_2 respectively.

Recall the definition of P_K, (chapter 2) in order to make sense of the 'minus' sign in the following.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Corollary 4.1 *A splitting of a 3-manifold M is stabilized iff there exist meridian circles m_i from H_1 and n_j from H_2, such that* $m_i \overline{n_j}\rceil - n_j \lceil\overline{m_i} = 0$

Pf. By Scharlemann's characterization of stabilized spittings, the splitting is stabilized iff there exist a pair of boundary reducing disks in H_1 and H_2 respectively, which intersect each other exactly once in $F = \partial H_1 = \partial H_2$. By Thm. 4.1), any such boundary reducing disks are connected sums of meridian disks m_i and n_j in their respective handlebodies. Then Lemma 3.1) (regarding Dehn twists of circles intersecting each other once) says that for such disks, $m_i \overline{n_j}\rceil - n_j \lceil\overline{m_i} = 0$ iff $|m_i \cap n_j| = 1$. Finally, Lemma 4.1) assures us that we only needed to consider the "meridian disk portions" (as opposed to the "tube portions") of the connected sums forming the boundary reducing disks, when we looked at their intersection in F. ///

To a certain extent, having written the quandle expression

$$m_i \overline{n_j}\rceil - n_j \lceil\overline{m_i}$$

as above, the corollary has the feel of a statement which says 'if there exist circles, bounding disks in their respective handlebodies, whose "bracket" equals 0, then the splitting is stabilized'. This is somewhat premature, as I have not yet figured out how, or if, it is possible to create a well-defined Lie-type bracket on P using $\overline{}\rceil$ and $\lceil\overline{}$ as above. Nonetheless, the next corollary has a similar flavor, describing consequences of interactions of disks from each handlebody, and relating these to a statement about the "bracket" of their bounding circles.

Definition 4.3 *A Heegaard splitting of a manifold M with splitting surface S, is said to be a* *__reducible__* *Heegaard splitting, if there exists a 2-sphere in M which intersects the splitting surface F in a single essential circle.*

Although it will not be used here, it is worth mentioning that Haken showed that if a *manifold* is reducible, i.e. if there exists a 2-sphere in M which fails to bound a 3-ball on one side, then there exists a reducible splitting for M. In effect, a reducing sphere for M may be isotoped so as to intersect the splitting surface F in a single essential circle. Note the difference in terminology. The definition above defines reducibility for a Heegaard splitting. The text of the first sentence of this paragraph defines reducibility for manifolds. M reducible $\Rightarrow \exists$ a reducible splitting. But existence of a reducible splitting does not insure that M is reducible. For example, the "standard" genus 2 splitting of S^3 is clearly reducible (take a sphere containing one handle, which intersects F in the "commutator" curve.), but S^3 is not a reducible manifold.

The second corollary will be formulated based upon the following theorem from [Casson, Gordon]. For the statement, we need the following definition.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Definition 4.4 *A surface $S \subset M$ is incompressible in M if there does not exist an imbedded disk $D \subset M$ with ∂D essential in S. Otherwise, S is compressible in M, and D is called a compression disk.*

The operation of cutting M (and the compressible surface S) along such a disk D and glueing two copies of the disk to the resulting surface to form a new surface of lower genus, is called compressing along the disk D.

Theorem. *(Casson, Gordon) Let (H_1, H_2, F) be a Heegaard splitting for a closed manifold M. If there exist boundary reducing disks D_1 in H_1 and D_2 in H_2 with $|D_1 \cap D_2| = 0$, then either the splitting is reducible, or M contains an incompressible surface of positive genus.*

Let $\mathcal{D}_1 = \{m_i : 1 \leq i \leq g\}$ and $\mathcal{D}_2 = \{n_j : 1 \leq j \leq g\}$ be sets of meridian loops as above, for H_1 and H_2 respectively.

Corollary 4.2 *Given a splitting of a manifold M, if M has no incompressible surface, then the splitting is reducible if there exist some m_i and n_j such that $m_i \overline{n_j}\rceil - n_j \lceil\overline{m_i} = 0$.*

Pf. We first appeal to the theorem of Casson-Gordon, and then apply Thm. 4.1) to show that the boundary reducing disks are connected sums of meridian disks. Now apply Lemma 3.2) (characterizing Dehn twists about nonintersecting loops in terms of quandle operations)and Lemma 4.2) to complete this, in a similar fashion to their usage in the corollary above.

5 Waves

5.1 Reducibility of Splittings

In this section we define a geometric object which may be associated to certain Heegaard splittings, allowing them to be simplified in a certain sense. These geometric objects, 'waves', will also have algebraic significance within the extended quandle. We will use the notion of a wave, and its algebraic interpretation to prove results about Heegaard splittings and to gather further geometric intuition regarding interactions of disks in splittings and their associated handlebodies. The first result we shall prove will be a kind of bridge between the results of the type exemplified by the theorem of Casson and Gordon, mentioned in the previous chapter, regarding reducibiliy of splittings, and the existence and use of waves. The method of proof will involve the operations of the extended quandle. Initially, we will be looking at waves arising from one handlebody of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the splitting, and the initial definition will reflect this. Later, we will refine the definition to accomodate waves arising from each handlebody. For the initial definition, and for a key theorem, we shall refer to [Starr]. First however, a preliminary definition;

Definition 5.1 *For a handlebody H of genus g, with $\partial H = F$, a* <u>*complete disk system*</u> *$\mathcal{D}$ for H, consists of g imbedded, mutually disjoint disks in H, whose boundaries are essential in F, with the property that if we cut H along $\mathcal{D}$, the result is a 3-ball.*

There are generally many complete disk systems for handlebodies of genus > 2, a fact that will be important later on. In particular, if we look at a standard presentation of a handlebody H, (see e.g. Fig. 1.1)), a choice of g meridian disks constitutes a complete disk system $\mathcal{D}$, for H. This will be called the "standard" complete disk system for H.

Definition 5.2 *A* <u>*wave*</u> *for H is an open arc in $\partial H - \partial\mathcal{D}$ whose endpoints lie on the same component $\partial D_1 \in \partial\mathcal{D}$, such that the arc approaches D_1 from the same side at both of its endpoints, but is not homotopic (rel. endpoints) to any arc in $\partial\mathcal{D}$.*

See Fig. 5.1)

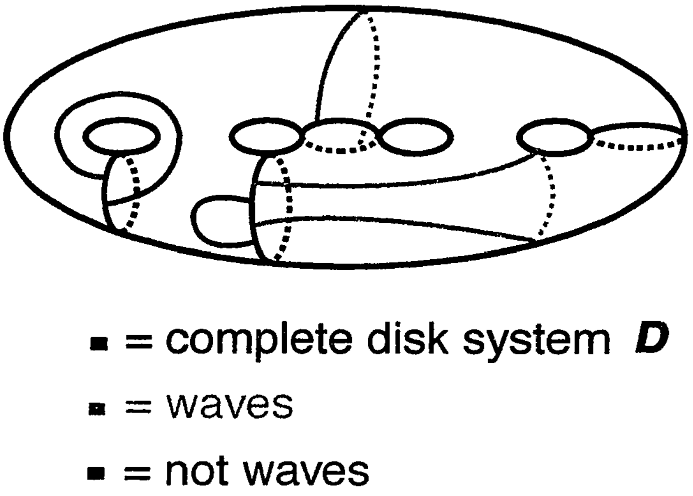


Fig. 5.1)

In proving the next theorem, we will make use of the following result.

Theorem *(Starr) Let H be a handlebody of genus g, and let A be an essential 1-manifold in $F = \partial H$. Let $N(A)$ denote a regular neighborhood of A. Then $F - N(A)$ is incompressible in H*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

iff there exists a complete disk system $\mathcal{D} = \{D_i | i = 1 \dots g\}$ *for* H*, having minimal intersections with* A*, such that* D *admits no waves disjoint from* A*.*

We will also use the following lemmas:

Lemma 5.1 *Let* M *be a 3-manifold, and suppose we have a Heegaard splitting where the glueing is given by the characteristic curves* $c_1 \dots c_g$ *on the splitting surface* F*. Then no pair of characteristic curves cobound a subsurface of* F*.*

Pf. We may think of the characteristic curves $c_1 \dots c_g$ as being gotten by applying a homeomorphism $\phi \in MCG(F)$ to the boundaries of the standard complete system of disks for H_1, where $F = \partial H_1$, shown in Fig. 1.1). Since no pair of these cobound a subsurface of F, no homeomorphic images of such cobound a subsurface. ///

Lemma 5.2 *Let* c *be a nonseparating circle on the surface* F *of genus* g*, and form* $(\mathcal{A} \otimes I)\Delta(c)$*, doubling the circle* c *and switching orientation on one component.* (see chapter 2). *Let* γ *be a path in* F *with* $\gamma \neq \tau$ *(where* τ *is the trivial path between the two components of* $\Delta(c)$*), so that* γ *joins the left side of one component to the right side of the other. Then the resulting loop* $c \#_\gamma c^{-1}$ *bounds a subsurface of* F *of genus 1.*

Pf. The two copies of c , between them, bound an annulus in F. Forming the nontrivial connected sum along γ adjoins an untwisted strip to this annulus, connecting its inner and outer boundary components. The result is a surface of genus 1 bounded by $c \#_\gamma c^{-1}$. ///

Theorem 5.1 *Let* M *be a closed orientable 3-manifold and suppose that* (H_1, H_2, F) *is a Heegaard splitting of genus* g *for* M*.*

1. *Suppose* $s \subset F$ *is a separating curve, bounding a disk in* H_2*. Suppose there exists a complete disk system* $\mathcal{D}$ *for* F*, which admits a collection of waves, missing* s*, which kill the genus of one of the subsurfaces of* F *bounded by* s*. Then the splitting is reducible. (Assume also that* $\mathcal{D} \cap s$ *is minimal)*
2. *A partial converse holds; if the splitting is reducible, there exists a complete disk system* $\mathcal{D}$*, and a wave on* F *for some* $D \in \mathcal{D}$*.*

Note that the arguments that will be used in the proof will work for an arbitrary separating curve on F, bounding a disk in H_2.

Pf. 1) We show first that it is always possible to create a separating curve s for F, which bounds a disk in H_2. In the Heegaard splitting of M, the meridian disks of H_2 are glued to g disjoint

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

characteristic curves $\{c_i | i = 1 \ldots g\}$ on $\partial H_1 = F$. Let $c \in \{c_i | i = 1 \ldots g\}$, and orient c somehow. Take a parallel copy of c and give it the opposite orientation, yielding c^{-1}.(This is just the "cobordism construction" of section 2.3)) Let γ be a path joining c to c^{-1}, with $\gamma \neq \tau$ (where τ is the canonical short arc joining the copies of c) such that γ does not intersect $\{c_i | i = 1 \ldots g\} - c$. To see that such an arc exists, let us suppose otherwise. Then any arc connecting the left side of c to the right side of c must pass through one of the other characteristic curves c_j. In other words, cutting along c_j separates the left side of c from the right side. But then c and c_j cobound a subsurface of F. This is not possible, by Lemma 5.1), so the desired path γ exists.

By Lemma 5.2) above, $c \#_\gamma c^{-1}$ bounds a genus 1 subsurface $F' \subset F$. Then since c bounds a disk in H_2, $c \#_\gamma c^{-1}$ bounds a disk in H_2, by Thm. 4.1). Since $c \#_\gamma c^{-1}$ is a connected sum of a meridian circle with another copy of itself, (along a path $\neq$ to the trivial path τ), it is an essential separating curve in F.

Now $c \#_\gamma c^{-1}$ is a 1-manifold in $F = \partial H$ which does not intersect any of the other c_i's, since c and γ were disjoint from the other c_i's. Consider complete disk systems for H_1, and choose one, call it $\mathcal{D}$, which intersects $c \#_\gamma c^{-1}$ minimally in the isotopy classes of the disk systems and of $c \#_\gamma c^{-1}$. Then for some disk $D \in \mathcal{D}$, if there are waves w_i with endpoints on ∂D such that $|w_i \cap c \#_\gamma c^{-1}| = 0$, which kill the genus of one of the subsurfaces of F cut off by s, Starr's theorem tells us that at least one subsurface of F bounded by $c \#_\gamma c^{-1}$ (recall it is separating in F) is compressible in H_1. This is because the proof of Starr's result shows that the existence of a wave implies the existence of a compression disk. In particular, let w be a wave with endpoints on $\partial D \in \mathcal{D}$. Then the compression disk arises as a disk bounded by one of the components of $\nu_w(\partial D)$. (Recall that ν was the operation in the extended quandle P which split a loop and an attendant wave into a pair of disjoint loops) Thus we may compress along this disk to lower the genus of the subsurface F', which is bounded by $c \#_\gamma c^{-1}$. Repeating this process of finding and compressing along compression disks arising from waves, for the remaining waves, we eventually reduce the genus of the subsurface bounded by $c \#_\gamma c^{-1}$ to 0. Again, this new subsurface (disk) lies in H_1. But $c \#_\gamma c^{-1}$ bounded a disk in H_2 as well, by the arguments above. So we get a 2-sphere which meets the original splitting surface in the essential circle $c \#_\gamma c^{-1}$, so the splitting is reducible. Similar arguments hold for arbitrary separating curves s, bounding disks in H_2./// See Fig. 5.2) for a picture of the process.

2) If the splitting is reducible, then there exists a 2-sphere which meets F in a single essential circle, c. This sphere intersects H_1 say, in a disk E. We may initially take $\mathcal{D}_1$, the complete disk system for H_1, to consist of the standard meridian disks for H_1. (see Fig. 1.5) Suppose $E \notin \mathcal{D}_1$.

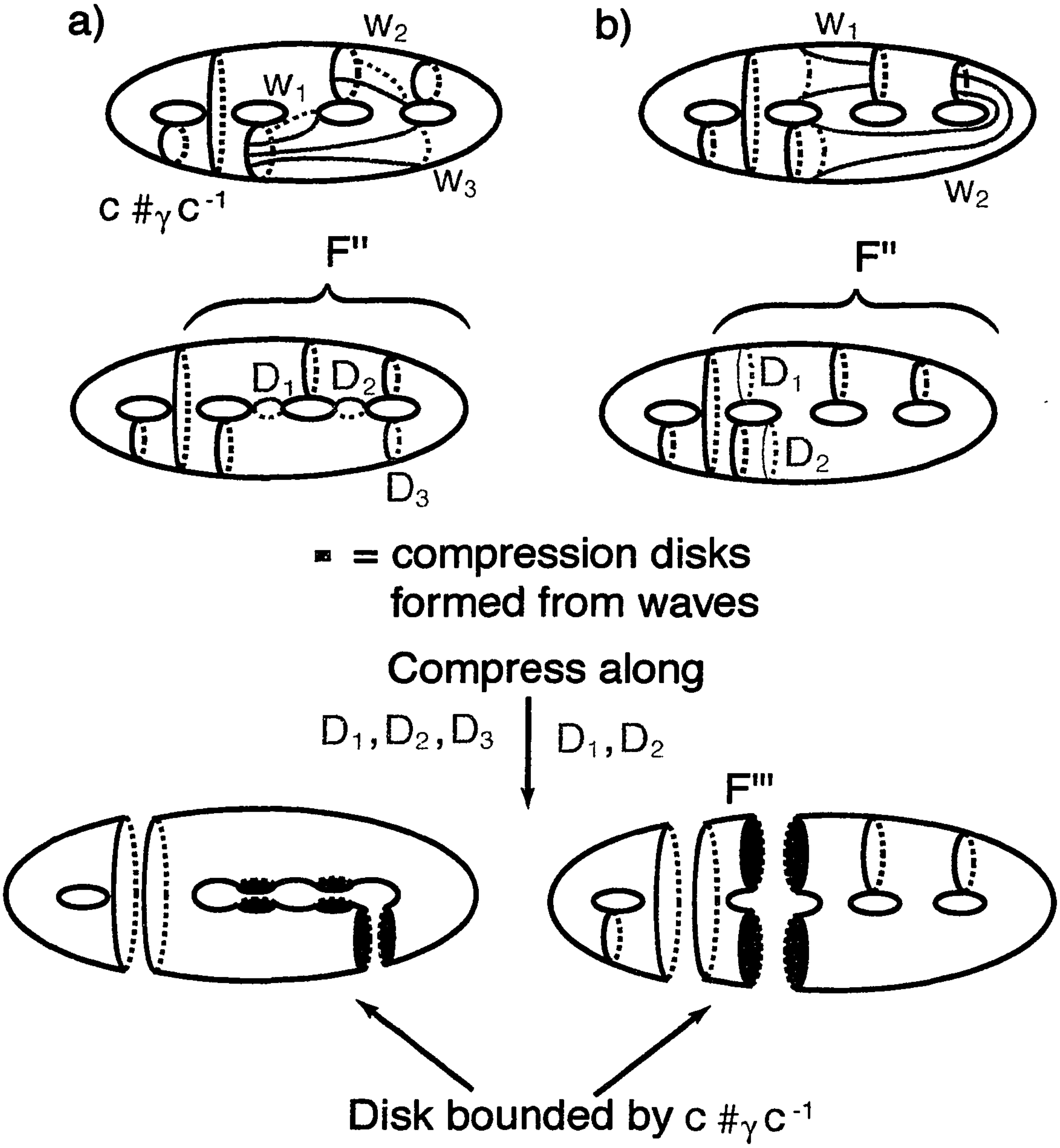
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

By Thm. 4.1), E is a connected sum of meridian disks in $\mathcal{D}_1$. Write $E = E_1 \#_\gamma E_2$. Now form a new disk system $\mathcal{D}'_1$ from $\mathcal{D}_1$, by replacing the meridian disk E_1 say in $\mathcal{D}_1$, by $E_1 \#_\gamma E_2$, and retaining E_2 and all the other disks of $\mathcal{D}_1$. Since $E_1 \#_\gamma E_2$ was the connected sum of E_1 and E_2 along the path γ (in the formation of E), Prop. 3.2) tells us that there is a path δ with endpoints on $\partial(E_1 \#_\gamma E_2)$ transverse to γ, and intersecting it once, such that

$$\nu_\delta \#_\gamma(\partial E_1 \otimes \partial E_2) = \partial E_1 \otimes \partial E_2$$

Here again, $\otimes$ corresponds to disjoint union of circles. The arc δ, with endpoints on the element $E_1 \#_\gamma E_2$ of $\mathcal{D}'_1$, is the wave we seek. On the other hand, if $E \in \mathcal{D}_1$, take a $\mathcal{D}'_1$ which does not contain E, and write E as a connected sum $E'_1 \#_\gamma E_2$, with E'_1, $E_2 \in \mathcal{D}'_1$. Then apply the argument above. ///

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



In a), number of waves = genus of subsurface

In b), number of waves < genus of subsurface

Fig. 5.2)

It is worth reiterating that the operations that were used to carry out the above proof are precisely those of the extended quandle. Taking a parallel copy (doubling) the characteristic curve c, and switching the orientation of one copy is achieved by $(\mathcal{A} \otimes I)\Delta(c)$ where A is the "antipode". We then took the connected sum along γ giving $\#_\gamma (\mathcal{A} \otimes I)\Delta(c)$, where I is the identity on P. We then applied ν to ∂D_i $(D_i \in \mathcal{D})$ to get compression disks in H_1. I am not certain how or if it is possible to give a completely algebraic proof of the previous argument, as perhaps an expression or an equation in the extended quandle, but the flavor of the extended quandle operations is apparent within it.

The process of sequentially reducing the genus of the subsurface bounded by $c\#_\gamma c^{-1}$ by compressing along disks is similar to the process used by Casson and Gordon in the theorem mentioned in chapter 4. They form a surface, capped off by the two nonintersecting disks in the hypothesis, and then using the other hypothesis of having no incompressible surfaces, proceed to lower the genus by compressing. Here, the existence of waves and their corresponding compression disks, allows us to proceed with the reduction of genus. To a certain extent, Thm. 5.2) and the theorem of Starr are complements of one another in the theorem of Casson-Gordon.

5.2 Further Results on Waves, and Relating Waves to a Subgroup of $MCG(F)$

We shall continue to look at questions of the existence and ramifications of waves in Heegaard splittings, and show how these further tie into the algebra of the extended quandle. In the mid '70's [Volodin, Kuznetzov, Fomenko] came out with what they believed was an algorithm which would enable them to distinguish Heegaard splittings of S^3 among splittings of all closed, orientable 3-manifolds. Their algorithm hinged on the ability, they thought, to find waves in the Heegaard splittings of S^3 for genus > 1, except the "standard" splitting in each genus, and to simplify the diagrams using these waves. This turned out not to be the case. [Ochiai], [Viro], and [Morikawa] were able to produce counterexamples of splittings which were not the standard splitting of a given genus, but which did not have waves. However, [Homma, Ochiai, Takahashi] did show that such an algorithm, using the existence of waves, did occur in genus 2. I will examine this in greater detail a bit later, and after having developed a bit more on the properties of waves, reprove their result using the machinery of the extended quandle. As promised, however, we need to expand our definition of a wave, in order to proceed. Suppose (H_1, H_2, F) is a Heegaard splitting of genus g for a closed, orientable 3-manifold M. We may consider this as being given by a pair of complete disk systems $\mathcal{D}_i$ for H_i, $(i = 1, 2)$. $\partial\mathcal{D}_2$ describes the characteristic curves on F.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

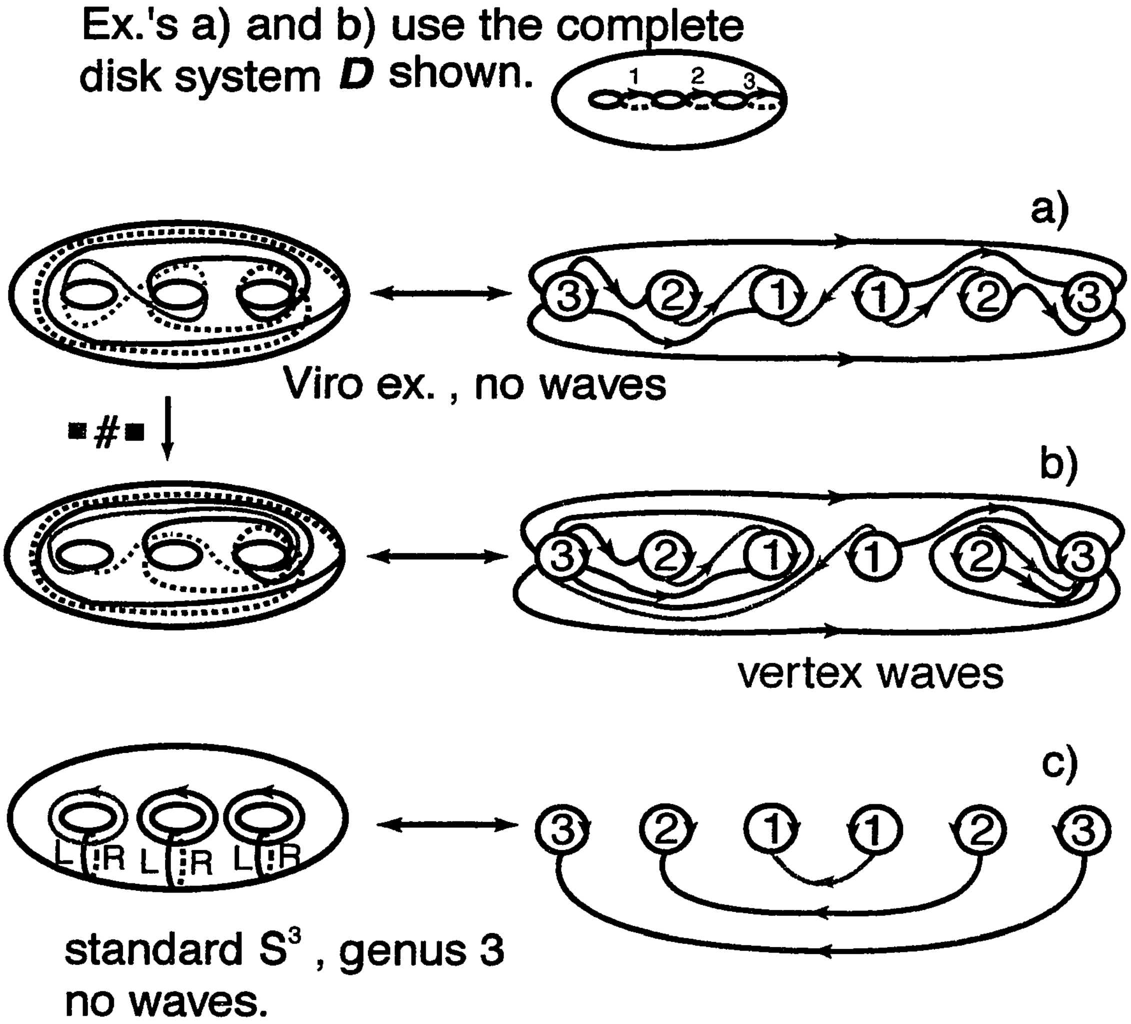
Definition 5.3 *A Heegaard diagram for a given splitting of genus g is a planar surface with 2g holes, joined by arcs. We get a Heegaard diagram by cutting H_1 along the disks of $\mathcal{D}_1$, and looking at the resulting surface. It is a sphere with 2g boundary components. The boundary component circles are joined to one another by arcs, which are remnants of the characteristic curves on F. Removing one more disk, which avoids all the other boundary components and arcs yields the planar surface with markings.*

We adapt the expanded definition of a wave from [Homma, Ochiai, Takahashi]. Consider a Heegaard splitting and associated diagram of genus g. $\partial\mathcal{D}_1$ is the collection of the 2g circular boundary components of the disk-with-holes, and $\partial\mathcal{D}_2$ yields the arcs corresponding to the characteristic curves of the splitting. The boundary circles together with the arcs divide the surface into regions, each bounded by some arcs, and some portions of boundary circles. See Fig. 5.3a) Assume all the circles in $\partial\mathcal{D}_1$ and $\partial\mathcal{D}_2$ are oriented.

Definition 5.4 *An arc in a region bounded by portions of circles from $\partial\mathcal{D}_1$ and $\partial\mathcal{D}_2$, with endpoints on this boundary, is a wave if it joins two segments of the same circle, which are components of the boundary of the region, and approaches these edges from the same side (according to the orientation of the circle). See Fig. 5.3b)*

We can see from this that the "standard" splitting of genus g for S^3 admits no waves (Fig. 5.3c) since we do not have any *pair* of edges appearing, which come from the same circle. Hence there is no possible arc fulfilling the requirements of the definition. The arcs and the boundary circles of the Heegaard diagram may be thought of as forming a graph, with arcs as edges and circles as (fat) vertices. Waves that connect arcs from $\partial\mathcal{D}_2$ will be "edge" waves, while those connecting portions of circles from $\partial\mathcal{D}_1$ will be "vertex" waves. See Fig. 5.4)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



Ex.'s a) and b) use the complete disk system ***D*** shown.

Viro ex. , no waves

vertex waves

standard S^3 , genus 3
no waves.

Fig. 5.3)

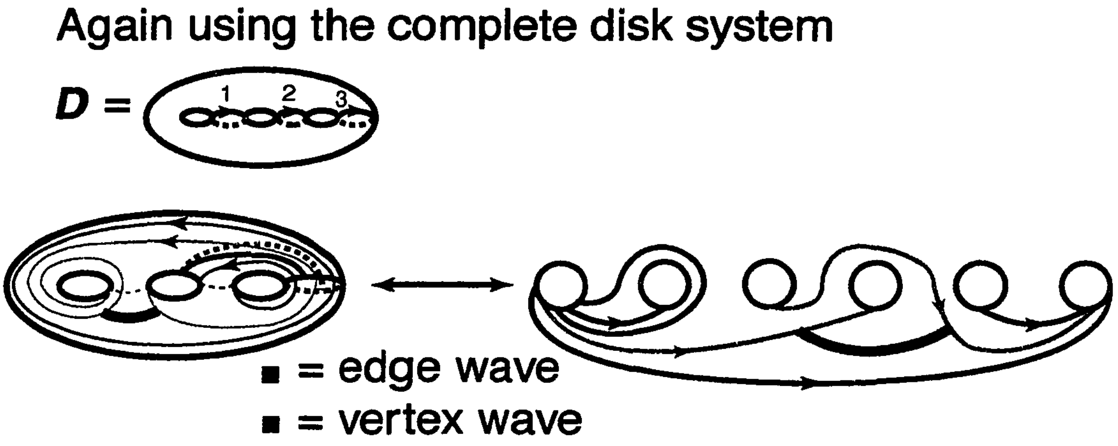


Fig. 5.4)

As we shall see (and it will be important in what follows), waves are defined with respect to a particular complete system of disks. If we change from one complete disk system to another, the waves which exist for one may not exist for the other. We may gain or lose some in the transition. In the section that follows, we'll look at a subgroup of $MCG(F)$ which preserves the number of waves for a given splitting.

Let (H_1, H_2, F) be a Heegaard splitting of genus g for a closed orientable manifold M. Let $\mathcal{D}$ be a chosen complete disk system for H_1. Generally, due to the niceties of the symmetry, I will take $\mathcal{D}$ so that

$$\partial\mathcal{D} = \{\textit{waist curves } w_i \,|\, i = 1 \ldots g-1\} \cup \{\textit{the meridian }, m_g\}$$

See Fig. 5.5)

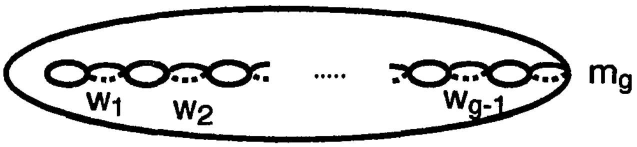


Fig. 5.5)

Recall (from Chapter 1) that $\mathcal{F}_g < MCG(F)$ is the subgroup which sends disks to disks in H_1. Let $G_{\mathcal{D}i} < \mathcal{F}_g$ be the subgroups of $MCG(F)$ which preserves the complete disk systems $\mathcal{D}_i$ in H_i, i = 1,2.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Theorem 5.2 *Suppose* $(H_1\,, H_2\,, F)$ *is a Heegaard splitting of genus* g *as above. We may consider its Hegaard diagram with respect to* $\mathcal{D}_1$.

1. *The elements of the subgroup* $G_{\mathcal{D}_i} < \mathcal{F}_g < MCG(F)$ *preserve the number of waves for the splitting. (Since* $G_{\mathcal{D}_i} < \mathcal{F}_g$, *elements* $\phi \in G_{\mathcal{D}_i}$ *applied to* $(H_1\,, H_2\,, F)$ *yield splittings of the same homeomorphism type)*

2. $G_{\mathcal{D}_i}$ *is generated by*

$$\{\{\overline{w_i}\rceil \mid i = 1\ldots g\} \quad , \{\overline{l_i}\rceil\,\overline{w_{i-1}}\rceil\;\overline{w_i}\rceil\,\overline{l_i}\rceil \mid i = 2\ldots g\} \quad , \overline{(w_1 \#_\gamma w_2)}\rceil\,, \textit{and}\;\; \overline{l_1}\rceil\,\overline{w_1}\rceil\,\overline{w_1}\rceil\,\overline{l_1}\rceil\,\}$$

where γ *is as shown in Fig. 5.6) and* w_g *is the meridian* m_g

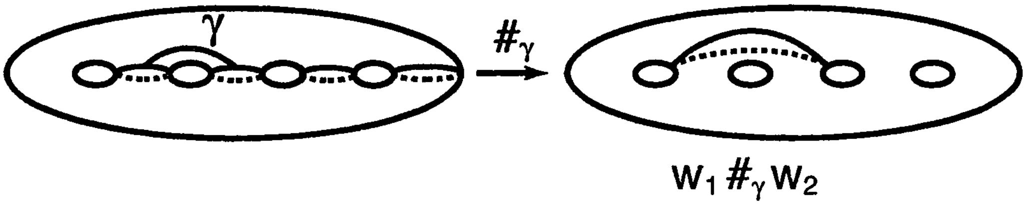


Fig. 5.6)

Pf. As above, we are taking

$$\mathcal{D}_1 = \{\text{waist curves}\, w_i \mid i = 1\ldots g\} \cup \{\text{meridian}\, m_g\},$$

so clearly twists around the w_i's and around m_g maintain the same arcs entering and leaving the boundaries of these disks, as well as maintaining the relative positions of their points of intersection with the disks. So arcs or strands that were parallel, remain parallel. Also the regions of the Heegaard diagram, which are created by the arcs of $\partial\mathcal{D}_2$ and the circles of $\partial\mathcal{D}_1$ may thus be twisted, but otherwise remain unchanged. In particular, pairs of boundary arcs which admitted a wave still do, while those that did not, still do not.

Performing $\overline{l_1}\rceil\;\overline{w_1}\rceil\;\overline{w_1}\rceil\;\;\overline{l_1}\rceil$ has the effect of reversing the orientation of $w_1 = \partial\mathcal{D}_1$, i.e. $\overline{l_1}\rceil\;\overline{w_1}\rceil\;\overline{w_1}\rceil\;\;\overline{l_1}\rceil$ sends D_1 to itself as a set, reversing the orientation on the boundary, but has no effect on the rest of the disks in $\mathcal{D}_1$. Hence it also leaves the Heegaard diagram unchanged, and so preserves the number of waves.

$\overline{(w_1 \#_\gamma w_2)}\rceil$ is a Dehn twist around a circle which does not intersect any of the disks in $\mathcal{D}_1$, and its effect on strands entering the boundary of a disk in $\mathcal{D}_1$ is to twist the strands but to keep

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

parallel strands parallel. So again, waves are neither created nor destroyed (essentially, regions are twisted, but preserved).

Finally, homeomorphisms of the form $\overline{l_i}| \; \overline{w_{i-1}}| \; \overline{w_i}| \; \overline{l_i}|$ interchange waist disks D_{i-1} and D_i while not affecting other disks in $\mathcal{D}_1$. Consider such a homeomorphism acting on a pair of such adjacent disks, and on arcs joining them (e.g. arcs coming from longitude curves). We may then observe that we can obtain the resulting Heegaard diagram, from the initial one, by sliding the boundary circles past one another, in the diagram, in a manner determined by the particular homeomorphism used. (or its inverse). We can combine many such "pair flips" to obtain any permutation of the disks in $\mathcal{D}_1$. Thus, such operations merely isotope the boundary circles around in the Heegaard diagram, without changing which strands enter or leave a given boundary circle. So the new diagram, waves or no waves, is isotopic to the original diagram, and the number of waves is preserved. Intuitively, that such homeomorphisms should generate the group of homeomorphisms that preserve $\mathcal{D}_1$ seems reasonable, since they merely permute or flip disks in $\mathcal{D}_1$, and flips of disks other than D_1 can be accomplished by moving them to the 1st position via applications of the relevant $\overline{l_i}| \; \overline{w_{i-1}}| \; \overline{w_i}| \; \overline{l_i}|$'s, then flipping by $\overline{l_1}| \; \overline{w_1}| \; \overline{w_1}| \; \overline{l_1}|$, and then moving back. Similarly, $\overline{(w_1 \#_\gamma w_2)}|$, the twist around the loop shown in Fig. 5.6), can be transported to twists around other loops which are connected sums of the disks in $\mathcal{D}_1$, which don't meet elements of $\mathcal{D}_1$. For an actual proof of generation, we note that $MCG(F)$ acts transitively on complete disk systems and it is shown in [Wajnryb] that for $\mathcal{D} =$ the standard set of meridian disks, $G_\mathcal{D}$ is generated by the conjugates of the homeomorphisms listed above by the homeomorphism

$$f = (\overline{l_i}| \, \overline{w_i}| \; |\overline{m_i} \, |\overline{l_i}) \cdot (\overline{l_{i-1}}| \; \overline{w_{i-1}}| \; |\overline{m_{i-1}} \; |\overline{l_{i-1}}) \cdots (\overline{l_1}| \, \overline{w_1}| \; |\overline{m_1} \, |\overline{l_1})$$

This is just a glorified application of the quandle relation Q3). Here $(\mathcal{D}_1)\, f = \mathcal{D}$. See Fig. 5.7) for justification.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

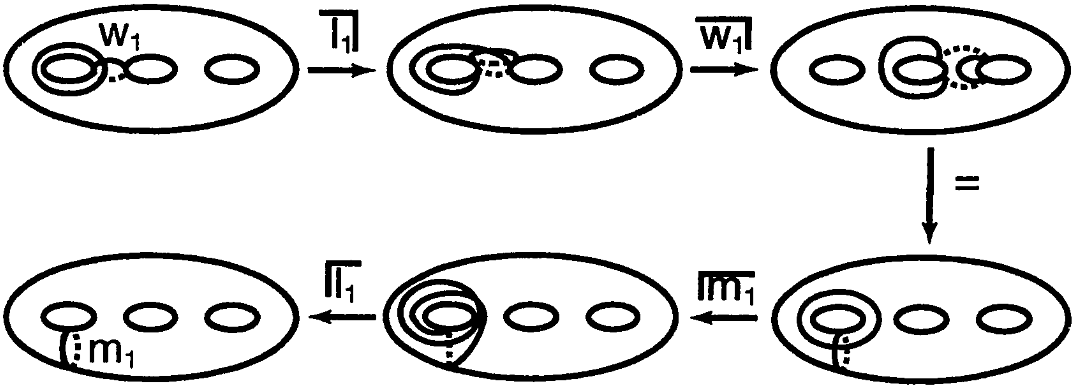


Fig. 5.7)

Fig. 5.7) shows $(w_1)\overline{l_1}\rceil\,\overline{w_1}\rceil\,\lceil\overline{m}_1\,\lceil\overline{l_1} = m_1$. The analogous results hold for the other w_i's. ///

Lest one harbor the erroneous notion, as I did for about a week, that $G_{\mathcal{D}1}$ is the entire collection of homeomorphims fixing the number of waves for a given splitting, consider the following. By symmetry, $G_{\mathcal{D}2}$ also preserves the number of waves. In fact any product in $G_{\mathcal{D}2} \cdot G_{\mathcal{D}1}$ will fix the number of waves. For a given splitting, $\phi \in G_{\mathcal{D}2}$ does not change the collection of characteristic curves, $\partial\mathcal{D}_2$, of the splitting. But then, since the splitting was taken with respect to the disk system $\mathcal{D}_1$, which is fixed, we need not consider the effect of ϕ on $\partial\mathcal{D}_1$. (i.e. we can apply $\phi \in G_{\mathcal{D}2}$, and then "insert" the chosen disk system $G_{\mathcal{D}1}$.) Then applying $\psi \in G_{\mathcal{D}1}$ fixes $\mathcal{D}_1$, and yields new characteristic curves $(\partial\mathcal{D}_2)\psi$. But by Thm. 5.6) above, it does not introduce or destroy waves. The situation is actually somewhat more subtle, as the next example will show.

We will look at a counterexample given by [Ochiai] to the conjecture of [Volodin, Kuznetsov, Fomenko], regarding whether all splittings of S^3 except the standard ones in each genus, admit waves. Ochiai actually gives a whole family of counterexamples in genus 4. To create this family, Ochiai considers "extended" complete disk systems for the handlebodies H_1 and H_2. For a handlebody of genus g, these are systems which consist of g+1 disks, such that cutting along them would separate the handlebody. So in this instance, an extended system will consist of 5 disks in each of H_1 and H_2, such that any choice of four of these, in each handlebody, yields a splitting of S^3. In particular, the example was constructed so that any such choice of four disks in each handlebody, coming from the extended disk system, gives a splitting without waves. See Fig. 5.8)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In Fig. 5.8),

a) shows the splitting surface with extended disk systems $\overline{\mathcal{D}}_1$ and $\overline{\mathcal{D}}_2$ where $\partial\overline{\mathcal{D}}_1 = \{d_i \mid i = 1\ldots5\}$ and $\partial\overline{\mathcal{D}}_2 = \{c_i \mid i = 1\ldots5\}$.

b) shows splittings of F and F' (with 4 disks in each handlebody) such that

$$F \leftrightarrow \partial\mathcal{D}_1 = \{d_2, d_3, d_4, d_5\} \text{ and } \partial\mathcal{D}_2 = \{c_2, c_3, c_4, c_5\}$$

and

$$F' \leftrightarrow \partial\mathcal{D}'_1 = \{d_2, d_3, d_4, d_5\} \text{ and } \partial\mathcal{D}'_2 = \{c_1, c_3, c_4, c_5\}$$

c) shows the Heegaard diagram associated with the spitting F shown in b). Note the absence of waves.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

With $\partial\overline{D}_1$ as shown, for the extended disk system $\overline{D}_1$

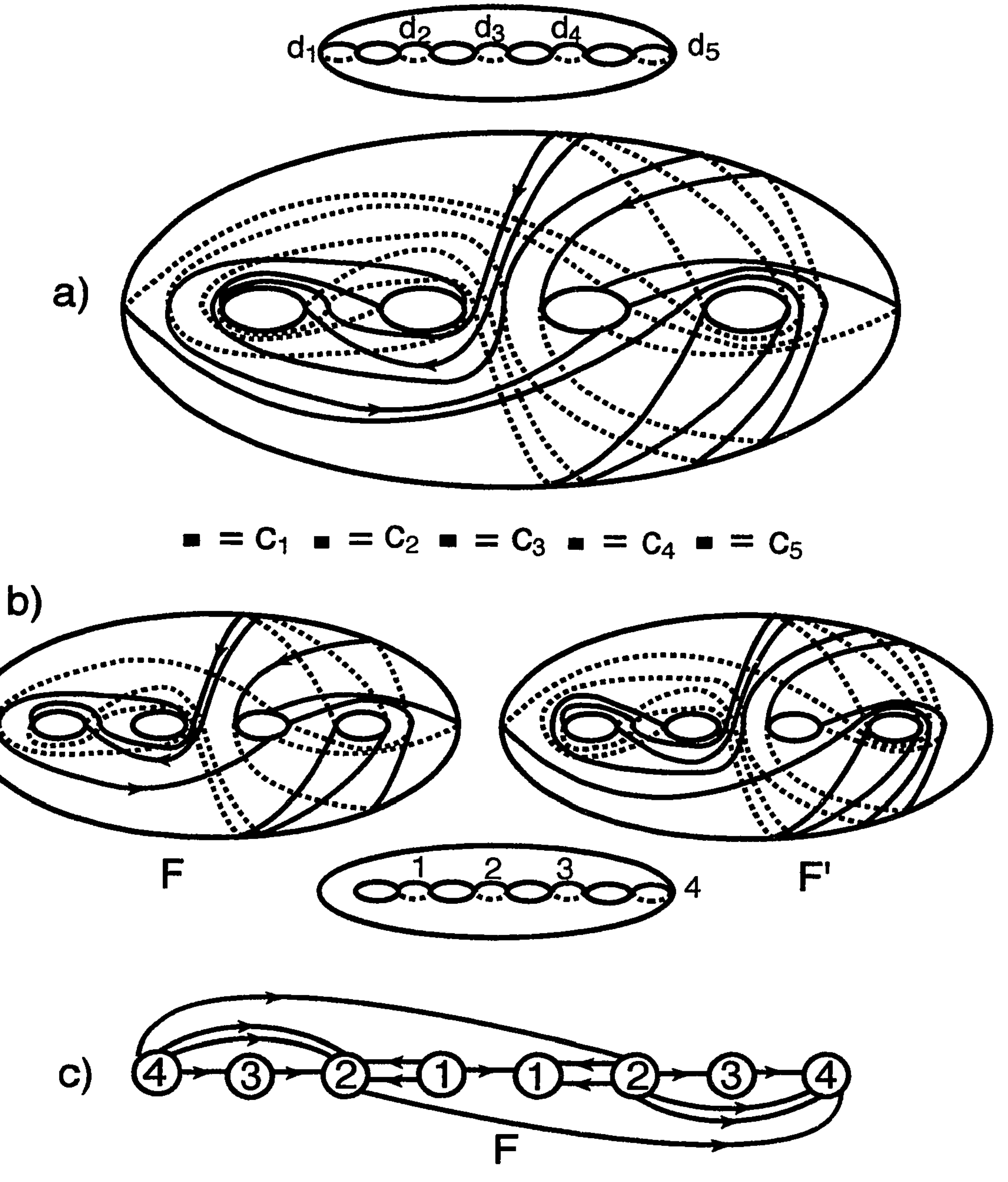


Fig. 5.8)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In Fig. 5.9) I show four disjoint loops $\alpha_1, \alpha_2, \alpha_3$, and α_4 such that a composition of Dehn twists around them sends the characteristic curves of F above to those of F', determined by $\partial\mathcal{D}'_1 = \{d_2, d_3, d_4, d_5\}$ and $\partial\mathcal{D}'_2 = \{c_1, c_3, c_4, c_5\}$, i.e. it replaces c_2 by c_1, leaving all other circles unchanged. Since the α's are mutually disjoint, the order of composition is unimportant. If we consider the group generated by the α's, $< \alpha_1, \alpha_2, \alpha_3, \alpha_4 > \cap G_{\mathcal{D}_2} = \emptyset$, since the product sends $\mathcal{D}_2$ to $\mathcal{D}'_2$, while individually, or in sub-products, the α's don't even preserve disks. Also, neither the product nor any sub-product lie in $G_{\mathcal{D}1}$, since circles bounding disks in $\mathcal{D}_1$ are sent to circles that no longer bound disks in H_1, and so the α's, in product or individually, are not even in the subgroup $\mathcal{F}_4 < MCG(F)$, whose elements send disks to disks in H_1. Since none of the splittings arising from picking disk systems from the extended disk systems $\overline{\mathcal{D}}_1$ and $\overline{\mathcal{D}}_2$ have waves, we have

$$(F)\,\lceil\overline{\alpha_1}\,\lceil\overline{\alpha_2}\,\overline{\alpha_3}\rceil\,\overline{\alpha_4}\rceil = F'$$

So in this instance, the homeomorphism $\lceil\overline{\alpha_1}\;\lceil\overline{\alpha_2}\;\overline{\alpha_3}\rceil\;\overline{\alpha_4}\rceil$ does preserve the number of waves. Furthermore, if some of the splittings arising from the extended disk systems $\overline{\mathcal{D}}_1$ and $\overline{\mathcal{D}}_2$ have waves, and some do not, then a homeomorphism taking one with waves to one without, would not preserve the number of waves. The example detailed above is relatively special in that all splittings arising from the extended disk systems had the same number of waves, zero. More generally, it would seem reasonable that if extended disk systems were created for the manifold M, with the following properties:

1. all complete disk systems coming from the extended system give rise to the same manifold
2. all complete disk systems coming from the extended system have equal numbers of waves

then the collection of homeomorphisms that preserved the extended disk system, for either H_1 or H_2, would preserve the number of waves. There. Lotsa time saved. Spend it wisely. Don't spend it in too many places at once.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

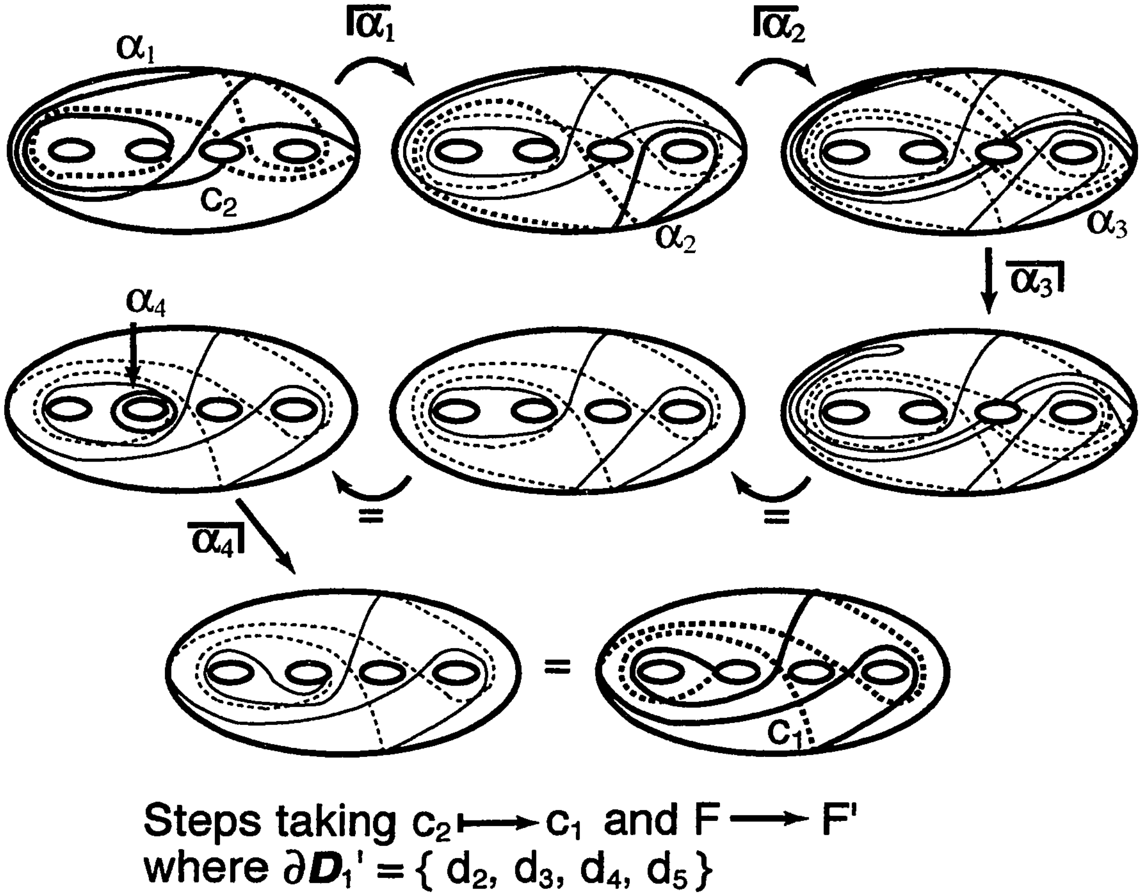


Fig. 5.9)

Apropos the foregoing discussion, I will use Thm. 5.6) and properties of the extended quandle to reprove the following result of [Homma, Ochiai, Takahashi]

Theorem *(Homma, Ochiai, Takahashi) Among the collection of genus 2 splittings of* S^3*, taken with respect to the standard complete disk system* $\mathcal{D}_1$*, with* $\partial\mathcal{D}_1 = \{m_1, m_2\}$*, all splittings except the standard splitting admit waves.*

So the conjecture of [Volodin, Kuznetzov, Fomenko] is true in genus 2.

To prove this, we will need to show that for any Heegaard diagram of genus 2 representing S^3, which is not the diagram of the standard splitting, there exists at least one wave w. If w has endpoints on one meridian, it must intersect neither the other meridian, nor the circles of $\partial\mathcal{D}_2$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

which form the characteristic curves of the splitting. In this way, it will be seen to lie in one of the "regions" mentioned by [Homma, Ochiai, Takahashi] and described in Definition 5.4).

Pf. of Theorem. Supposing as per hypothesis, that F is of genus 2 and m_1 and m_2 are standard meridians so that for the complete disk system $\mathcal{D}_1$, we have $\partial\mathcal{D}_1 = \{m_1, m_2\}$. We consider the following cases;

Case 1)

Suppose $\psi \in G_{\mathcal{D}1}$. Then by Lemma 5.5) the only effect ψ has on the Heegaard diagram of the standard genus 2 splitting is to isotope the boundary circles around. In particular, it will not change the number of pairs of edges available for the formation of waves. But since the standard splitting of genus 2 (see Fig. 5.10)) has 2 components with one edge apiece, neither admits waves. (recall that a wave required a pair of distinct edges on the boundary of a region in the splitting diagram, and an arc between them.) Hence, elements of $G_{\mathcal{D}1}$ applied to the standard genus 2 spitting yield diagrams isotopic to the original diagram, which has no waves.

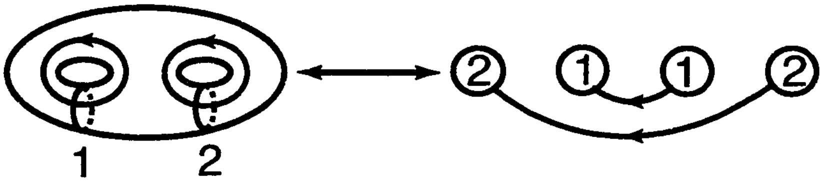


Fig. 5.10)

Case 2)

Now assume that $\psi \in \mathcal{F}_2 - G_{\mathcal{D}1}$

Claim: The image $(m_i)\psi$ is either disjoint from m_j, or can be pushed off m_j $\forall\, i, j \in \{1, 2\}$.

Pf. of Claim. If $(m_i)\psi = m_i$ for $i \neq j$, then this is clear. On the other hand, if $(m_i)\psi = m_j$ for $i \neq j$, we may isotope the image copy of m_j off another copy of m_j. Finally, if $(m_i)\psi$ is not a meridian, then by Thm. 4.1), it is a connected sum of meridians. Such connected sums on the surface of genus 2 can be only of the following types (see Fig. 5.11a)), since if a path γ connects m_1 and m_2, it cannot go around a hole, since then we could not form a legal connected sum (there would be self-intersections). Also, the image $(m_i)\psi$ cannot be a connected sum of 2 copies of a single meridian circle, since then it would be a separating curve on F. This would contradict the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

fact that m_i is nonseparating, and ψ is a homeomorphism. See Fig. 5.11b). From the figure, it is clear that all such connected sums may be pushed off m_1 or m_2. This proves the claim.

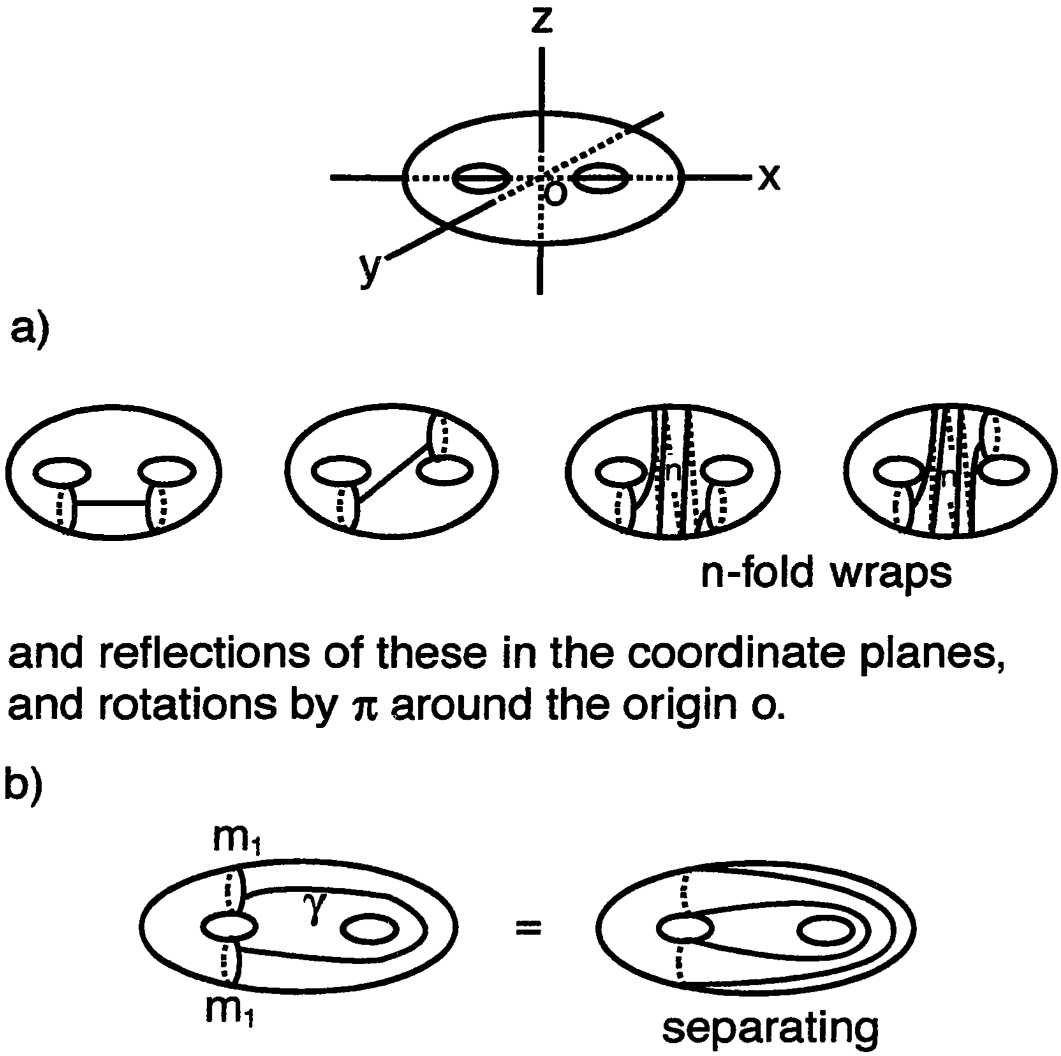


Fig. 5.11)

Now, since $|l_j \cap m_j| = 1$ for $j \in \{1, 2\}$, then $|(l_j)\psi \cap (m_j)\psi| = 1$, where l_j represents the j-th longitude circle. Thus, in the connected sum $(m_j)\psi = m_i \#_\gamma m_j$, we may use Lemma 4.1) to slide the intersection point $x = (l_j)\psi \cap (m_j)\psi$ so that it lies in the m_j "portion" of the connected sum and not in the tube connecting m_i and m_j. But now Prop. 3.2) assures us that since we have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$(m_j)\psi$ represented as a connected sum, with γ the connecting path, there exists another short path τ, transverse and perpendicular to γ, with endpoints on $(m_j)\psi = m_i \#_\gamma m_j$. We may think of τ as connecting the two inside walls of the tube. It is a wave for $(m_j)\psi$. I particular, since we have pushed the intersection point $x = (l_j)\psi \cap (m_j)\psi$ off the tube, τ avoids $(l_j)\psi$. It is also the case that $(l_j)\psi \cap (m_i)\psi = \emptyset$ for $i \neq j$, since $l_j \cap m_i = \emptyset$. Thus the wave τ misses the images of the characteristic curves l_1 and l_2 as well as avoiding the other meridian m_i, as required.

Now, following [Birman], we give the next definition.

Definition 5.5 *Two genus g Heegaard splittings F and F′ are said to be strongly equivalent, denoted* $F \sim F'$*, if there exists an orientation preserving homeomorphism* $h : H_1 \cup_\psi H_2 \to H_1 \cup_\phi H_2$ *such that* $(H_1)h = H_1$ *and* $(H_2)h = H_2$.

Here, homeomorphisms ψ, $\phi \in MCG(F)$ yield the characteristic curves on F and F' respectively. Birman proves that $F \sim F'$ iff there exist homeomorphisms f_1, $f_2 \in \mathcal{F}_g$ such that $\psi = f_1 \phi f_2$.

[Waldhausen] and later [Scharlemann, Thompson] prove that all splittings of S^3 are strongly equivalent. If ψ is the homeomorphism on $F = \partial H_1$ yielding the standard splitting F, with characteristic curves given by the longitudes l_1 and l_2, then we may think of ψ as having created these characterstic curves by acting on the initial configuration of the handlebody H_1. Recall from chapter 1 that the initial configuration consisted of the two (in this case) meridian curves m_1 and m_2. See Fig 5.12).

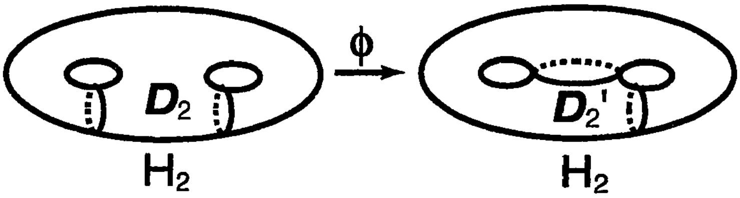


Fig. 5.12)

From the discussion above, any other configuration of characteristic curves on the surface, corresponding to F' and the homeomorphism $\phi \in MCG(F)$, must be related by

$$\phi = f_1 \psi f_2.$$

Now as elements of $\mathcal{F}_2$, the f_i $i = 1, 2$ merely permute disks in the complete disk system $\mathcal{D}_1$, where $\partial\mathcal{D}_1 = \{m_1, m_2\}$. After applying f_1, m_1 and m_2 still bound disks in H_1. Then $\psi : m_i \to$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

l_i $i = 1, 2$, but does not change the inside of H_1. (recall that $\psi \in (MCG(F) - \mathcal{F}_2)$, so it does not extend to H_1). Now if $f_2 \in G_{\mathcal{D}1}$, we showed above that it created no waves, and the Heegaard diagram it yields is in fact the same as the one associated to ψ. On the other hand, if $f_2 \in \mathcal{F}_2 - G_{\mathcal{D}1}$, then it did indeed yield waves. Thus for all new configurations of characteristic curves, that is those which do not merely differ from the standard configuration by elements of $G_{\mathcal{D}1}$ (which give isotopic Heegaard diagrams anyway), the homeomorphisms giving rise to them also create waves. ///

The proof just given depends heavily upon the relative scarcity of disks and possible connected sums of meridians in the handlebody of genus 2. This is quite different from the situation for higher genus. In particular in higher genus, we would be unable to limit the types of possible connected sum as we have in Fig. 5.11). The dearth of disks in the handlebody of genus 2 also yields another somewhat unusual property; any homeomorphism in $MCG(F)$ that sends a disk in the handlebody to another disk, sends all disks to disks, and thus lies in $\mathcal{F}_2$. So whether a homeomorphism lies in $\mathcal{F}_2$ is determined by its behavior on a single disk. This is not the case for genus > 2.

6 Dessert: Quirks, Oddities, and some Post-Prandial Speechifying

In this section, we look at some other structures that arise from the extended quandle P, and also pose some further questions.

To this end, let us assume our surface F is of genus g, and let

$$S_x = \{b \in P \mid x\,\overline{b}\rceil = x\lceil\overline{b} = x\}$$

So S_x is the collection of circles on F which don't intersect the circle x. (If a circle intersected x nontrivially, Dehn twisting around it would insert a copy of it in x, yielding a result different from x. Dehn twists about elements of S_x form a subset of $Stab(x) < MCG(F)$. S_x is closed under $\lceil\overline{\ \ }$ and $\overline{\ \ }\rceil$. For if $a\,, b \in S_x$, then $|a \cap x| = 0 = |b \cap x|$, and thus none of $a\lceil\overline{b}$, $a\,\overline{b}\rceil$, $b\lceil\overline{a}$, $b\,\overline{a}\rceil$ intersect x. One can see this by realizing that inserting one circle into another via Dehn twist is a purely local operation. Thus, all of the four resulting circles are in S_x. Furthermore, S_x is closed under connected sum, assuming that for $a\,, b \in S_x$, with $a \cap b = \emptyset$, any connected sum $a \#_\gamma b$ is formed along a path γ which either doesn't meet x or may be homotoped off x. Thus for each such x, S_x is a sub-extended quandle of P, with group $< S_x > < MCG(F)$, somewhat analogous, for a given x, to $Stab(x) < MCG(F)$. Note that I am not insisting that S_x is a subgroup of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$MCG(F)$, or even of $Stab(x)$; it is not, since concatenations of Dehn twists do not generally yield a homeomorphism which is a Dehn twist about some circle.

Now consider triples of elements of P of the form $(x\,,\,b\,,\,p)$, where $b \in S_x$, and p is any old element in P, and x is a test element. According to the quandle relation Q3),

$$x\lceil\overline{b}\;\overline{p}\rceil\overline{b}\rceil = x\,\overline{b\overline{p}\rceil}\rceil \qquad ,$$

but for $b \in S_x$, since $\lceil\overline{b}$ has no effect on x, this relation turns into

$$x\,\overline{p}\rceil\overline{b}\rceil = x\,\overline{b\overline{p}\rceil}\rceil \qquad .$$

This is an expression of associativity of products of Dehn twists in P. So the general relation Q3) becomes <u>true</u> associativity when $b \in S_x$. In some sense, we may consider that for triple multiplications of the type shown on the left side of Q3), whether b lies in S_x or not is an obstruction to true associativity. The full formula in Q3) is an expression of "quasi-associativity", which I think of as being "local" with respect to a given element x. So for each x, there is a collection of truly associative multiplications associated (pardon the pun) to x, amongst all possible triple multiplications, namely those multiplications with first Dehn twist taken from S_x.

Now let $N = \{x \in P \,|\, x \,\text{is nonseparating on}\, F\}$. For elements $x \in N$, the S_x's and the attendant truly associative multiplication triples have an additional feature. They satisfy a type of cocycle condition. To see this, note that $MCG(F)$ acts transitively on the elements of N. Let $x\,,\,y\,,\,z \in N$ and let $g\,,\,h\,,\,k \in MCG(F)$ so that

$$\begin{aligned}(x)g &= y\\(y)h &= z\\(x)k &= z\end{aligned}$$

where the actions by elements of $MCG(F)$ are written on the right of elements of P. Define S_x, S_y, and S_z as above. Then for $a \in S_x$ we have

$$(y)g^{-1}\,\overline{a}\rceil g = y,$$

and

$$g^{-1}\,\overline{a}\rceil g = \overline{(a)g}\rceil,$$

where this last equation is just a reformulation of Q3) for a general homeomorphism $g \in MCG(F)$, where g is not necessarily a single Dehn twist, and $(a)g \in S_y$. Then we may define an action of $MCG(F)$ on the collection of S_x's by $(a, g) \to (a)g$, where $a \in S_x$ and $g \in MCG(F)$. Let us

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

denote this action by $(S_x)^g$. The action is 1-1 and onto since it is action via homeomorphism, carrying for example, S_x to S_y. More particularly,

$$(S_x)^{gh} = (S_x)^k.$$

Thus to a given $x \in N$ we may associate its unique sub-quandle, S_x, and the collection of triples (X, b, p), with $b \in S_x$, such that the multiplications of these triples (in the specified order, $x \lceil \overline{b \quad p} \rceil \, \overline{b} \rceil$ and variants) are associative. We also have the further structure of the cocycle condition, mentioned above. This agglomeration of structure suggests an analogy to the structure of a vector bundle over a manifold. There, cocycle conditions mediate interactions of locally trivial bundles over coordinate patches, glueing them together to yield the entire bundle, which is globally "quasi-trivial". Here, we have the cocycle condition relating the interactions of S_x's and the attendant collections of truly associative multiplications (for each $x \in N$), and glueing them to form a globally "quasi-associative" whole. So for each $x \in N$, S_x and the associative multplications play the same role in the global structure as the local patches and locally trivial bundles over them play in the vector bundle. The "size" of the collection of associative multiplications, for a given element $x \in N$, carries some (obvious) topological information; namely it is a measure of the genus of F, in that the greater the genus, the greater the number of circles on F which don't intersect x, so the greater the number of possible truly associative multiplications. The only exception is the case when $F = S^2$, where all circles are in S_x for a given circle x on F.

We may glean some further properties of the S_x's.

Proposition 6.1 *The assignment $[a] \to S_a$ is 1-1, where $[a]$ denotes the isotopy class of $a \in P$.*

Pf. There are a number of cases to consider. Let $a, b \in P$ such that $a \cap b = \emptyset$, and let $\text{genus}(F) = g$. Assume $[a], [b] \neq [\tau]$ where τ is a representative of the trivial class.

Case 1)

Suppose a is not homologous to b, and that both a and b are nonseparating imbedded circles in F. Since a and b are not homologous, they do not bound a subsurface of F, so $F - \{a, b\}$ is still path connected. Choose points $x, y \in F$, near a, lying on opposite sides of a, so that x, y are the endpoints of a short segment $\overline{xy}$, transverse and perpendicular to a. See Fig. 6.1). Since $F - \{a, b\}$ is path connected, there is an imbedded path $\sigma \subset F - \{a, b\}$ which joins the points x and y. Then $c = \sigma \cup \overline{xy}$ is a simple closed curve in $F - \{b\}$, such that $|c \cap a| = 1$, while $|c \cap b| = 0$. Thus, Dehn twisting around c fixes b but acts nontrivially on a. So $c \in S_b$ but $c \notin S_a$, so $S_a \neq S_b$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

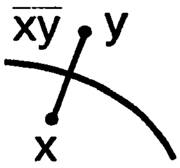


Fig. 6.1)

Case 2)

All other suppositions as above, now assume a is homologous to b in F. If a and b bound an annulus in F, then they are isotopic, and $[a] = [b] \Rightarrow S_a = S_b$. So now assume that a and b do not bound an annulus. Then they bound some subsurface $F' \subset F$ where genus$(F') < g$. By the classification of surfaces, (with boundary) all surfaces with the same genus and same number of boundary components are homeomorphic. Now $MCG(F)$ acts transitively on $N =$ {separating circles}, so there is a homeomorphism $f \in MCG(F)$ with $(a)f = a'$ and $(b)f = b'$, where these new circles bound a subsurface homeomorphic to F' in F, which is "generic" or standard. For this new generic picture, we now use the fact that genus$(F') < g$ to enable us to find a curve c' involving a hole "unused" by one of a' or b', so that e.g. $|c' \cap b'| = 0$. Then $\overline{(c')f^{-1}}\rceil = \overline{c}\rceil$ moves a but fixes b, so $c \in S_b$ but $c \notin S_a$, and $S_a \neq S_b$.

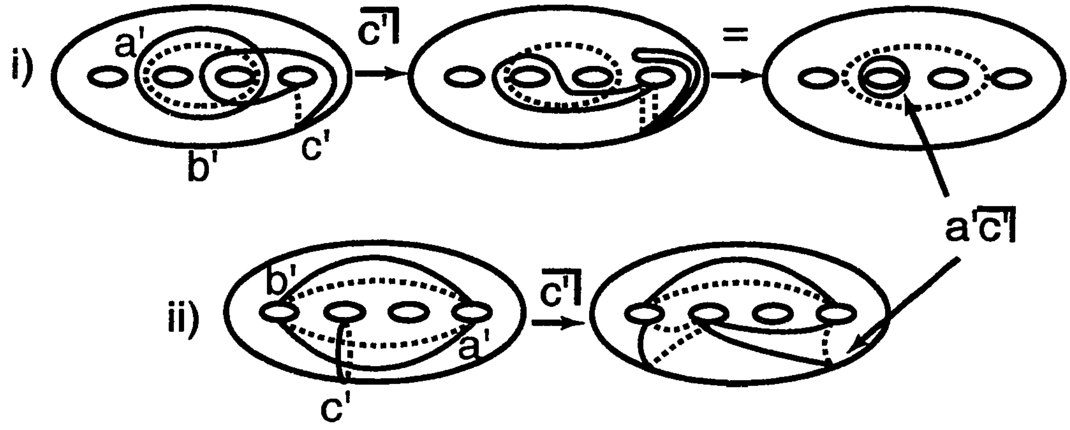


Fig. 6.2)

Case 3)

Now suppose $a \cap b = \emptyset$, and let a be a nontrivial separating curve on F, while b is nonseparating. Since $a \cap b = \emptyset$, and a is separating, b lies on one side of a, i.e. cutting along a yields b lyng in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

a component $F' \subset F$, where genus$(F') < g$. Again we may choose points x, $y \in F'$ and form the segment $\overline{xy}$, transverse and perpendicular to b, with x on one side of b, and y on the other. Since b does not separate F, it does not separate F', so $F' - b$ is path connected. As in case 1) we may again form the path σ and define the circle $c = \sigma \cup \overline{xy}$. We then proceed as in case 1).

In each of the three cases above,we found a circle that intersected one of a or b once, by construction.

Case 4)

Now suppose that both a and b are separating, and $a \cap b = \emptyset$. Cut F along a, say. One of the resulting components, F' contains b. If genus$(F') = 0$, then $[b] = [\tau]$, but this contradicts the assumption that $[b] \neq [\tau]$. On the other hand, if genus$(F') = 1$, then $[b] = [a]$, also contrary to assumption. Hence, genus$(F') \geq 2$, and b is a separating curve for F'. If we cut F' along b, we get two components of genus $<$ genus(F'). Each of these components has genus ≥ 1. Pick two nontrivial circles c_1 and c_2 in F', one from each of the components gotten by cutting along b. Form the connected sum $c = c_1 \#_\gamma c_2$ along a path γ which intersects b nontrivially. Then c crosses b (an even number of times) and thus Dehn twists about c move b but fix a, so $c \in S_b$ but $c \notin S_a$, and $S_a \neq S_b$. ///

By the previous proposition, since the assignment $[a] \rightarrow S_a$ is 1-1, the operations $\#_\gamma$, $\urcorner$, and $\ulcorner$ induce operations on the S_x's, so that e.g. $\overline{(S_x)S_y}| = S_{\overline{x\,y}|}$, and $S_x \#_\gamma S_y = S_{x \#_\gamma y}$, wherever $x \#_\gamma y$ is defined. I do not know what meaning this has, if any. It would be nice, as well, to have some idea of whether the structure analogous to the vector bundle, that was mentioned above, can be used to any topological benefit, with regards to 3-manifolds and their Heegaard splittings.

In the previous sections, I attempted to show some of the ways in which the algebraic structure of the extended quandle may be used to look at various aspects of the topology of 3-manifolds, specifically regarding representations via Heegaard splittings. Now allow me to indulge in some questions and speculations as to further properties and avenues of study.

As I mentioned briefly at the end of section 2, it would be nice to know whether the connected sum admits a well defined, canonical or unique (possibly shortest) path, and whether this would allow us to create connected sums using any applicable paths, yet specialize down to to a unique, well defined multiplication and an actual Hopf algebra structure. Again, the desire for this is motivated by some notion of completeness, and the comfort of finding somewhat more familiar structures in a blooming algebraic wilderness. Whether such motivation is justified is also a question.

There are other questions which arise somewhat naturally from the consideration of the extended quandle and comparison to well known structures. Consider the homology group $H_1(F)$ for F a surface of genus g. For a pair of homology classes $\{a\}$ and $\{b\}$ with representatives which can be made to intersect in 0 or 1 point, the sum operation in homology is accomplished either by connected sum, in the first case, or by one of two Dehn twists, in the second. E.g. if $|a \cap b| = 1$ then $\{a\} + \{b\} = \{a\overline{b}\rceil\}$ or $\{a\lceil\overline{b}\}$, depending on the signs. See Fig. 6.3)

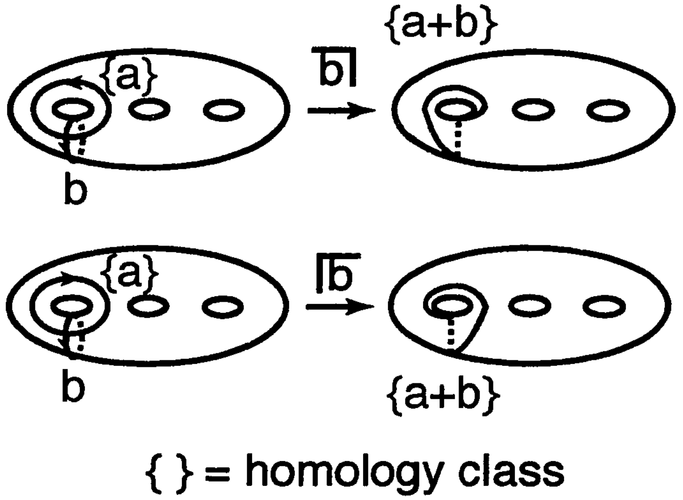


Fig. 6.3)

So to this extent, the operations in the extended quandle generalize the sum in homology. The question then arises: if connected sum and Dehn twists are manifestations of the same underlying operation, sum, in homology, are they also manifestations of some analogous single operation in the extended quandle? To a certain extent, this would be a pleasant development, if it were the case. For as it now stands, the connected sum is defined only for nonintersecting pairs of circles, so it is not, strictly speaking, an algebraic operation. The Dehn twist is of course defined in all cases. If it were true that they were manifestations of some underlying and as yet unspecified operation, we might be able to reasonably make connected sum into fully algebraic operation, adding a sort of completeness (not in the mathematical sense) to the structure of the extended quandle.

One possibility for realizing this might be to find a way to think of connected sum, defined by a path γ joining the two circles, as degenerating into the Dehn twist of two intersecting circles (intersecting once) as length(γ) $\to 0$. In other words, considering the intersection point of a pair of circles a and b with $|a \cap b| = 1$, as a path between the circles of length 0. However, I have not yet figured out a way to correctly and coherently formalize this. A related complementary

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

question might be: is it possible or useful to create a notion of connected sum between pairs of circles which intersect?

This actually leads to a rather more profound question of whether in formulating the operations of the extended quandle, and in examining some of their interactions, it would have been more correct and/or more useful to have defined them for immersed, rather than merely imbedded circles in F? It is not unreasonable to expect that the addition of the further information of intersection and self-intersection on the circles, into the "constitution" of the extended quandle might be a useful generalization, and that the extra structure might enhance the "coherence" (not in the categorical sense) and usefulness of the quandle. There are algebraic structures defined for closed curves on surfaces (see [Turaev]) based on this. Alternatively, it might be more fruitful to consider paths or arcs in F as the basic building blocks in the definitions, rather than circles. How might this be accomplished? Could this be used to eliminate the non-algebraicness of the connected sum, for instance?

Another avenue of pursuit which I looked into briefly, without terribly much success, was whether it might be possible to find some type of "function" associated to F and the extended quandle, which might give a simple, closed form expression for the Dehn twist about a given connected sum of two circles, in terms of "generating" Dehn twists. In other words, given a, $b \in P$, with $|a \cap b| = 0$, and given a path γ between them, is there a relatively simple function of a, b, and γ which determines the Dehn twists $\overline{|(a \#_\gamma b)}$ and $\overline{(a \#_\gamma b)|}$? Such a development would provide a host of new relations in $MCG(F)$, or perhaps they might be better termed "meta-relations", since the notion of connected sum is not native to $MCG(F)$, though it is to the extended quandle. In some sense, such relations would "factor through" the extended quandle. Indeed, there already exists methods for describing Dehn twists about connected sums, or any circle, for that matter. If, for instance, the circle $a \#_\gamma b$ is nonseparating in F, since $MCG(F)$ acts transitively on $N = \{x \in P \mid x \text{ is nonseparating in } F\}$, one may choose a circle $c \in N$, and find a homeomorphism $\phi \in MCG(F)$ such that $(c)\phi = a \#_\gamma b$ where ϕ is expressed as a product of Dehn twists. Then the quandle relation Q3) allows the twist about $a \#_\gamma b$ to be expressed as a conjugate by ϕ , of the corresponding twist about c. This method of yielding an expression for a twist about $a \#_\gamma b$ depends on the choice of the circle c. In particular, the length (number of twists about circles) and amount of complication of the resulting expression, also depend on the choice of c. There are many such c's from which to choose. Sticking with this particular point of inquiry, one might ask: for a given $a \#_\gamma b$, what choice of c yields a conjugation expression, (i.e. a word in $MCG(F)$) for the twist about $a \#_\gamma b$, which is least complicated? Here, complicated might mean 'is of minimal length' or perhaps 'is made up of Dehn twists about circles which are themselves,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

uncomplicated', homotopically for instance. Again the choices for the element c provide a large measure of ambiguity to this formulation. This was precisely the reason for wanting to have a "function" only dependent on a, b, and γ.

Another idea which crops up rather naturally, but which I have not used explicitly in the discussions above, is the notion of a duality in P, dovetailing neatly with the idea of P as an "algebra" for Heegaard splittings. More explicitly, take P as the free module over some commutative ring K, generated by the circles on F, as was done in section 2. If we have a Heegaard splitting given by (H_1, H_2, F), it is reasonable to consider the boundary circles of a complete disk system $\mathcal{D}_1$ of H_1 and those of a complete disk system $\mathcal{D}_2$ of H_2, (these latter being the characteristic circles of the splitting) as being dual to one another. One impetus for doing so is that when we consider a handle decomposition of M compatible with the splitting, the disks in $\mathcal{D}_1$ are exactly the co-cores of the 1-handles, while those in $\mathcal{D}_2$ are cores of the 2-handles. So $\partial\mathcal{D}_1$ and $\partial\mathcal{D}_2$ may be thought of as dual bases of submodules associated specifically with the handlebodies of a given Heegaard splitting. Changes to the respective complete disk systems, for instance by use of waves (as in section 5) or by application of some homeomorphism preserving a handlebody, may be thought of as changes of basis. The homology intersection form $< a, b >$ for $H_1(F)$, where $a \in \partial\mathcal{D}_1$ and $b \in \partial\mathcal{D}_2$ provides a method of evaluation; $< , >: P^* \otimes P \to K$. Since by definition, the $\mathcal{D}_i$'s (i=1,2) are disks with nonseparating boundary in F, and since again $MCG(F)$ acts transitively on N, there are explicit endomorphisms of P_K arising from the homeomorphisms $\phi \in MCG(F)$ with $(\partial\mathcal{D}_2)\phi = \partial\mathcal{D}_1$, which carry the dual basis to the basis, and vice versa. Perhaps for a more complete (and conceivably useful) notion of the extended quandle, such a natural construction ought to be incorporated into the structure.

One final project that I would like to realize is the use of the algebraic structure to form invariants of 3-manifolds. Up until now, I had not looked into this. My hope would be that since the algebra here is tied to the topology in a relatively straightforward fashion, that thus, any invariants derived from it would be easily seen to be descriptive of particular aspects of the topology and the geometry. This might distinguish such invariants from many of the present ones, quantum and Vassiliev, for which the connection to the topology is not immediately evident. One alternative thought is that perhaps some of the present invariants, in particular the quantum invariants, might turn out to be invariants of the quandle formed by the action of $MCG(F)$ on $\pi_1(M)$.

I have tried to show some of the interplay between algebraic and topological structures related to curves on surfaces, and to point out some possible directions for further inquiry. I feel that placing some of the topological ideas within an algebraic framework lends a dynamic quality to the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

geometric topology, which I appreciated; in particular, the pushing around of collections of disks by elements of $MCG(F)$. Also, I have attempted to suggest possible additions and ameliorations to the structures which might prove useful and/or aesthetic. I hope I have been somewhat successful. The area which I have briefly explored seems quite rich in structure and possibility.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

7 References

[Birman] Birman, J., *On the Equivalence of Heegaard Splittings of Closed Orientable 3-Manifolds*, Annals of Math. Studies, 84, N.J. (1975) 137-164

[Casson, Gordon] Casson, A., Gordon, C., *Reducing Heegard Splittings*, Top. and Appl. 27, (1987) 275-283

[Hatcher, Thurston] Hatcher, A., Thurston, W., *A Presentation of the Mapping Class Group of a Closed, Orientable Surface*, Topology 19, (1980) 221-237

[Homma, Ochiai, Takahashi] Homma, T., Ochiai, H., Takahashi, M., *An Algorithm for Recognizing S^3 in 3-Manifolds with Heegaard Splittings of Genus Two*, Osaka J. of Math., 17, (1980) 625-648

[Kauffman] Kauffman, L., *Knots and Physics*, World Scientific (1991)

[Morikawa] Morikawa, O., *A Counterexample to a Conjecture of Whitehead*, Math. Sem. Notes, Kobe Univ. vol. 8, (1980) 295-298

[Ochiai] Ochiai, M.,*A Counterexample to a Conjecture of Whitehead and Volodin- Kuznetzov-Fomenko* J. Math Soc. of Japan, vol. 31, no. 4, (1979) 287-291

[Scharlemann] Scharlemann, M., *Handbook of Geometric Topology*, Preprint, (1996)

[Scharlemann, Thompson] Scharlemann, M., Thompson, A., *Thin Position and Heegaard Splittings of the 3-Sphere*, J. Diff. Geom. 39, (1994) 343-357

[Starr] Starr, E., *Curves in Handlebodies*, Ph.D Thesis, Berkeley, (1992)

[Turaev] Turaev, V., *Algebras of Loops and Surfaces; Algebras of Knots, and Quantization*, Braid Group, Knot Theory, and Statistical Mechanics, Ed. Yang and Ge, World Scientific Publishing, Advanced Series in Mathematical Physics 9, (1989) 59-96

[Viro] Viro, O., *The Volodin-Kuznetzov-Fomenko Conjecture on Heegaard Diagrams is False*, Uspehi Math. Nauk. 32,(1977) 175-176

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[Volodin, Kuznetzov, Fomenko] Volodin, I., Kuznetzov, V., Fomenko, A., *The Problem of Discriminating Algorithmically the Standard Three Dimensional Sphere*, Uspekhi Math. Nauk. 29:5, (1974) 72-168

[Wajnryb] Wajynryb, B., *Mapping Class Group of a Handlebody*, Preprint (1997)

[Waldhausen] Waldhausen, F., *Heegaard-Zerlegungen de 3-Sphäre*, Toplogy 7, (1968) 195-203

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.